

Aus dem Institut für  
Transfusionsmedizin und Immunhämatologie  
des Blutspendedienstes Hessen des DRK, Frankfurt am Main  
(Ärztlicher Direktor: Prof. Dr. med. E. Seifried)

**Populationsgenetische Untersuchungen der  
Tetranukleotid Tandem Repeat Polymorphismen  
D11S488, D18S51, D19S246,  
HUMFIBRA und HUMVWFA31/A**

INAUGURAL-DISSERTATION  
zur Erlangung des Doktorgrades der Medizin  
des Fachbereichs Humanmedizin der  
Johann Wolfgang Goethe-Universität Frankfurt am Main

vorgelegt von  
**Sandra Müller**  
aus Dieburg

Frankfurt am Main, 2000

Dekan: Prof. Dr. med. G. von Jagow

Referent: Prof. Dr. med. E. Seifried

1. Korreferent: Prof. Dr. med. U. Langenbeck

2. Korreferent: Prof. Dr. med. H.J. Bratzke

Tag der mündlichen Prüfung: 19. März 2001

**für Edith und Miriam**



# Inhaltsverzeichnis

## Abbildungsverzeichnis

## Tabellenverzeichnis

## Abkürzungsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Bedeutung von genetischen Markern in der Abstammungsbegutachtung	1
1.1.1	Konventionelle Blutgruppen und Serumenzym polymorphismen	2
1.1.2	Molekulargenetische Marker	4
1.1.2.1	Variable Number of Tandem Repeats	4
1.1.2.2	Short Tandem Repeats	5
1.2	Untersuchungsgegenstand und Ziel der Arbeit	7
<b>2</b>	<b>Material und Methoden</b>	<b>9</b>
2.1	Geräte	9
2.2	Populationsstichprobe	10
2.3	Untersuchungsmaterial	10
2.3.1	DNA-Isolierung	11
2.3.1.1	Reagentien	12
2.4	Short Tandem Repeat Loci	13
2.4.1	Chromosomenlokalisierung	13
2.4.2	Primersequenzen und Farbstoffmarkierung	15
2.4.3	Berechnung des Primerschmelzpunktes	16
2.5	Polymerasekettenreaktion	16
2.5.1	Prinzip der PCR	16
2.5.2	Durchführung der PCR	18
2.5.2.1	Geräte	18
2.5.2.2	Bedingungen	19
2.5.3	Nachweis der amplifizierten DNA	20
2.5.3.1	Prinzip	20
2.5.3.2	Durchführung	21
2.5.3.3	Reagentien	22
2.6	Detektion von STR-Markern	23
2.6.1	Prinzip der Fluoreszenzmarkierung und automatischen Detektion	23
2.6.2	Durchführung	25
2.6.3	Reagentien	28
2.6.4	Auswertung der erhobenen Daten	29
2.7	Sequenzierung von STR-Allelen	30
2.7.1	Prinzip der Sanger Sequenzierung	30
2.7.2	Durchführung der Sequenzierung mit Sequenase	31
2.7.3	Taq Cycle Sequencing	32
2.7.4	Gelextraktion	34
2.7.5	Auswertung der Sequenzen	35
2.7.6	Allelbezeichnung	36
2.8	Statistische Auswertung	37

<b>3</b>	<b>Ergebnisse</b>	<b>39</b>
3.1	Vorversuche zur Optimierung der PCR-Bedingungen	39
3.2	Etablierung eines Multiplex Amplifikationssystems	42
3.3	Populationsgenetische Untersuchungen	45
3.3.1	HUMD11S488	45
3.3.1.1	Allelfrequenzen und Genotypen	46
3.3.1.2	Sequenzstruktur der Allele	48
3.3.1.3	Neue Allele	49
3.3.1.4	Varianten	50
3.3.2	HUMD18S51	50
3.3.2.1	Allelfrequenzen und Genotypen	52
3.3.2.2	Sequenzstruktur der Allele	54
3.3.2.3	Neue Allele	55
3.3.2.4	Varianten	55
3.3.3	HUMD19S246	55
3.3.3.1	Allelfrequenzen und Genotypen	56
3.3.3.2	Sequenzstruktur der Allele	58
3.3.3.3	Neue Allele	59
3.3.3.4	Varianten	59
3.3.4	HUMFIBRA	60
3.3.4.1	Allelfrequenzen und Genotypen	60
3.3.4.2	Sequenzstruktur der Allele	62
3.3.4.3	Neue Allele	63
3.3.4.4	Varianten	64
3.3.5	HUMVWFA31/A	64
3.3.5.1	Allelfrequenzen und Genotypen	64
3.3.5.2	Sequenzstruktur der Allele	66
3.3.5.3	Neue Allele	67
<b>4</b>	<b>Diskussion</b>	<b>69</b>
	<b>Zusammenfassung</b>	<b>79</b>
	<b>Publikationsnachweis</b>	<b>81</b>
	<b>Literaturverzeichnis</b>	<b>83</b>
	<b>Anhang</b>	<b>91</b>
	<b>Danksagung</b>	<b>93</b>
	<b>Lebenslauf</b>	<b>95</b>
	<b>Ehrenwörtliche Erklärung</b>	<b>97</b>

# Abbildungsverzeichnis

Abb. 2.1:	Lokalisation der Marker auf ihren Chromosomen	14
Abb. 2.2:	Schematische Darstellung der PCR	17
Abb. 2.3a:	Schematische Darstellung des Strahlengangs von Laser- und Fluoreszenzlicht im 373A DNA-Sequencer (Blick von oben)	24
Abb. 2.3b:	Gel im Blick von vorn	24
Abb. 2.4:	Darstellung des internen Standards (Version 2) als Elektropherogramm	28
Abb. 3.1:	Ausschnitt aus der Versuchsreihe 3 für D19S246	40
Abb. 3.2:	Darstellung von Amplifikationsprodukten der fünf Marker unter drei Bedingungen im Agarosegel	41
Abb. 3.3:	Darstellung des Triplex als Elektropherogramm	44
Abb. 3.4:	Darstellung des Quadruplex als Elektropherogramm	44
Abb. 3.5:	Darstellung der DNA-Fragmente von D11S488 als Elektropherogramm	45
Abb. 3.6:	Verteilung der Allelfrequenzen von D11S488	46
Abb. 3.7:	D11S488 mit seinen variablen Regionen VR I und VR II	49
Abb. 3.8:	Darstellung der DNA-Fragmente von D18S51 als Elektropherogramm	50
Abb. 3.9:	Verteilung der Allelfrequenzen von D18S51	52
Abb. 3.10:	D18S51 mit seiner variablen Region VR	54
Abb. 3.11:	Darstellung der DNA-Fragmente von D19S246 als Elektropherogramm	55
Abb. 3.12:	Verteilung der Allelfrequenzen von D19S246	56
Abb. 3.13:	D19S246 mit seinen variablen Regionen VR I und VR II	58
Abb. 3.14:	Darstellung der DNA-Fragmente von HUMFIBRA als Elektropherogramm	60
Abb. 3.15:	Verteilung der Allelfrequenzen von HUMFIBRA	61
Abb. 3.16:	HUMFIBRA mit seiner variablen Region VR	63
Abb. 3.17:	Darstellung der DNA-Fragmente von HUMVWFA31/A als Elektropherogramm	64
Abb. 3.18:	Verteilung der Allelfrequenzen von HUMVWFA31/A	65





# Tabellenverzeichnis

Tabelle 2.1:	Liste der verwendeten Primer	15
Tabelle 2.2:	Liste der Fluoreszenzfarbstoffe	16
Tabelle 2.3:	PCR-Bedingungen	20
Tabelle 2.4:	123 bp DNA-Leiter	23
Tabelle 2.5:	Interner Standard	27
Tabelle 2.6:	PCR-Bedingung für Taq Cycle Sequencing	33
Tabelle 3.1:	Amplifikationscharakteristika der untersuchten Systeme	43
Tabelle 3.2:	Allel- und Genotypfrequenzen von D11S488	47
Tabelle 3.3:	Chi-Quadrattest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D11S488	48
Tabelle 3.4:	weitere statistische Parameter für D11S488	48
Tabelle 3.5:	Allele mit DNA-Sequenzen der Repeatregion von D11S488	51
Tabelle 3.6:	Allel- und Genotypfrequenzen von D18S51	53
Tabelle 3.7:	Chi-Quadrattest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D18S51	53
Tabelle 3.8:	weitere statistische Parameter für D18S51	54
Tabelle 3.9:	Allel- und Genotypfrequenzen von D19S246	57
Tabelle 3.10:	Chi-Quadrattest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D19S246	57
Tabelle 3.11:	weitere statistische Parameter für D19S246	57
Tabelle 3.12:	Allele mit DNA-Sequenzen der Repeatregion von D19S246	59
Tabelle 3.13:	Allel- und Genotypfrequenzen von HUMFIBRA	61
Tabelle 3.14:	Chi-Quadrattest zur Abweichung vom Hardy-Weinberg Gleichgewicht für HUMFIBRA	62
Tabelle 3.15:	weitere statistische Parameter für HUMFIBRA	62
Tabelle 3.16:	Allel- und Genotypfrequenzen von HUMVWFA31/A	65
Tabelle 3.17:	Chi-Quadrattest zur Abweichung vom Hardy-Weinberg Gleichgewicht für HUMVWFA31/A	66
Tabelle 3.18:	weitere statistische Parameter für HUMVWFA31/A	66
Tabelle 4.1:	pM-Werte für die fünf individuellen Loci und pM-Werte für Triplex, Quadruplex und möglichen Pentaplex	75



# Abkürzungsverzeichnis

A	Adenin
AmpFLP's	Amplifikations-Fragment-Längen-Polymorphismen ( <i>amplifiable fragment length polymorphisms</i> )
bp	Basenpaare
C	Cytosin
CEPH	Centre d'Etude du Polymorphisme Humain
ddNTP	Didesoxynukleosidtriphosphat
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleosidtriphosphat
EDTA	Ethylen-Diamin-Tetraessigsäure
FAM	5-carboxy-fluorescein
G	Guanin
HEX	4,7,2',4',5',7'-hexachloro-6-carboxy-flourescein
PCR	Polymerasekettenreaktion ( <i>polymerase chain reaction</i> )
PMT	Photomultiplier ( <i>photo multiplier tube</i> )
ROX	6-carboxy-X-rhodamin
RT	Raumtemperatur
STR	kurze Tandem-Wiederholung ( <i>short tandem repeat</i> )
T	Thymin
TAMRA	N,N,N',N'-tetramethyl-6-carboxyrhodamin
<i>T<sub>m</sub></i>	Schmelztemperatur
UV	Ultraviolett
VNTRs	variable Anzahl von Tandem-Wiederholungen ( <i>variable number of tandem repeats</i> )
vWF	von Willebrand Faktor



# 1 Einleitung

Während der letzten Jahre sind auf dem Gebiet der Chromosomenanalyse zahlreiche DNA-Marker in das Zentrum wissenschaftlicher Studien gerückt. Aufgrund von Längenvariationen, die durch einfache immer wiederkehrende Sequenzen verursacht werden, weisen genetische Marker teilweise einen erheblichen Polymorphismus auf [Weber et al. 1991]. Dabei lassen sich zwei Gruppen von DNA-Markern unterscheiden. Zum einen die VNTRs (*variable number of tandem repeats*), die vornehmlich in der Telomerregion der Chromosomen vorkommen und zum anderen die STRs (*short tandem repeats*), welche nahezu überall im menschlichen Genom verteilt und im Durchschnitt alle sechs bis zehn Kilobasen auftreten [Beckmann u. Weber 1992]. STRs lassen sich leicht mit Hilfe der Polymerasekettenreaktion (PCR) amplifizieren und finden Anwendung in der Genkartenerstellung [Edwards et al. 1991, Stallings et al. 1991, Tomfohrde et al. 1992, Beckmann u. Weber 1992] und der pränatalen Diagnostik vieler genetischer Erbkrankheiten, wie z.B. der Zystischen Fibrose [Morall u. Estivill 1995] oder der Muskeldystrophie Duchenne/Becker [Oudet et al. 1991]. Mit der Funktion eines „genetischen Fingerabdrucks“ werden STR-Marker in der forensischen Medizin zur Identifizierung von Individuen eingesetzt und können daher auch für die Abstammungsbegutachtung von großer Bedeutung sein.

## 1.1 Bedeutung von genetischen Markern in der Abstammungsbegutachtung

Zur Feststellung des biologischen Vaters eines Kindes bedient sich das Gericht im Streitfall abstammungsgenetischer Gutachten. Dazu gehören das Blutgruppengutachten, das Anthropologisch-erbbiologische Gutachten und die Begutachtung des Chromosomenpolymorphismus. Wichtigstes Hilfsmittel bei der Klärung strittiger Abstammung sind derzeit die Blutgruppengutachten. Mit ihrer Hilfe läßt sich die Vaterschaft eines Mannes mit einer bestimmten Wahrscheinlichkeit feststellen oder ausschließen.

Seltener werden Anthropologisch-erbbiologische Gutachten erstellt. Es handelt sich um einen Ähnlichkeitsvergleich zwischen Kind und möglichem Erzeuger, wobei die gesamte

körperliche Erscheinung der Menschen berücksichtigt wird. Das Mindestalter für eine solche Untersuchung beträgt allerdings drei Jahre.

Erst seit kurzer Zeit gewinnen DNA-Analysen zur Identifizierung des biologischen Vaters zunehmend an Bedeutung. Diese Untersuchungen werden vor allem bei Blutsverwandtschaft durchgeführt und beruhen auf Polymorphismen genetischer Marker. Eine Anwendung des Verfahrens kann schon vor der Geburt des Kindes durch Amniozentese erfolgen. Chromosomenanalysen wurden bisher meist nur als zusätzliche Gutachten gefordert. Seit populationsgenetische Studien zu zahlreichen Markern mit Bestimmung von Allelverteilung und Sequenzstruktur durchgeführt werden und sich einfache Verfahrenstechniken zur genauen Genotypenermittlung etablieren, kommt diese Art von Gutachten immer mehr zum Einsatz.

### **1.1.1 Konventionelle Blutgruppen und Serumenzym polymorphismen**

Blutgruppeneigenschaften können an Erythrozyten, Leukozyten, Thrombozyten, Serum-eiweißen und Enzymen nachgewiesen werden. Voraussetzung für die forensische Verwertbarkeit ist die Reproduzierbarkeit dieser Nachweise.

Daher müssen Fehlerquellen, wie beispielsweise Verunreinigungen von Testseren und Reagentien, längeres Liegen der Blutprobe, ungenügende Identitätsprüfung und das fehlende Mitführen von Kontrollblutproben mit bekannten Eigenschaften, dringend vermieden werden.

Wichtig ist eine Konstanz der Merkmale. Während Blutgruppenmerkmale zeitlebens konstant bleiben, können Merkmale des Hp- oder Gm-Systems vorübergehend oder dauerhaft fehlen. Zuverlässige Blutgruppenuntersuchungen sind allerdings erst bei Kindern jenseits des achten Lebensmonats möglich. Nach Bluttransfusionen muß wegen Verfälschung der Ergebnisse mit einer solchen Untersuchung drei Monate gewartet werden.

In Abstammungsfragen dürfen nur Eigenschaften des Blutes verwertet werden, welche übersichtlich vererbt werden und deren Erbgang gesichert ist. Je größer die Anzahl der Allele, desto höher ist der Informationsgehalt eines Blutgruppensystems im Rahmen der Vaterschaftsbegutachtung.

#### **Erythrozytensysteme**

Zu den Erythrozytensystemen gehören das ABO-System, das MNSs-System, das P-System, das Kell-System, das Duffy-(Fy-)System, das Kidd-System und das CDE-(Rhesus-)System. Am

längsten bekannt sind die um die Jahrhundertwende von dem Nobelpreisträger Karl Landsteiner entdeckten Gruppen des *ABO-Systems*. Zur Merkmalsbestimmung werden die Erythrozyten mit Testseren (Agglutinogene) und das Blutserum mit Test-Erythrozyten (Gegenprobe, Agglutinine) untersucht. Das Verfahren kann lediglich den Phänotyp, nicht aber den Genotyp (Untergruppen) ermitteln. Beim *MNSs-System* sind die Merkmale M und N genetisch mit den Erbmerkmalen S und s gekoppelt. Der Nachweis erfolgt mit menschlichen Antiseren. Von allen bekannten Blutgruppensystemen hat das MNSs-System die höchste Ausschlußchance bei Nichtvaterschaft (ca. 32 %). Das Merkmal Kell (K) des *Kell-Systems* wird bei etwa 7 % der Menschen in Europa gefunden. Es ist ein dominantes Erbmerkmal mit einem kodominanten Allel-Gen (k). Der Nachweis erfolgt im Coombs-Test. Im *Duffy-System* finden sich zwei kodominant vererbte Merkmale, Fy (a) und Fy (b). Sie werden ebenfalls mit dem Coombs-Test nachgewiesen. Das *CDE-(Rhesus-)System* besteht aus den Allelpaaaren D und d, C und c und E und e, die in Dreierkomplexen getrennt weitervererbt werden. Die Nomenklatur Rh-positiv und rh-negativ richtet sich nach dem Vorhandensein bzw. Fehlen des Merkmals D. Für den Nachweis der Rhesus-Merkmale wird der Coombs-Test angewendet.

### **Serumsysteme**

Zu den wichtigen vererbten Eigenschaften des Serums gehören: *Haptoglobin-(Hp-)System* (2 bzw. 5 häufige Allele je nach Methodik), *Gc* Gruppenspezifische Komponente (2 bzw. 3 häufige Allele, je nach Methodik), *C3* (2 häufige Allele), *Tf* Transferrin (3 häufige Allele), *Pi* Alpha<sub>1</sub>-Antitrypsin (5 häufige Allele), *PLG* Plasminogen (2 häufige Allele), *F XIII a* (2 häufige Allele), *F XIII b* (3 häufige Allele), *Bf* (2 häufige Allele), *AHS* Alpha<sub>2</sub>HS-Glykoprotein (2 häufige Allele).

### **Enzymsysteme**

Zusätzlich zu den genannten Gruppen kommt den vererbten Enzymgruppen der roten Blutkörperchen eine wichtige Rolle in der Vaterschaftsbegutachtung zu. Als Systeme werden verwendet: saure Erythrozytenphosphatase (*ACP* acid phosphatase), Phosphoglukomutase<sub>1</sub> (*PGM<sub>1</sub>*), Adenylatkinase (*AK*), Adenosindesaminase (*ADA*), 6-Phosphoglukonat-Dehydrogenase (*6-PGD*), Glutamat-Pyruvat-Transaminase (*GPT*), Glyoxalase (*GLO*), Esterase D (*EsD*) und andere.

Durch die Art der Vererbung und den in einer Bevölkerung vorkommenden Polymorphismus der genannten Merkmale erreichen die einzelnen Systeme Vaterschaftsausschlußchancen

zwischen 10 % und 30 %, welche in Kombination Werte zwischen 91 % und 96 % annehmen können. Dabei ist die Aussagekraft um so höher, je seltener ein Merkmal ist.

## 1.1.2 Molekulargenetische Marker

Um eine Vaterschaft mit an Sicherheit grenzender Wahrscheinlichkeit nachweisen zu können, ist entsprechend den Richtlinien für die Erstattung von Abstammungsgutachten eine kombinierte Allgemeine Vaterschafts-Ausschluß-Chance (AVACH) von 99,99 % notwendig [Robert Koch-Institut 1996]. Solch ein Wert kann, wenn über die konventionellen Gutachten noch kein Ausschluß nachgewiesen werden konnte, nur mit zusätzlichen Systemen erreicht werden, wie sie bei den molekulargenetischen Markern zu finden sind. In Zukunft wird die DNA-Analyse wachsende Bedeutung erlangen. Die in allen Zellkernen enthaltene Desoxyribonukleinsäure (DNA) läßt sich aus kernhaltigen Blutzellen gewinnen. Unter Verwendung locuspezifischer Primer können Fragmente molekulargenetischer Marker, die von Mensch zu Mensch meist unterschiedlich sind, über die Polymerasekettenreaktion (PCR) vervielfältigt und elektrophoretisch aufgetrennt werden. Damit werden Erbmerkmale sichtbar gemacht, die bei Mutter, Kind und Vater mit bisherigen serologischen Verfahren nicht zu erfassen sind. Mit der DNA-Analyse sollen Fälle fraglicher Vaterschaft gelöst werden, deren Klärung heute noch Schwierigkeiten bereitet. Von besonderer Bedeutung sind zwei Klassen genetischer Marker, die VNTRs (*variable number of tandem repeats*) und die STRs (*short tandem repeats*), auf die im weiteren näher eingegangen wird. Diese DNA-Marker weisen im Gegensatz zu den serologischen Systemen einen beachtlichen Polymorphismus auf, der durch Variationen in der Allellänge und -sequenz hervorgerufen wird.

### 1.1.2.1 Variable Number of Tandem Repeats

VNTRs (*variable number of tandem repeats*) sind Genloci mit kurzen sich in unterschiedlicher Anzahl wiederholenden Sequenzmotiven. Wegen ihrer hohen Heterozygotenraten und großen Allelzahlen kommen VNTRs als genetische Marker in der biologischen und biomedizinischen Forschung zur Anwendung [Jeffreys et al. 1988, Burke et al. 1991, Pena et al. 1993]. Basierend auf der Größe der Repeateinheit lassen sich VNTRs in drei Gruppen



einteilen: Mikrosatelliten (1-2 bp), Short Tandem Repeats (STRs; 2-6 bp) und Minisatelliten (>10 bp) [Shriver et al. 1995]. Weiterhin können VNTRs nach ihrer Allellänge klassifiziert werden. Zunächst wurden vor allem Allele der langen VNTRs, die bevorzugt in der Telomerregion menschlicher Chromosomen vorkommen [Beckmann u. Weber 1992, Kimpton et al. 1993, Newton u. Graham 1994], über den Restriktionsfragment-Längenpolymorphismus (RFLP) bestimmt. In den letzten Jahren hat sich jedoch auch die Untersuchung kürzerer VNTRs mittels Polymerasekettenreaktion (PCR) zunehmend durchgesetzt [Boerwinkle et al. 1989, Weber u. May 1989]. Von Vorteil ist hierbei, daß bei den Mikrosatelliten bzw. STRs eine präzisere Allelbestimmung als bei den längeren VNTRs möglich ist [Gill et al. 1990, Evett u. Gill 1991].

### **1.1.2.2 Short Tandem Repeats**

Short Tandem Repeat (STR) Loci [Edwards et al. 1991, Edwards et al. 1992, Polymeropoulos et al. 1990] sind eine Untergruppe der VNTRs [Budowle et al. 1996] und bestehen aus zwei bis sechs Basenpaar langen Sequenzmotiven, die in unterschiedlicher Anzahl wiederholt werden [Stallings et al. 1991, Tomfohrde et al. 1992, Beckmann u. Weber 1992, Engelstein et al. 1993, Fregeau u. Founey 1993]. Ihre Verteilung reicht nahezu über das gesamte menschliche Genom, in dem sie als Di-, Tri- oder Tetranukleotidrepeats etwa alle 20 kb (Kilobasen) vorkommen [Beckmann u. Weber 1992]. Über die Hälfte der STR-Loci, die Edwards et al. [1992] untersucht hat, weisen einen Polymorphismus in der Anzahl ihrer Repeats auf, was sie zu einer wertvollen Quelle von genetischen Markern macht [Sprecher et al. 1996]. Polymorphe STR-Sequenzen sind sowohl in genetischen als auch extragenetischen Regionen zu finden. Innerhalb genetischer Regionen treten sie in Introns, flankierenden Sequenzen und kodierenden Abschnitten auf [Edwards et al. 1991].

Aufgrund ihrer geringen Allellängen (in der Regel 130-300 bp) [Brinkmann 1992, Kimpton et al. 1992, Wiegand et al. 1993, Möller et al. 1994a] sind STRs ideal zur Amplifikation mittels Polymerasekettenreaktion (PCR) unter Verwendung locuspezifischer Primer geeignet. Selbst der spezifische Nachweis eines individuellen Allel-Profiles durch Analyse alter oder schlecht gelagerter Proben, welche nur degradierte genomische DNA enthalten, ist möglich [Hagelberg et al. 1991, Gill et al. 1992, Jeffreys et al. 1992, Kimpton et al. 1993].

Die Analyse von Short Tandem Repeat (STR) Loci mittels PCR hat gezeigt, daß diese Marker aufgrund ihres hohen Repeatpolymorphismus sehr informativ für eine Vielzahl von Anwendungsbereichen sind [Edwards et al. 1992]. Wegen ihres häufigen Auftretens innerhalb des Genoms werden Dinukleotidrepeat Loci zur Erstellung von Genkarten genutzt [Fornage et al. 1992, Lareu et al. 1996]. Derzeit werden Studien durchgeführt, die Tetranukleotide für den Gebrauch in der forensischen Medizin und der abstammungsgenetischen Begutachtung untersuchen [Urquhart et al. 1995, Fregeau u. Fournay 1993].

Die PCR-Amplifikation der STR-Systeme erfolgt unter Verwendung fluoreszenzmarkierter Primer. Anschließend werden die Produkte mittels Polyacrylamidgelelektrophorese in einem automatischen DNA-Sequenzer aufgetrennt [Ziegle et al. 1992, Kimpton et al. 1993, Fregeau u. Fournay 1993]. Anhand der Längenzuordnung eines internen Standards können DNA-Fragmente bis zu einem Basenpaar genau bestimmt werden [Kimpton et al. 1993].

Neben der Anwendung als Singleplex lassen sich STR-Marker auch als Multiplex bearbeiten [Edwards et al. 1991, Morral et al. 1992, Gill et al. 1990]. Hierbei erfolgt die Amplifikation mehrerer Genorte in einem gemeinsamen Reaktionsgefäß. Durch Verwendung unterschiedlicher Fluoreszenzfarbmarkierungen der Primer können sogar Loci mit überlappenden Allelgrößen zusammen bestimmt werden [Gill u. Evett 1995].

Bei Untersuchungen von Dinukleotid STRs ist häufig ein „enzymatischer Slippage“ während der Amplifikation beobachtet worden, welcher sich in Form artifizieller *stutter peaks* oder Stotterbanden im Elektropherogramm darstellte [Litt et al. 1993, Schlotterer u. Tautz 1992, Hauge u. Litt 1993]. Diese Peaks sind um ein oder mehrere Repeateinheiten kürzer als das Hauptallel und machen eine eindeutige Allelbestimmung schwierig. Tri- und Tetranukleotid-repeats dagegen, welche einen größeren Abstand zwischen den einzelnen Allelen besitzen, scheinen offensichtlich weniger zum „Slippage“ zu neigen und sind daher besser zur Identifikation von Personen geeignet [Edwards et al. 1991, Edwards et al. 1992, Urquhart et al. 1995].

Die Repeatstruktur von STR-Allelen kann in unterschiedlicher Weise variieren. Variationen können in der Länge der individuellen Repeateinheit, in der Anzahl von Repeateinheiten und durch Unterschiede im Repeatmuster einzelner Allele auftreten [Puers et al. 1993, Urquhart et al. 1993, Möller u. Brinkmann 1994]. *Simple* repeats enthalten Einheiten identischer Länge und Sequenz, *compound* repeats bestehen aus einem oder mehreren zusammengesetzten einfachen (*simple*) repeats und *complex* repeats sind durch unterschiedliche Blöcke von variabler Länge, die zusätzlich noch von mehr oder weniger variablen Sequenzen unterbrochen werden können, gekennzeichnet [Urquhart et al. 1994]. Voraussetzung zur

genauen Allelbestimmung sind Kenntnisse zur Sequenzstruktur der STR-Loci. Erst nach eingehenden Untersuchungen ist ein Einsatz der Marker zur Identifikation von Individuen und Vaterschaftsbegutachtung möglich.

## **1.2 Untersuchungsgegenstand und Ziel der Arbeit**

Ziel dieser Arbeit war es, fünf STR-Loci (D11S488, D18S51, D19S246, HUMFIBRA, HUMVWFA31/A) auf ihre Populationsgenetik und Sequenzstruktur zu untersuchen, sowie deren Eignung für den Einsatz in der Personenidentifikation und abstammungsgenetischen Begutachtung zu überprüfen.

Aufgrund relativer Unempfindlichkeit gegenüber enzymatischen Slippage wurde die Studie mit Tetranukleotidrepeatmarkern durchgeführt, deren Primersequenzen aus der Literatur entnommen wurden [Browne et al. 1993, Straub et al. 1993, Weber et al. 1993, Mills et al. 1992, Kimpton et al. 1992]. Bei der Auswahl der Systeme wurde auf eine unterschiedliche Chromosomenlokalisierung, ähnliche Amplifikationsbedingungen wie vergleichbare Annealingtemperaturen der Primer im Hinblick auf die Möglichkeit der Multiplex-Amplifikation und auf das Vorhandensein von mehreren häufigen Allelen geachtet.

Für die populationsgenetischen Untersuchungen wurden 100 gesunde, unverwandte kaukasische Blutspender aus der Region Hessen herangezogen.

Zunächst wurden die PCR- und Amplifikationsbedingungen der einzelnen Systeme in Vorversuchen optimiert. Im Hinblick auf die Etablierung eines Multiplex-Amplifikationssystems wurde die Annealingtemperatur angeglichen. Durch standardisierte Anwendung von PCR und Fragmentanalyse konnten reproduzierbare Ergebnisse erreicht werden.

Am automatischen 373A DNA-Sequenzierer (Applied Biosystems, Inc.) wurden die fluoreszenzfarbmarkierten Amplifikate in einem 6 %igen denaturierenden Polyacrylamidgel elektrophoretisch aufgetrennt. Die Längenbestimmung der DNA-Fragmente erfolgte anhand eines internen Längenstandards. Daraufhin konnten die Allel- und Genotypfrequenzen in der Populationsstichprobe ermittelt werden.

Zur Erfassung der exakten Basenpaarlänge wurden ausgewählte Allele der Loci mit Ausnahme von HUMVWFA31/A auf ihre Sequenz untersucht. Die Sequenzstruktur der Fragmente gab einen näheren Einblick in den Repeatbereich und sollte zur Klärung von Allelvariationen herangezogen werden.



## 2 Material und Methoden

### 2.1 Geräte

Analysenwaage	(Sartorius-Werke GmbH, Göttingen)
Computer	Macintosh II ci (Apple Computer, Inc., USA)
Elektrophoresekammer	BRL Horizon <sup>®</sup> 58 (Gibco-BRL, Life Technologies GmbH, Eggenstein)
Glasküvetten	(Pharmacia Biotech GmbH, Freiburg)
Greiner-Röhrchen	50 ml (Greiner GmbH, Frickenhausen)
Grobwaage	Modell 1200 (Acculab <sup>®</sup> , USA)
Kamera	The Imager <sup>™</sup> (Appligene Oncor, Heidelberg)
Magnetkonzentrator	für Standard PCR tubes (0,5 ml) DYNAL MPC-P-12 (Dynal, Oslo, Norwegen)
Magnetrührer	Ikamag <sup>®</sup> RCT (Kühn + Bayer, Nidderau-Heldenbergen) Modell MR 2002 (Heidolph, Kelkheim)
Membranfilter	0,2 µm Ø 47 mm RC 58 (Schleicher & Schüll GmbH, Dassel)
Netzgerät	Electrophoresis Power-Supply (Appligene, Illkirch, Frankreich)
PCR Strips	á 8 tubes (0,2 ml) (Biozym Diagnostik GmbH, Hameln)
PCR Cap Strips	á 8 Deckel (Biozym Diagnostik GmbH, Hameln)
pH-Meter	Mikroprozessor-pH-Meter (Knick, Mannheim)
Pipetten	Eppendorf Varipetten 0,5-10 µl; 2-20 µl; 10-100 µl; 50-250 µl; 200-1000 µl (Eppendorf-Netheler-Hinz GmbH, Hamburg)
Pipettenspitzen	(Sarstedt GmbH, Nümbrecht)
Reaktionsgefäße	0,5 ml; 1,5 ml; 2,0 ml (Eppendorf-Netheler-Hinz GmbH, Hamburg)
Schüttler (Vortexer)	Modell REAX 2000 (Heidolph, Kelkheim)
Schüttelwasserbad	Modell SW 21C (Julabo, Seelbach)
Sequenzier	ABI 373A DNA-Sequencer (Applied Biosystems, Inc., Foster City, USA)
Spektralphotometer	Ultrospec Plus (Pharmacia Biotech GmbH, Freiburg)
Speedvac	Univapo 150 H (Uniequip, Martinsried)

Thermocycler	Perkin Elmer 9600 (Perkin Elmer, Applied Biosystems, Inc., Weiterstadt) The PTC-100™ (MJ Research, Inc. Watertown, USA)
Tischinkubator	Eppendorf Thermomixer 5436 (Eppendorf-Netheler-Hinz GmbH, Hamburg)
Tischzentrifuge	Eppendorf 5415 C (Eppendorf-Netheler-Hinz GmbH, Hamburg)
Transferpipetten	(Sarstedt GmbH, Nümbrecht)
UV-Transilluminator	(Renner, Dannstadt)
Zentrifugen	Minifuge 2 Heraeus (Christ GmbH, Osterode) Servall® RC-3B (Du Pont de Nemours GmbH, Bad Homburg)

## 2.2 Populationsstichprobe

Als Populationsstichprobe dienten 100 gesunde, nicht verwandte kaukasische Blutspender aus der Region Hessen. Die Geschlechtsverteilung innerhalb des Personenkollektivs war ausgeglichen. Bei den Blutspendern handelte es sich sowohl um Hausspender des Blutspendedienstes Hessen des DRK in Frankfurt am Main, als auch um Erst- oder Mehrfachspender, die von den in ganz Hessen eingesetzten Blutentnahmeteams erfaßt wurden.

## 2.3 Untersuchungsmaterial

Untersucht wurde genomische DNA, die aus kernhaltigen Blutzellen, den Leukozyten, isoliert wurde. Dazu war den Blutspendern aus einer gestauten peripheren Armvene mit üblichem Entnahmebesteck 10 ml Blut entnommen worden, das zur Verhinderung der Blutgerinnung in EDTA-(Ethylen-Diamin-Tetraessigsäure-)Röhrchen gefüllt wurde.

### 2.3.1 DNA-Isolierung

#### **Leukozytenisolierung aus EDTA-Blut**

Das peripher entnommene EDTA-Blut (10 ml) wurde in ein 50 ml Greiner-Röhrchen überführt und mit der dreifachen Menge Ery-Lysis-Puffer (4 °C) versetzt. Gut gemischt kam der Ansatz für 15-20 Minuten in den Kühlschrank, wo er zwischenzeitlich erneut gemischt wurde. Anschließend sollte das Blut hämolysiert sein.

Nach einem Zentrifugationsschritt von 10 Minuten bei 2400 U/min bildete sich am Boden ein Leukozytensediment. Mit der Wasserstrahlpumpe wurde der Überstand abgesaugt und verworfen.

Das verbleibende Pellet wurde aufgeschüttelt und in 4 ml Ery-Lysis-Puffer resuspendiert. Unter gleichen Bedingungen folgte erneut ein Zentrifugationsschritt, zur möglichst vollständigen Entfernung der Erythrozyten. Anschließend wurde der Überstand verworfen. Die gewonnenen Zellen konnten bei -40 °C eingefroren oder direkt weiterverarbeitet werden.

#### **DNA-Isolierung nach Dykes**

Zum aufgeschüttelten Leukozytensediment wurden 3 ml Nuclei-Lysis-Puffer (4 °C) gegeben und gemischt. Der Ansatz wurde mit 200 µl 10 % SDS und 600 µl 2 mg% Proteinase K versetzt, kräftig geschüttelt und über Nacht bei 37 °C in ein Schüttelwasserbad gestellt.

War das Gemisch am nächsten Tag noch geliert, erfolgte eine nochmalige Zugabe von 200 µl Proteinase K und ein ein- bis zweistündiger Verbleib im Wasserbad. Bei erfolgreicher Inkubation wurden 1,2 ml gesättigte NaCl-Lösung hinzugegeben und kräftig geschüttelt (Aussalz-Methode [Miller et al. 1988]). Der Ansatz wurde für zehn Minuten bei 2500 U/min zentrifugiert und der Überstand mit einer Transferpipette in ein neues Röhrchen überführt.

Aufschütteln, Zentrifugieren und Überführen wurden noch zwei weitere Male wiederholt bis der Überstand ganz klar war.

Ein Teil des klaren abpipettierten Überstandes wurde mit 2,5 Teilen eiskaltem absoluten Ethanol versetzt. Das neue Röhrchen wurde verschlossen und vertikal rechtsdrehend geschwenkt bis die DNA ausfiel. Mit Hilfe einer Eppendorf-Pipettenspitze wurde die DNA vorsichtig aus der Flüssigkeit gefischt und in 70 %igem Alkohol gewaschen.

Je nach DNA-Menge erfolgte die Lösung in einem mit 100-500 µl 1xTBE-Puffer (50 °C) gefüllten Eppendorf-Tube (1,5 ml). Im Anschluß an eine etwa zweistündige Inkubation bei 50 °C im Wärmebad konnte die gelöste DNA bei -40 °C gelagert werden.

### **Photometrische Messung der DNA**

Die DNA wurde 1:100 mit Aqua dest. verdünnt (5 µl DNA + 500 µl Aqua dest.) und in eine saubere Küvette gegeben. Bei 260 nm und 280 nm erfolgte die photometrische Messung mit dem Spektralphotometer. Da das Extinktionsmaximum der Nukleinsäuren bei 260 nm und das der verunreinigten Proteine bei 280 nm liegt, wurde die Reinheit der Nukleinsäuren anhand des Quotienten 260/280 beurteilt. Er soll über 1,6 liegen.

### **Konzentrationsmessung der DNA**

Die Konzentration der DNA erhielt man durch folgende Berechnung:

Konzentration x in µg/µl

$$x = \frac{\text{Extinktion 260 nm} \times \text{Verdünnung (100)}}{\text{optische Dichte (20)}} = \text{Extinktion 260 nm} \times 5$$

#### **2.3.1.1 Reagentien**

Stammlösung 0,5 mol EDTA pH 8,0

186,12 g EDTA (Fluka Chemie, Buchs, Schweiz) in 800 ml Aqua dest. lösen, pH 8,0 mit 5 mol NaOH (Merck, Darmstadt) einstellen und auf 1000 ml mit Aqua dest. auffüllen. Das EDTA löst sich erst nach Zugabe von NaOH.

Stammlösung 1 mol Tris-HCl pH 7,5

121,14 g Tris (Merck, Darmstadt) in 800 ml Aqua dest. lösen, pH 7,5 mit konzentrierter HCl (Merck, Darmstadt) einstellen, auf 1000 ml mit Aqua dest. auffüllen.

TE-Puffer

10 ml der 1 mol Tris-HCl pH 7,5 Stammlösung + 2 ml der 0,5 EDTA pH 8,0 Stammlösung auf 1000 ml mit Aqua dest. auffüllen (bei -20 °C aufbewahren).

Gesättigte NaCl Lösung

350 g NaCl (Carl Roth GmbH, Karlsruhe) auf 1000 ml Aqua dest. geben.



#### 10 % SDS pH 7,2

100 g SDS (Oncor, Heidelberg) in 900 ml Aqua dest. unter Erwärmung bei 68 °C lösen, mit 0,5 mol NaOH (Merck, Darmstadt) pH 7,2 einstellen und auf 1000 ml mit Aqua dest. auffüllen.

#### Proteinase K Puffer

5 g SDS (Oncor, Heidelberg) + 0,744 g EDTA (Fluka Chemie, Buchs, Schweiz) auf 500 ml mit Aqua dest. auffüllen.

#### Proteinase K 2 mg%

100 mg Proteinase K (QIAGEN GmbH, Hilden) auf 50 ml mit Proteinase K Puffer auffüllen, proportionieren und bei -20 °C einfrieren.

#### Nuclei-Lysis-Puffer

1,21 g Tris (Merck, Darmstadt) in ca. 800 ml Aqua dest. lösen, mit HCl (Merck, Darmstadt) pH 8,2 einstellen und mit Aqua dest. auf 1000 ml auffüllen. Davon 104 ml abnehmen, verwerfen und durch 100 ml der 4 mol NaCl Stammlösung + 4 ml der 0,5 mol EDTA Stammlösung ersetzen.

#### Ery-Lysis-Puffer

8,29 g NH<sub>4</sub>Cl (Merck, Darmstadt) + 1 g KHCO<sub>3</sub> (Merck, Darmstadt) + 0,037 g EDTA (Fluka Chemie, Buchs, Schweiz) auf 1000 ml mit Aqua dest. auffüllen.

## **2.4 Short Tandem Repeat Loci**

### **2.4.1 Chromosomenlokalisierung**

Abbildung 2.1 zeigt die Lokalisation der in dieser Arbeit untersuchten fünf STR-Marker D11S488 (11q24.1-q25), D18S51 (18q21.3), D19S246 (19q13.3-q13.4), HUMFIBRA (4q28) und HUMVWFA31/A (12p12pter) auf ihren Chromosomen.

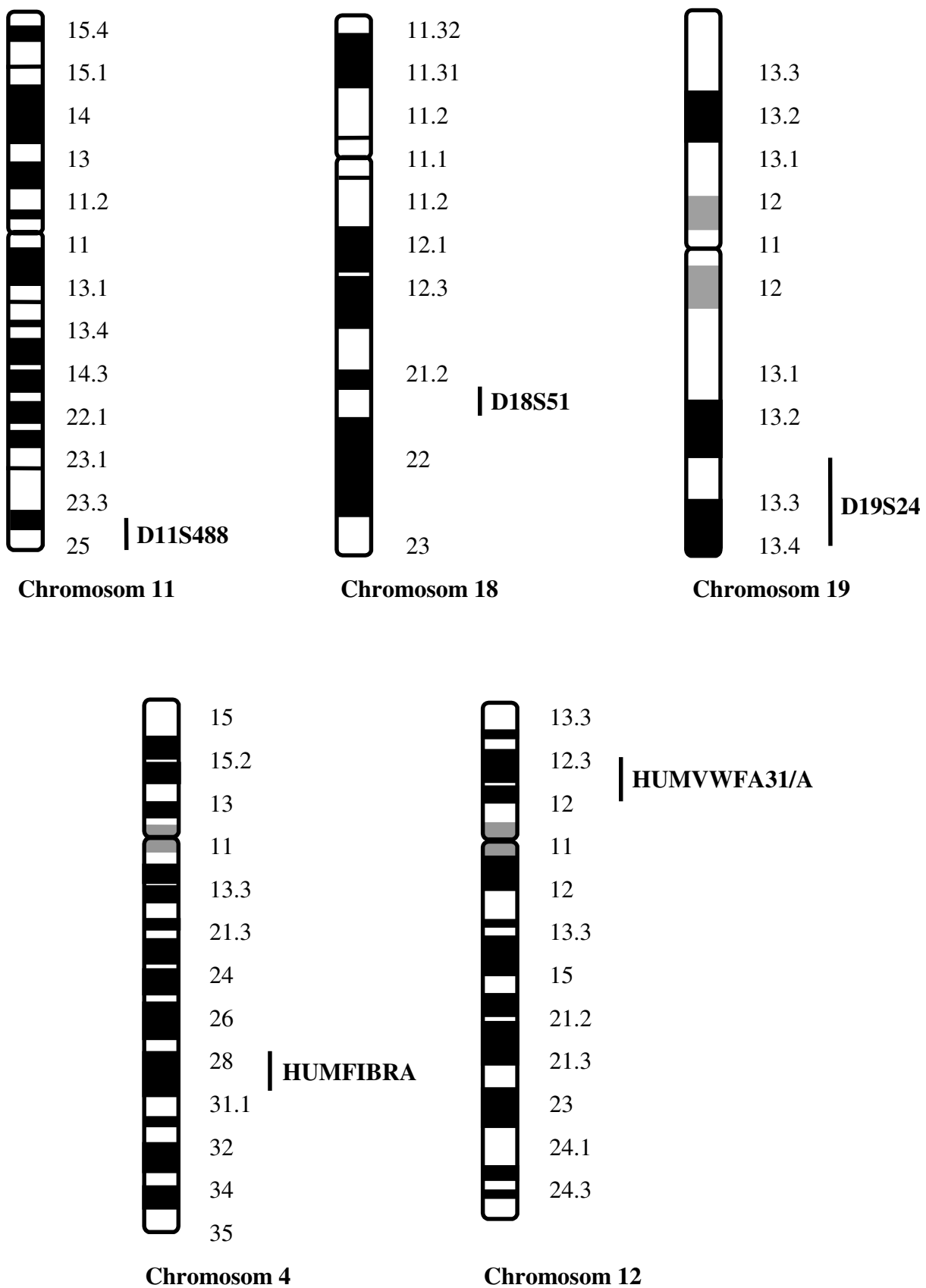


Abb. 2.1: Lokalisation der Marker auf ihren Chromosomen

## 2.4.2 Primersequenzen und Farbstoffmarkierung

Die Primersequenzen der STR-Marker sind in Tabelle 2.1 aufgelistet. P1 (Primer 1) entspricht dem Forward-Primer und P2 (Primer 2) dem Reverse-Primer. Auf die Berechnung der Primerschmelztemperatur  $T_m$  wird unter Kap. 2.4.3 näher eingegangen.

Tabelle 2.1: Liste der verwendeten Primer

Bezeichnung	Lokalisation	Sequenz 5'-3'	$T_m$
<b>D11S488</b> (Browne et al. 1993)	11q24.1-q25		
P1		AGGCAATAGAGACCCTGTG	58 °C
P2 <i>FAM</i>		GATGATGAATTGTTACTGAGAG	60 °C
<b>D18S51</b> (Straub et al. 1993)	18q21.3		
P1		GAGCCATGTTCATGCCACTG	62 °C
P2 <i>HEX</i>		CAAACCCGACTACCAGCAAC	62 °C
<b>D19S246</b> (Weber et al. 1993)	19q13.3-q13.4		
P1 <i>TAMRA</i>		GAAACACATCATTTACCCAC	56 °C
P2		AGAGTGAGATTCCACCTTTC	58 °C
<b>HUMFIBRA</b> (Mills et al. 1992)	4q28		
P1 <i>TAMRA</i>		GCCCCATAGGTTTTGAACTCA	62 °C
P2		TGATTTGTCTGTAATTGCCAGC	62 °C
<b>HUMVWFA31/A</b> (Kimpton et al. 1992)	12p12pter		
P1 <i>HEX</i>		CCCTAGTGGATGATAAGAATAATCAGTATG	82 °C
P2		GGACAGATGATAAATACATAGGATGGATGG	84 °C

Alle Primer wurden handelsüblich synthetisiert und von Applied Biosystems Inc., Weiterstadt bezogen. Jeweils einer der Oligonukleotide jedes Systems wurde mit einem fluoreszierenden Farbstoff versehen und in lyophilisierter Form geliefert. Die Lösung erfolgte in Aqua dest. (HPLC). Eine Zusammenstellung der Fluoreszenzfarbstoffe ist in Tabelle 2.2 wiedergegeben.

Tabelle 2.2: Liste der Fluoreszenzfarbstoffe

Markierung	Farbstoff	Farbe
HEX	4,7,2',4',5',7'-hexachloro-6-carboxy-fluorescein	grün
FAM	5-carboxy-fluorescein	blau
TAMRA	N,N,N',N'-tetramethyl-6-carboxyrhodamin	gelb
ROX	6-carboxy-X-rhodamin	rot

### 2.4.3 Berechnung des Primerschmelzpunktes

Die Schmelztemperatur  $T_m$  eines Primers läßt sich näherungsweise mit einer Vielzahl verschiedener Formeln berechnen. Primerpaare werden für jede Amplifizierung möglichst so konstruiert, daß beide Primer eine ähnliche  $T_m$  haben. Die Schmelztemperatur wurde in dieser Arbeit wie folgt berechnet [Newton u. Graham 1994]:  $T_m = (\text{Anzahl von A} + \text{T}) \times 2 \text{ } ^\circ\text{C} + (\text{Anzahl von C} + \text{G}) \times 4 \text{ } ^\circ\text{C}$ . Sie wird jedoch bei Primern, die länger als 20 Nukleotide sind, ungenau. Meist wählt man in der Praxis für die Anlagerung der Primer Temperaturen (Annealingtemperaturen), die 3-5  $^\circ\text{C}$  unter der berechneten  $T_m$  liegen. Geeignete Bindungstemperaturen lassen sich nur in etwa aus der Schmelztemperatur ableiten. Die berechnete  $T_m$  diene in den Vorversuchen zu Beginn der Experimente lediglich als Anhaltspunkt.

## 2.5 Polymerasekettenreaktion

### 2.5.1 Prinzip der PCR

Bei der Polymerasekettenreaktion (*polymerase chain reaction*; PCR) handelt es sich um eine *in vitro*-Technik, die 1985 von Kary Mullis [Saiki et al. 1985, Mullis et al. 1986] entwickelt wurde. Mit ihr kann man gezielt Desoxyribonukleinsäure-(DNA-)Abschnitte vervielfältigen, die von zwei bekannten DNA-Sequenzen, den Primern, eingerahmt werden. Die DNA, eine Doppelhelix, besteht aus zwei DNA-Strängen, die antiparallel verlaufen und durch komplementäre Basen miteinander verbunden sind. Über Wasserstoffbrücken bindet Adenin

(A) mit Thymin (T) und Guanin (G) mit Cytosin (C). Jede Base ist mit einem Zuckermolekül, einer Desoxyribose, verknüpft. Die 5'- und 3'-Kohlenstoffatome benachbarter Zucker sind über Phosphatgruppen miteinander verknüpft. Dadurch erhält die einzelsträngige DNA ein 5'- und ein 3'-Ende.

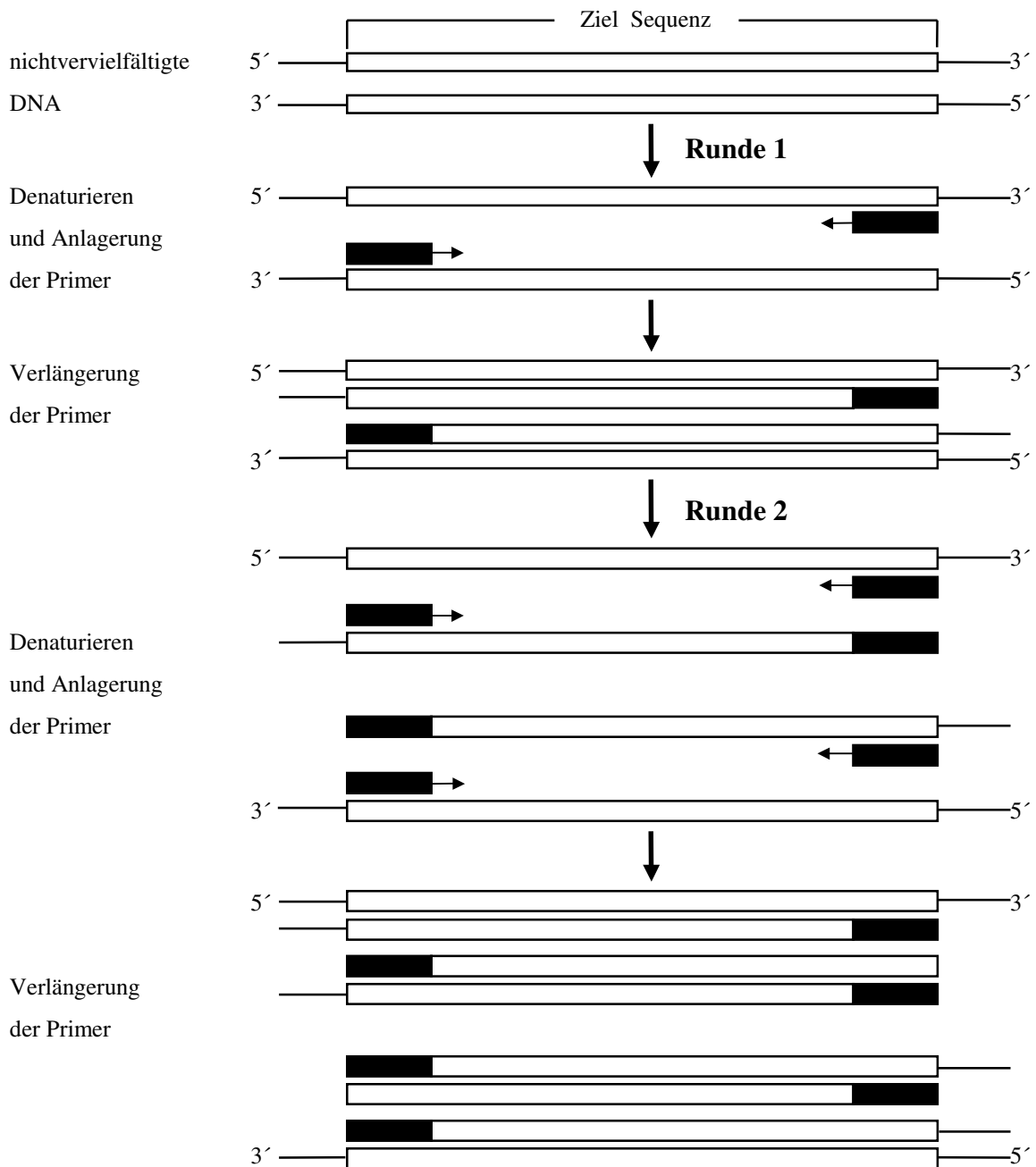


Abb. 2.2: Schematische Darstellung der PCR

Um mit Hilfe einer PCR die DNA amplifizieren zu können, werden als Starthilfe Oligonukleotidprimer benötigt. Es handelt sich dabei um kurze, einzelsträngige DNA-Moleküle, die komplementär zu den Enden einer definierten Sequenz der DNA-Matrize (*template*) sind. In Abbildung 2.2 zeigt ein Schema die ersten Runden einer PCR. Unter den richtigen Reaktionsbedingungen verlängert eine DNA-Polymerase in Gegenwart von Desoxynukleosidtriphosphaten (dNTPs) die Primer entlang der einzelsträngigen denaturierten DNA-Matrize. Somit werden neue DNA-Stränge, deren Sequenz komplementär zur Matrize ist, synthetisiert. Am Ende der Reaktion liegen die DNA-Moleküle als Doppelstränge vor. Bei einer Wiederholung der Synthese müssen die Doppelstränge erneut durch Hitze aufgeschmolzen werden (*denaturing*), damit sich die Primer nach dem Abkühlen an die DNA-Einzelstränge anlagern (*annealing*) können. Sobald die richtige Temperatur für die Enzymreaktion erreicht ist, verlängert die DNA-Polymerase die Primer (*extension*). Die neusynthetisierten DNA-Stränge können nun ebenfalls als Matrize dienen. Die Ziel-Sequenz vermehrt sich exponentiell, so daß nach 20 PCR-Runden eine Vermehrung der gesuchten DNA etwa um das  $2^{20}$ fache stattgefunden hat. In der Regel wird der Zyklus, bestehend aus Denaturierung, Primeranlagerung und -verlängerung, 20- bis 40mal wiederholt.

## 2.5.2 Durchführung der PCR

### 2.5.2.1 Geräte

Für die PCR Amplifikation werden Geräte benutzt, die die Temperatur zyklisch verändern können. Dabei handelt es sich um sogenannte Thermocycler. Sie führen die Denaturierung der DNA, die Anlagerung der Primer und die Kettenverlängerung mit einer DNA-Polymerase durch, wobei die Zyklen mehrmals wiederholt werden. Alle hier beschriebenen PCRs sind mit dem PTC-100 (Programmable Thermal Controller) von MJ RESEARCH, INC., Watertown, Massachusetts, USA durchgeführt worden.

### 2.5.2.2 Bedingungen

Für jeden STR-Marker wurden in Vorversuchen (Kap. 3.1) die Reaktionsbedingungen optimiert. Ein PCR-Ansatz bestand aus einem Gesamtvolumen von 50 µl, das sich wie folgt zusammensetzte:

2 µl	Primer 1 (10 pmol) <sup>1</sup>
2 µl	Primer 2 (10 pmol) <sup>1</sup>
5 µl	5 mM dNTP-Mix ( <i>Stratagene, Heidelberg</i> )
5 µl	10x PCR-Puffer (1xC) ( <i>Appligene Oncor, Heidelberg</i> ) <sup>2</sup>
10 µl	genomische DNA (100 ng)
2 µl	Taq-DNA-Polymerase (1 Unit) ( <i>Appligene Oncor, Heidelberg</i> )
24 µl	Aqua dest.

Nach Zugabe der Taq-DNA-Polymerase wurde der Ansatz in den Thermocycler gestellt und das entsprechende Programm gestartet. Die PCR-Bedingungen der einzelnen Systeme sind in Tabelle 2.3 für Single-Locus Amplifikationen und die entsprechenden Multiplex-Systeme zusammengefaßt.

Der Denaturierungsschritt zu Beginn (94 °C/95 °C) führte zur Trennung der DNA-Doppelstränge in jeweils zwei Einzelstränge. Die Anlagerung der Primer an die DNA fand bei einer für jeden Primer speziellen Temperatur, der Annealingtemperatur (hier 54 °C bzw. 60 °C), statt, auf die bereits oben unter der Berechnung der Primerschmelztemperatur eingegangen wurde. Während der Elongation oder Extension (72 °C) konnte mit Hilfe der hitzestabilen Polymerase ein komplementärer Strang entlang der Matrix gebildet werden. Die Wiederholung dieses Schrittes am Ende des Programms über einen längeren Zeitraum sollte für die vollständige Synthetisierung aller Stränge sorgen.

---

<sup>1</sup> bei HUMVWFA31/A wurde eine Primermenge von 25 pmol eingesetzt

<sup>2</sup> Zusammensetzung des 10x PCR-Puffers: 10 mM Tris HCl (pH 9,0); 50 mM KCl; 1,5 mM MgCl<sub>2</sub>; 0,1 % TritonX100; 0,2 mg/ml Gelatine

Tabelle 2.3: PCR-Bedingungen

System	Temperatur	Zeit	Zyklen
<b>D11S488</b>	95 °C	5 min	1
<b>D19S246</b>	95 °C	30 sec	30
	54 °C	40 sec	
	72 °C	30 sec	
	72 °C	5 min	1
<b>D18S51</b> <b>HUMVWFA31/A</b> <b>HUMFIBRA</b>	94 °C	5 min	1
	94 °C	60 sec	28
	60 °C	60 sec	
	72 °C	60 sec	
<b>Multiplex</b>	72 °C	10 min	1
	95 °C	5 min	1
	95 °C	30 sec	28
	54 °C	40 sec	
	72 °C	30 sec	
72 °C	5 min	1	

## 2.5.3 Nachweis der amplifizierten DNA

### 2.5.3.1 Prinzip

Anhand eines Agarosegels kann der Erfolg der PCR überprüft werden. Die amplifizierte DNA wird dazu in kleine Geltaschen des Agarosegels, sogenannte slots, pipettiert. Das Gel besteht aus einer elektrisch neutralen, festen Gelmatrix. Aufgrund ihres Zucker-Phosphat-Rückrats sind Nukleinsäuren bei allen pH-Werten negativ geladen. Folglich wandern sie während einer Elektrophorese von der positiv geladenen Kathode zur Anode. Je kleiner die Fragmente sind, desto schneller ist ihre Wanderungsgeschwindigkeit im Gel. Der Trennbereich linearer Doppelstrang-DNA liegt bei 0,1 bis drei Kilobasen in einem 2 %igen Agarosegel. Die aufgetrennten Fragmente werden nach der Elektrophorese angefärbt und stellen sich in Form von Banden dar. Als Farbstoff dient Ethidiumbromid, das sich zwischen die Basenpaare der DNA einlagert und unter UV-Licht (254 nm) sichtbar gemacht werden kann.



### **2.5.3.2 Durchführung**

#### **Herstellung des Agarose-Minigels (2 g%)**

Zur Herstellung des Gels wurden 2 g Agarose abgewogen und zusammen mit 100 ml 1x TBE-Puffer in ein Becherglas gegeben. Das Glas wurde abgedeckt und auf den mit 250 °C eingestellten Magnetrührer plaziert bis die Agarose vollständig gelöst war. Nach kurzem Abkühlen wurden jeweils 20 ml der gelösten Agarose mit einer Pipette luftblasenfrei in einen Gelschlitten überführt, der mit zwei Kämmen á 14 Zähnen ausgelegt war. Etwa 20-30 Minuten später war das Gel erstarrt. Der Gelschlitten wurde in die mit TBE-Puffer gefüllte Elektrophoresekammer gesetzt, wobei der Puffer das Gel bedeckte. Anschließend konnten die Kämmen entfernt und die Gelschlitze beladen werden.

#### **Beschicken des Agarosegels**

Mindestens eine Geltasche pro Reihe wurde mit 2,5 µl des DNA-Längenstandards (123 bp DNA-Leiter) beladen. Von den PCR Amplifikaten wurden jeweils 5 µl entnommen, mit 1 µl Gelladepuffer gemischt und in eine freie Geltasche pipettiert. Auf diese Weise wurde mit weiteren Proben verfahren bis alle Geltaschen des Agarosegels gefüllt waren. Anschließend wurde das Gel in der Elektrophoresekammer an ein Gleichstromnetz von etwa 60-90 Volt angeschlossen. Die Dauer der Elektrophorese war abhängig von der angelegten Spannung und betrug zwischen 20 und 40 Minuten.

#### **Färben und Fotografieren**

War die Gelelektrophorese beendet, wurde das Gel für etwa 10 Minuten in ein Färbebad mit Ethidiumbromid gegeben. Somit konnten unter UV-Licht die DNA-Banden sichtbar gemacht und fotografiert werden.

### 2.5.3.3 Reagentien

#### Agarose

Agarose NA (Pharmacia AB, Uppsala, Schweden)

#### 10x TBE-Puffer

Tris-Borat-EDTA-Puffer (pH 8,0) Rotiphorese<sup>®</sup> (Carl Roth GmbH, Karlsruhe)

108 g Trometamol (Tris), 55 g Borsäure, 8,3 g EDTA ad 1000 ml Aqua dest.

#### Gelladepuffer

0,1 g Bromphenolblau (Sigma, Deisenhofen), 0,1 g Xylene Cyanol

(Sigma Deisenhofen), 8,0 g Ficoll (Sigma, Deisenhofen) ad 40 ml Aqua dest.

#### Ethidiumbromid-Stammlösung

0,1 g Ethidiumbromid (Sigma, Deisenhofen) ad 40 ml Aqua dest.

#### DNA-Längenstandard

1 Teil 123 bp DNA-Leiter (Gibco-BRL, Eggenstein), 1 Teil Reaktionspuffer (10 mM Tris HCl (pH 7,5), 50 mM NaCl, 0,1 mM EDTA) (Gibco-BRL, Eggenstein), 5 Teile Gelladepuffer, 4 Teile Aqua dest.

#### 123 bp DNA-Leiter

Bei der 123 bp DNA-Leiter handelt es sich um einen Marker, mit dessen Hilfe die Länge von doppelsträngiger DNA zwischen 123 bp und 3075 bp bestimmt werden kann. Die einzelnen Fragmente des Markers bestehen aus Längen von 123 bp bis 4182 bp, siehe Tabelle 2.4. Bis zu 25 einzelne Banden können auf dem Agarosegel nach Ethidiumbromidbehandlung sichtbar gemacht werden. Bezugsquelle ist Gibco-BRL, Eggenstein.

Tabelle 2.4: 123 bp DNA-Leiter (Gibco-BRL)

Fragment	Größe in bp	Fragment	Größe in bp
1	4182	18	2091
2	4059	19	1968
3	3936	20	1845
4	3813	21	1722
5	3690	22	1599
6	3567	23	1476
7	3444	24	1353
8	3321	25	1230
9	3198	26	1107
10	3075	27	984
11	2952	28	861
12	2829	29	738
13	2706	30	615
14	2583	31	492
15	2460	32	369
16	2337	33	246
17	2214	34	123

## 2.6 Detektion von STR-Markern

### 2.6.1 Prinzip der Fluoreszenzmarkierung und automatischen Detektion

Mit dem automatischen DNA-Sequencer 373A von Applied Biosystems kann in einer Fragmentanalyse die Länge von fluoreszenz-farbmarkierten DNA-Fragmenten genau bestimmt werden. Der Farbstoff in den PCR-Produkten läßt sich während der Amplifizierung durch einen fluoreszenz-markierten Primer einführen, der am 5'-Ende eingebaut wird. Die Primer können dazu an einen der Farbstoffe HEX, ROX, FAM oder TAMRA (Applied Biosystems, Inc.) gekoppelt sein.

Der Sequencer beinhaltet die Elektrophorese- und Detektionseinheit. Zur automatischen Detektion werden vorher bearbeitete DNA-Proben auf ein denaturierendes 0,4 mm dickes Polyacrylamidgel aufgetragen, das zwischen zwei fluoreszenzfreie, planare Glasplatten gegossen wird, und elektrophoretisch der Länge nach aufgetrennt. Ein Argonlaser, dessen Hauptemission bei Wellenlängen von 514 nm und 488 nm liegt, ist das Kernstück der Detektionseinheit (Abb. 2.3a). Der Laserstrahl wird über drei Spiegel in einem Winkel von 42° auf das zwischen den Glasplatten polymerisierte Gel (Abb. 2.3b) gerichtet.

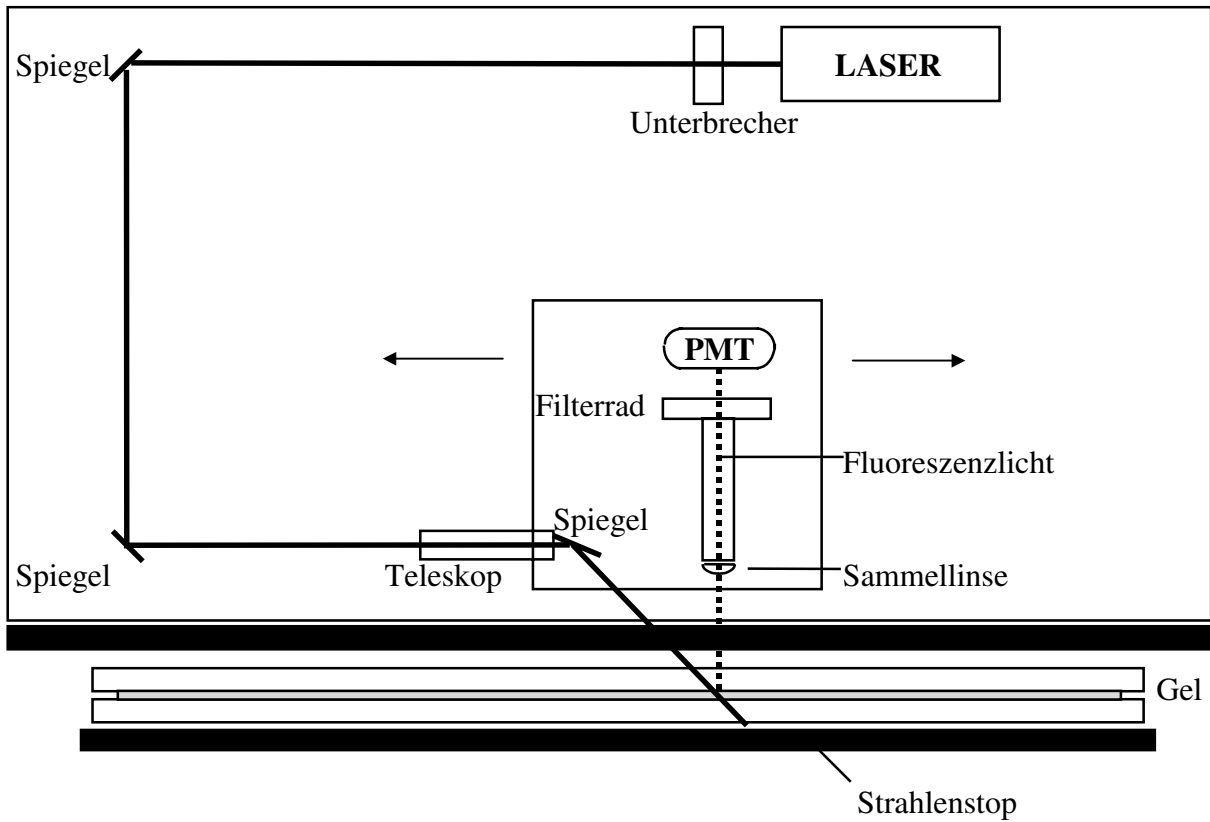


Abb. 2.3a: Schematische Darstellung des Strahlengangs von Laser- und Fluoreszenzlicht im 373A DNA-Sequenzier (Blick von oben)

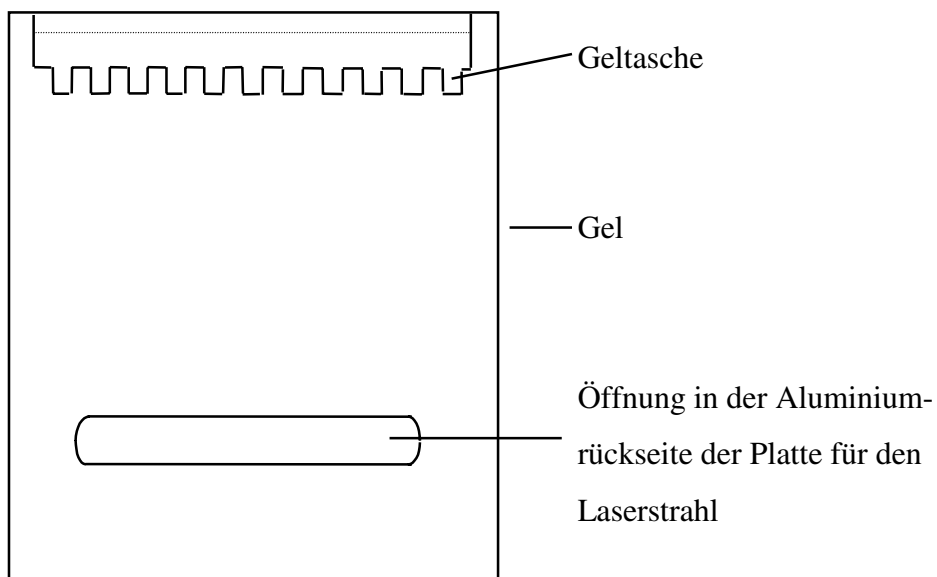


Abb. 2.3b: Gel im Blick von vorn

Der letzte der drei Spiegel befindet sich auf einem beweglichen Schlitten. Dieser enthält außerdem ein Filtrerrad, das zwischen den einzelnen Farben diskriminiert und einen Photomultiplier. Der Photomultiplier (PMT) wandelt das Lichtsignal in ein elektrisches, digitales Signal um, welches direkt auf den Computer übertragen wird und dort ein Gelbild generiert.

Zur Aufnahme der Rohdaten tastet (*scanned*) der Laser während der Elektrophorese kontinuierlich einen definierten Bereich des Gels für alle vier Farben ab. Passiert ein markiertes DNA-Fragment den Scan-Bereich, wird der an den Primer gekoppelte Farbstoff zur Sekundärstrahlung angeregt. Die von einer Linse fokussierte Strahlung wird durch das aus verschiedenen Filtern bestehende Filtrerrad geschickt und von dem Photomultiplier detektiert. Dieser registriert Zeitpunkt, Farbe und Intensität der Lichtemission.

Die automatische Längenbestimmung von DNA-Fragmenten erfolgt mit Hilfe eines internen Längenstandards. Der Standard ist farbstoff-markiert und kann von anders markierten Proben unterschieden werden, die in der gleichen Spur aufgetragen werden. Mit der Genescan™ 672 Software wird in der „*Southern Local*“-Methode aus den bekannten Fragmentlängen des Standards eine Kalibrierungskurve erzeugt, wobei die drei Fragmente, die dem unbekanntem Fragment am nächsten liegen, in die Berechnung einbezogen werden. Durch Extrapolation können so die Längen unbekannter PCR-Produkte bestimmt werden. Da der Standard in jeder Spur (*Lane*) des Gels mitläuft und jede Spur für sich analysiert wird, sind Fehlberechnungen aufgrund von Variationen in der elektrophoretischen Mobilität innerhalb des Gels praktisch ausgeschlossen.

Der Computer erstellt am Ende zu jeder Lane ein Elektropherogramm, in dem die gemessenen Fluoreszenzintensitäten gegen die Zeit (Scan-Nummer bzw. bp-Länge) aufgetragen werden. Dadurch werden die Fragmente in Form von „Peaks“ sichtbar. Die Fläche unter den Spitzen (*peak area*) ist direkt proportional zur eingesetzten DNA-Menge.

## **2.6.2 Durchführung**

### **Polyacrylamidgel**

Zur Fragmentanalyse wurde ein denaturierendes 6%iges Polyacrylamidgel verwendet. Zunächst mußten zwei 24 cm well-to-read Glasplatten (25 x 42 cm) mit einem Detergenz (Alconox) und Aqua dest. gründlich gereinigt werden. Danach wurden sie mit Isopropanol

nachbehandelt und getrocknet. Nacheinander wurden die Glasplatten zum Probescannen in den DNA-Sequencer eingebaut und auf ihre Reinheit überprüft. Wiesen die Platten noch Fluoreszenzhintergrund auf, waren sie erneut zu reinigen. Anschließend erfolgte das Zusammensetzen der Glasplatten zur Gelkassette und die Vorbereitung des Gels. Zur Gelherstellung wurden 30 g Urea abgewogen, in ein Becherglas gegeben, mit 23,5 ml Aqua dest. (HPLC) aufgefüllt und bei 50 °C mit einem Rührkern gemischt bis eine klare Lösung entstand. Diese war darauf mit 6,0 ml 10x TBE und 9,0 ml 40 % Acrylamid versetzt, vacuum filtriert (0,2 µ Filter) und entgast worden. Nach Zugabe von 45 µl TEMED und 400 µl 10 % APS wurde das Gel rasch blasenfrei in die Gelkassette gegossen. Für die Geltaschen wurde ein Kamm mit 24 Zähnen eingesetzt. Etwa 90-120 Minuten später war das Gel in horizontaler Lage polymerisiert, die Gelkassette wurde im Scan-Bereich erneut geputzt und in den Sequencer eingebaut. Nachdem die Pufferkammern mit 1x TBE-Laufpuffer gefüllt, der Kamm vorsichtig gezogen und die Geltaschen mit Laufpuffer gespült waren, wurde der Vorlauf von 60 Minuten gestartet.

### **Probenvorbereitung und Elektrophoresebedingungen**

Zunächst wurden die PCR-Produkte mit Aqua dest. verdünnt. Die Verdünnung richtete sich nach der Fluoreszenzfarbe des Primers (blau ist kräftiger als gelb) und nach der Stärke der angefärbten Banden im Agarosegel. Amplifikate mit kräftigen Fluoreszenzfarben und dicken Banden wurden dementsprechend höher verdünnt als andere. In der Regel reichte die Palette von unverdünnt bis zu Verdünnungen von 1:20.

Beim Beladen der Geltaschen galt es folgendes zu beachten: Eine Geltasche durfte mit mehreren unterschiedlich farb-markierten Proben beladen werden. Sollten zwei Proben, deren Primer die gleiche Fluoreszenzfarbe besaßen, in der gleichen Lane laufen, so mußten die STR-Loci unterschiedliche Fragmentlängen besitzen, um bei der Auswertung unterschieden werden zu können.

Für jede Geltasche wurden 1,5 µl Probengemisch entnommen und zusammen mit 1,0 µl des Längenstandards und 2,5 µl Formamid gut gemischt in ein 0,5 ml Eppendorf-Cup gegeben. Nach dreiminütiger Denaturierung bei 95 °C im Thermomixer wurden die Proben auf Eis gelegt, um eine Zusammenlagerung der Stränge zu vermeiden.

Die Geltaschen wurden noch einmal mit Laufpuffer durchgespült und mit 4,7 µl Aliquot beladen. Die Elektrophoresebedingungen lagen bei 30 Watt (W), 1600 Volt (V) und

28 Milliampere (mA). Die Laufdauer betrug zehn Stunden und konnte abhängig von der Fragmentlänge variieren.

### **Der Standard**

Zur Verbesserung der Fragmentlängenberechnung wurde ein laboreigener Längenstandard verwendet. Im Vergleich zu den kommerziell erhältlichen Standards (GENESCAN 2500) ist dieser Standard speziell für den molekularen Bereich zwischen 50-400 bp geeignet und zeichnet sich durch eine gleichmäßige Verteilung der Standardfragmente aus (Abb. 2.4). Bei den Fragmenten dieses Standards handelt es sich um PCR-Amplifikate, welche unter Verwendung verschiedener Primerkombinationen vom Vektor pGL2-basic (Promega) gewonnen und auf ihre Längen mittels Sequenzierung untersucht wurden. Der interne Standard war mit der Fluoreszenzfarbe ROX (rot) markiert. Durch Weiterentwicklung kamen in den Analysen verschiedene Versionen zum Einsatz (Tabelle 2.5).

*Tabelle 2.5: Interner Standard*

<b>Interner Standard (Version 1)</b>		<b>Interner Standard (Version 2)</b>	
Fragment	Größe in bp	Fragment	Größe in bp
1	105	1	41
2	133	2	64
3	142	3	78
4	170	4	105
5	206	5	119
6	234	6	133
7	266	7	142
8	294	8	156
9	342	9	170
10	370	10	206
11	407	11	220
12	435	12	234
		13	266
		14	280
		15	294
		16	342
		17	356
		18	370
		19	407
		20	421
		21	435

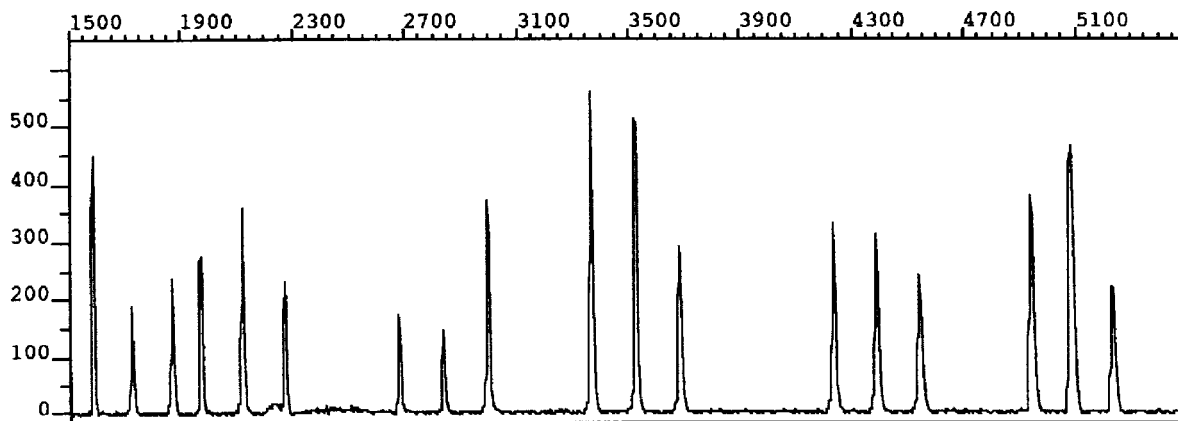


Abb. 2.4: Darstellung des internen Standards (Version 2) als Elektropherogramm. Standard-Peaks von links nach rechts: 105, 119, 133, 142, 156, 170, 206, 220, 234, 226, 280, 294, 342, 356, 370, 407, 421 und 435 bp (x-Achse: Scan-Nummer (=Funktion der Zeit), y-Achse: relative Fluoreszenzintensität).

### 2.6.3 Reagentien

#### Urea

Rotiphorese<sup>®</sup> NF-Harnstoff (Carl Roth GmbH, Karlsruhe)

#### Aqua dest. (HPLC)

Wasser HPLC (J.T. Baker, Deventer, Holland)

#### 40 % Acrylamid

Rotiphorese<sup>®</sup> NF-Acrylamid/Bis-Lösung 40 % (19:1) (Carl Roth GmbH, Karlsruhe)

#### 10x TBE-Puffer

Rotiphorese<sup>®</sup> NF-10x TBE-Puffer, fluoreszenzfrei (Carl Roth GmbH, Karlsruhe)

108 g Tris Base, 55 g Borsäure, 8,3 g Na<sub>4</sub>EDTA, lichtgeschützt bei RT aufbewahren.

Arbeitspuffer 1x entspricht: 89 mM Tris Base, 89 mM Borsäure, 2 mM EDTA pH ~8,3

#### 10 % APS

Ammoniumpersulfat: 1,0 g APS Molecular (Aldrich Chemical, Gillingham, England)

ad 10,0 ml Aqua dest. (HPLC), lichtgeschützt max. 5 Tage bei +4 °C haltbar



TEMED

N,N,N',N'-Tetramethylethylendiamin (BIO-Rad Laboratories, Richmond, USA)

Isopropanol

2-Propanol (Merck, Darmstadt)

Formamid

CH<sub>3</sub>NO (Eastman Chemical, New Haven, USA)

#### **2.6.4 Auswertung der erhobenen Daten**

Die Analyse der erhobenen Rohdaten erfolgte wie unter Kap. 2.6.1 beschrieben mit der Genescan™ 672 Software von Applied Biosystems. Als eine Zusatzoption zum 373A DNA-Sequenzier ist sie insbesondere für klinische und rechtsmedizinische Fragestellungen geeignet. Mit der Software ließ sich die Länge von DNA-Fragmenten im Bereich von 40-1000 Basen bis auf eine Base genau bestimmen. Die Genauigkeit der Längenbestimmung lag bei 99,5 %. Zunächst wurden alle Daten, die während des Elektrophoreselaufs anfielen, registriert und aus den Rohdaten ein COLLECTION FILE erstellt. Da Peaks, die auf nicht verbrauchte Primer bzw. Primer-Dimere zurückzuführen waren, nicht im Elektropherogramm erscheinen sollten, wurde der Collection File nachbereitet und an Stellen, die größer oder kleiner als der Standardbereich waren geschnitten. Daraus resultierte der GEL FILE. Visuell wurde überprüft, ob der Rechner die Lanes bzw. Spuren richtig zugeordnet hatte. Für den Standard wurden manuell die entsprechenden Basenpaarwerte eingegeben, sie dienten zur Berechnung der unbekannt Fragmentlängen. Über die Analysis Parameter wurde die Schwelle für die Peakhöhe so gewählt, daß der Computer nur die beiden interessierenden Peaks jedes Allels analysierte und kleinere Peaks, die möglicherweise auf eine unspezifische Bindung der Primer oder „Stutter peaks“ zurückzuführen waren, unbeachtet blieben. Wurde ein Peak vom Computer in mehrere Flächen unterteilt, so addierten sich die *peak areas*. Auf diese Weise wurde jede Lane analysiert und vom Genescan Programm ein RESULTS FILE erstellt. Er beinhaltete die Elektropherogramme der Proben, in denen die DNA-Fragmente als farbige Peaks dargestellt waren.

## 2.7 Sequenzierung von STR-Allelen

### 2.7.1 Prinzip der Sanger Sequenzierung

Mit der Methode von Sanger [Sanger et al. 1977] kann die Basenabfolge des Amplifikats bestimmt werden. Sie beruht auf dem Prinzip der Desoxymethode, bei der eine DNA-Polymerase die Primer mit den vorhandenen vier dNTP-Molekülen (dATP, dCTP, dGTP und dTTP) entlang einer einzelsträngigen Matrize verlängert. Zur Anwendung kommt die T7-DNA-Polymerase [Tabor u. Richardson 1987]. Vor der Sequenzierreaktion muß der DNA-Doppelstrang durch Hitzebehandlung denaturiert und so voneinander getrennt werden.

Der Einbau von einem Didesoxynukleosidtriphosphat (ddNTP), einem dNTP-Analogen, führt zu einem zufälligen Abbruch bei der Sequenzierreaktion. Der Wahrscheinlichkeit nach wird der DNA-Strang an jeder Stelle mindestens einmal durch ein ddNTP nicht mehr weiter amplifiziert und bricht ab. Folglich entstehen unterschiedlich lange DNA-Ketten, die alle mit einem Didesoxynukleosidtriphosphat enden. In einer hochauflösenden Polyacrylamid-gelelektrophorese werden die DNA-Stränge aufgetrennt. Unter Zuhilfenahme unterschiedlich farbmarkierter ddNTP-Moleküle (ddATP, ddCTP, ddGTP, ddTTP) (DyeDeoxy-Terminatoren<sup>TM</sup>, Applied Biosystems) wird der Nachweis der Fragmente erbracht. Jede der vier Didesoxybasen erhält eine eigene Fluoreszenzfarbe. Wie bereits unter Kap. 2.6.1 beschrieben tastet während des Sequenzierlaufs ein Laser den Scanbereich kontinuierlich ab und bringt fluoreszenzmarkierte Basen zum Leuchten. Durch den zufälligen Kettenabbruch sind die Ketten unterschiedlich lang geworden. Während der Elektrophorese wandern zunächst die kurzen und später die längeren Ketten durch das Gel. Die farbmarkierten Basen, die zeitlich aufeinanderfolgend in einer Spur von dem Laser erfaßt werden, ergeben am Ende des Sequenzlaufs die Basensequenz für den untersuchten DNA-Abschnitt.

#### **Fluoreszenz-markierte ddNTP's**

Ebenso, wie sich Primer mit Fluoreszenzfarbstoffen markieren lassen (Kap. 2.6.1), können die Farbstoffe (JOE, ROX, FAM und TAMRA) mit einer speziellen patentrechtlich geschützten Technik an ddNTP-Moleküle (DyeDesoxy-Terminatoren<sup>TM</sup>, Applied Biosystems) gekoppelt werden. Die Farbe des Fluoreszenzfarbstoffs signalisiert eine der vier Didesoxyreaktionen, so daß man jede Base anhand ihrer Fluoreszenz identifizieren kann. Die Terminatoren ermöglichen es, mit jedem Primer zu sequenzieren, ohne diesen extra mit einem Farbstoff

markieren zu müssen. Jede Base hat ihren eigenen Farbstoff, der jeweils bei einer anderen Wellenlänge fluoresziert. Das hat den Vorteil, daß alle Reaktionen in einem Gefäß durchgeführt werden können und für jede Probe nur eine Spur im Gel benötigt wird. Nur Fragmente, die ein Didesoxynukleotid an ihrem 3'-Ende eingebaut haben, können von dem Laserstrahl, der das Gel abtastet, einzeln erfaßt werden.

### **Biotinylierte PCR-Primer**

Amplikons lassen sich sequenzieren, indem einer der beiden Stränge an eine feste Phase bindet, die DNA anschließend denaturiert und der fixierte Strang sequenziert wird. Dazu ist einer der PCR-Primer am 5'-Ende mit Biotin markiert. Während der Amplifikation wird Biotin in einen Strang des DNA-Produktes eingebaut. Nach Zugabe von DYNABEADS<sup>®</sup> zum PCR-Produkt kann das Biotin in einer 15 minütigen Reaktion kovalent an Streptavidin binden [Hultmann et al. 1991] und mit einem Magneten fixiert werden. Bei den DYNABEADS<sup>®</sup> handelt es sich um mit Streptavidin besetzte Eisenoxyd-Partikel, die paramagnetischen Charakter besitzen. Der nicht mit Biotin markierte Strang wird mit NaOH eluiert und zurück bleibt die einzelsträngige DNA, die an ihrem 5'-Ende fixiert ist. Mit Hilfe eines Sequenzierprimers kann dann eine Sanger-Sequenzierung an der Festphase gestartet werden.

Die Primersequenzen der STR-Loci entsprechen der Auflistung in Kap. 2.4.3. Statt der Farbmarkierung wurde eine Biotinmarkierung von Applied Biosystems Inc., Weiterstadt vorgenommen. Die PCR-Bedingungen unterschieden sich in der Menge des eingesetzten Primers (4 pmol für den biotinylierten und 12 pmol für den nicht markierten Primer). Um Sequenzierungen in beide Richtungen zu ermöglichen, wurden Primersets sowohl mit als auch ohne Biotinmarkierung angefertigt.

Der Erfolg der PCR wurde im Agarosegel überprüft.

## **2.7.2 Durchführung der Sequenzierung mit Sequenase**

### **Polyacrylamidgelelektrophorese**

Die Auftrennung der DNA-Fragmente bei der Sequenzierung erfolgte wie auch bei der Fragmentanalyse auf einem Polyacrylamidgel. Die Herstellung des Gels wurde bereits unter Kap. 2.6.2 beschrieben. Allerdings kam bei der Sequenzierung ein anderer Kamm mit 36 Zähnen zur Anwendung.

### **Probenvorbereitung und Elektrophoresebedingungen**

Die biotinylierten PCR-Produkte wurden an magnetische mit Streptavidin beschichtete Kügelchen (DYNABEADS<sup>®</sup>, Oslo, Norwegen) gebunden und mit 0,15 N NaOH/Triton (2,5 ml aus 10N NaOH in 47,5 ml Aqua dest. und 37,5 µl Triton X100) denaturiert. Daraufhin konnte der Gegenstrang entfernt und verworfen werden. Mit Hilfe eines Magneten wurden die DNA-Einzelstränge des biotinmarkierten Primers an der Wand des Reaktionsgefäßes fixiert und in zwei neutralisierenden Schritten mit TT-Puffer (10 mM Tris/HCL, pH 8,0, 0,01 % Tween 20) gewaschen.

Die Durchführung der Sequenzierreaktion erfolgte mit 1 pmol nicht biotinmarkierten Primer unter Verwendung der T7-DNA-Sequenase und DyeDeoxy-Terminatoren entsprechend den Angaben des Herstellers (Prism Sequenase und DyeDeoxy-Terminator Sequencing Kit, Perkin Elmer).

Die Produkte der Sequenzierreaktion wurden in einem 6 %igen Polyacrylamidgel auf einer Distanz von 24 cm getrennt.

Für den 373A DNA-Sequenzer (Applied Biosystems, Inc., Perkin Elmer) lagen die Elektrophoresebedingungen bei 2500 V, 40 mA und 30 W. Die Laufzeit war abhängig von der Fragmentlänge und belief sich in der Regel auf 8-10 Stunden. Unter Verwendung von SeqEditor<sup>™</sup> und Sequence Navigator<sup>™</sup> Software (Applied Biosystems, Inc., Perkin Elmer) wurden die Daten der Sequenzierung analysiert.

### **2.7.3 Taq Cycle Sequencing**

#### **Prinzip**

Zur Sequenzierung von DNA nach der Sanger-Methode gibt es bei den Polymerasen neben der oben beschriebenen T7-Polymerase (Sequenase) auch die Taq-Polymerase [Innis et al. 1988] (Perkin Elmer), welche Grundlage zum *Taq Cycle Sequencing* ist. Die Taq-Polymerase ist ein hitzestabiles Enzym aus dem Bakterium *Thermus aquaticus*. Vorteile des Taq Cycle Sequencing sind die geringeren Mengen an DNA, die im Gegensatz zur Sequenzierung mit Sequenase erforderlich sind. Statt einer alkalischen Denaturierung, bei der viele Pipettierschritte mit eventuellen Probenverlust nötig sind, wird thermisch vorgegangen. Da die Strangtrennung nicht über magnetische Kügelchen erfolgt, sind keine Biotin-markierten Primer nötig. Das Prinzip des Taq Cycle Sequencing sieht folgendermaßen aus. Zum DNA-

Template werden dNTP's, farbmarkierte Terminatoren (ddNTP's), AmpliTaq<sup>®</sup> DNA-Polymerase und Primer hinzugegeben. Der Reaktionsansatz wird in einen Thermocycler gegeben, in dem quasi erneut eine PCR stattfindet, und unter Anwesenheit von ddNTP's markierte Sequenzierungsprodukte unterschiedlicher Länge entstehen. Nach einem initialen Denaturierungsschritt folgen in mehreren Zyklen Anlagerungs-, Verlängerungs- und Denaturierungsschritte. Im Anschluß muß das Produkt von überschüssigen farbmarkierten Terminatoren gereinigt werden. Diese Abtrennung erfolgt an speziellen Säulen.

Der Unterschied zur Sequenase-Sequenzierung liegt in der größeren Ausbeute an Sequenzierungsprodukten, was durch den zyklischen Vorgang erreicht wird. Damit ist dem Taq Cycle Sequencing der Vorzug zu geben, wenn nur geringe Mengen an DNA-Template zur Verfügung stehen.

### Durchführung

Zunächst erfolgte die Entfernung überschüssiger PCR-Primer über die S 400 Säule (Pharmacia Biotech GmbH, Freiburg). In 0,2 ml PCR tubes wurde folgender Reaktionsansatz mit einem Endvolumen von 20 µl zusammenpipettiert:

- 4 µl Terminator ready mix (ABI PRISM<sup>™</sup> Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq<sup>®</sup> DNA Polymerase, FS)
- 70-200 ng DNA (PCR-Produkt)
- 10 pmol/µl Primer
- mit Aqua dest. auf ein Gesamtvolumen von 20 µl auffüllen

Entsprechend der zu sequenzierenden Richtung wurde der Forward- oder Reverse-Primer eingesetzt. Die Taq Cycle Sequencing-PCR lief im Perkin Elmer 9600 Thermocycler mit folgendem Programm (Tabelle 2.6).

*Tabelle 2.6: PCR-Bedingung für Taq Cycle Sequencing*

	Temperatur	Zeit	Zyklen
<b>Taq Cycle Sequencing</b>	96 °C	10 sec	26
	50 °C	5 sec	
	60 °C	4 min	
	4 °C	endlos	1

Nach der Sequenzierungs-PCR wurden die überschüssigen ddNTP's entfernt. Dies geschah mit Hilfe der MicroSpin™ G-50 Säule (Pharmacia Biotech GmbH, Freiburg). Das Amplifikat konnte sorgsam auf die vorbereiteten Säulen aufgetragen werden. Bei einer Zentrifugation für 1 min bei 2000x g erhielt man das gereinigte Produkt, welches über weitere 30 min in einer vorgeheizten Vakuumzentrifuge getrocknet wurde bis nur noch ein Pellet zurückblieb. Das Pellet wurde mit 3,5 µl Ladepuffer (5 Teile Formamid, 1 Teil EDTA) aufgefüllt, über 2 min bei 95 °C auf dem Thermoblock denaturiert und anschließend auf Eis gestellt. Nun konnten die Proben auf das Gel aufgetragen werden. Die Elektrophoresebedingungen entsprachen denen der Sequenase-Sequenzierung.

### **2.7.4 Gelextraktion**

#### **Prinzip**

Mit Hilfe des QIAEX® II Gel Extraction Kits (QIAGEN GmbH, Hilden) ist es möglich, DNA-Fragmente aus Agarosegelen zu extrahieren und so zu reinigen, daß sie zur Sequenzierung geeignet sind. Dazu wird das PCR-Produkt auf ein Agarosegel aufgetragen. Um Kontaminationen zu vermeiden, sollte jedes Mini-Gel nur mit einer einzigen Probe in ausreichender Menge beladen werden. Nach elektrophoretischer Trennung des Amplifikates können durch sachgemäßes Ausschneiden der Banden unter UV-Licht die einzelnen Fragmente voneinander getrennt werden, d.h. bei einem Heterozygoten können die zwei Banden (Allele) unterschiedlicher Länge getrennt ausgeschnitten und weiterverarbeitet werden. Nichtgebundene Primer und andere überschüssige Reagentien befinden sich an anderer Stelle im Gel und werden nicht herausgeschnitten. Mit dem Kit besteht die Möglichkeit DNA-Fragmente von 40 bp bis zu 50 kb unter einer hohen Salzkonzentration an QIAEX II Partikel zu binden. In mehreren Waschsritten sollen Verunreinigungen, wie Agarose, Proteine, Salz und Ethidiumbromid entfernt werden. Am Ende steht die gereinigte DNA, die weiter in PCR oder Sequenzierung eingesetzt werden kann.

#### **Durchführung**

Die Gelextraktion wurde mit dem QIAEX® II Gel Extraction Kit der Firma QIAGEN GmbH, Hilden durchgeführt. Die einzelnen Schritte wurden an das vom Hersteller mitgelieferte Protokoll angelehnt.

Das PCR-Produkt wurde für jede Probe auf einem 2 %igen Agarosegel in 1x TBE-Puffer elektrophoretisch aufgetrennt und mit Ethidiumbromid gefärbt. Die zu eluierende DNA-Bande wurde unter UV-Licht mit einem sterilen Skalpell aus dem Gel ausgeschnitten und in ein 1,5 ml Eppendorftube überführt. Das Gelstück wurde gewogen und pro 100 mg mit 300 µl QX1-Puffer (QIAGEN) versetzt. Die hohe Salzkonzentration dieses Puffers bewirkte das Lösen der Agarose. Der Zugabe von 10 µl resuspendiertem QIAEX II, einem Gemisch aus Kieselgel-Körnchen, schloß sich eine zehnmünütige Inkubation bei 50 °C an. Dabei vollzog sich das Schmelzen des Agaroseblöckchens und die Absorption der Nukleinsäuren an die Kieselgelpartikel. Die Bindung der DNA an das QIAEX II war abhängig von den großen Anionen im Salzpuffer QX1 und dem pH-Wert. Deshalb wurde nach der Inkubation der pH-Wert überprüft, war er größer als pH 7,5, erfolgte die Korrektur mit 10 µl 3M Natriumacetat. Der Ansatz wurde bei 13000 U/min für 30 Sekunden zentrifugiert und der Überstand, bestehend aus Agarose, Proteinen, Ethidiumbromid und Salzresten, verworfen. Nach einem erneuten Waschschrift mit 500 µl QX1-Puffer wurden die letzten Agarosereste über die Zentrifugation entfernt. Anschließend erfolgte zweimaliges Waschen mit Ethanol enthaltendem PE-Puffer, so daß nach Zentrifugation die Salzreste mit dem Überstand verworfen werden konnten. Nun ließ man das Pellet in den nächsten 10-15 Minuten bei Raumtemperatur trocknen. Durch Zugabe von 20 µl Aqua dest. wurde die DNA aus dem Kieselgel herausgelöst, über einen Zentrifugationsschrift mit dem Überstand abpipettiert und in ein sauberes Eppendorfcup überführt.

## **2.7.5 Auswertung der Sequenzen**

### **Sequence Analysis Software**

Eine erste Analyse der DNA-Sequenzierungsreaktionen wurde mit dem Programm 373A Analysis (Version 1.2) von ABI (Applied Biosystems Inc., Weiterstadt) durchgeführt. Die Software ist in der Lage, 24 bzw. 36 verschiedene Proben auf einem Sequenzierungsgel zu bearbeiten. Eine Ablage der Daten des 373A DNA-Sequenzers erfolgt in Form eines Rohdaten Gelfiles. Das Analysis Programm startete automatisch, nachdem alle Daten gesammelt waren. Im Anschluß an das *Lane tracking* wurde für jede Sequenzierreaktion ein Sample file angelegt, der die Rohdaten (RAW DATA) enthielt. Diese wurden analysiert und als ANALYZED DATA in Form eines Chromatogramms ausgegeben. Eine weitere

Bearbeitung der Daten erfolgte mit dem SEQUENCE NAVIGATOR™ Programm von ABI (Applied Biosystems Inc., Weiterstadt).

### **Sequence Navigator Software**

Nachdem die Sequenzierungsdaten gesammelt und von der Analysis Software analysiert worden waren, erfolgte die Weiterbearbeitung der Daten mit dem SEQUENCE NAVIGATOR™. Dazu wurden die Sequenzen in ein Layout geladen. Innerhalb des Layouts konnten Sequenzen durch die gleichzeitige Einblendung mehrerer Elektropherogramme miteinander verglichen werden. Identische oder unterschiedliche Basen wurden besonders kenntlich gemacht und waren über das Elektropherogramm näher zu betrachten. Die Darstellung der Basen erfolgte mit den entsprechenden Anfangsbuchstaben A/G/C/T, ein N erschien bei Basen die nicht zugeordnet werden konnten. Kamen mehrere Basen an einer Position in Frage, konnten diese im IUB-Code (siehe Anhang) beschrieben werden. Ursache dafür war meist die Anwesenheit einer Heterozygoten-Sequenz. Reverse-Sequenzen ließen sich komplementär umkehren und mit der Forward-Sequenz zusammenlegen. Verschiedene Fenster erlaubten es, Ansichten zum Elektropherogramm oder der Sequenz herzustellen.

### **2.7.6 Allelbezeichnung**

Die Allelbezeichnung wurde in Anlehnung an die Empfehlungen der ISFH (International Society for Forensic Haemogenetics) [DNA Commission of the ISFH 1992, 1994, Bär et al. 1997] vorgenommen, deren Benennung auf der Anzahl kompletter Repeateinheiten beruht. Interallele waren mit ihrer Anzahl Basen des inkompletten Repeats hinter einem Dezimalpunkt vermerkt worden. Damit war die Sequenzierung Voraussetzung für eine gemeinsame Namensgebung von Allelen. Dieses Übereinkommen ist unabhängig von der angewandten Methode und läßt einen Vergleich und Austausch von Ergebnissen aus unterschiedlichen Laboratorien zu.



## 2.8 Statistische Auswertung

Die *Allelfrequenz* errechnet sich aus der Anzahl der beobachteten Allele dividiert durch die Zahl der untersuchten Individuen.

Die Anzahl der beobachteten Allele dividiert durch die Zahl der untersuchten Chromosomen (Individuen  $\times$  2) entspricht der *Genfrequenz*. Hierbei werden homozygote Merkmalsträger doppelt gezählt.

Zur Berechnung der *Heterozygotenrate* wurde die Anzahl der Heterozygoten durch die Gesamtindividuenzahl geteilt.

Unter der *sample gene diversity* (GeneD) versteht man die Heterozygotenrate, welche unter dem Hardy-Weinberg-Equilibrium erwartet wird. Ihr Wert wurde entsprechend der Beschreibung von Kimpton [Kimpton et al. 1993] berechnet.

Die Beurteilung des Hardy-Weinberg Gleichgewichts basiert auf dem Vergleich von beobachteten und erwarteten Genotypen einer Population unter Verwendung des Chi-Quadrat-Tests „chi-squared goodness of fit test“ [Kirby 1992]. Bei weniger als fünf Ereignissen innerhalb erwarteter Genotypklassen wurden Allelgruppen gebildet [Rand et al. 1992].

Der PIC-Wert (*polymorphic information content*) wurde wie von Botstein [Botstein et al. 1980] vorgeschlagen berechnet. Er beschreibt den Informationsgehalt eines Markers bei einer gemeinsamen Vererbung mit einem bestimmten Allel.

Die Wahrscheinlichkeit, daß zwei Individuen in ihrem Phänotyp übereinstimmen, wird mit dem pM-Wert (*probability of match*) ausgedrückt, dessen Berechnung nach Jones [Jones 1972] ausgeführt wurde. Der *discrimination index* (DI) entspricht der Wahrscheinlichkeit, mit der sich zwei Individuen in ihrem Phänotyp unterscheiden, er errechnet sich aus der Differenz  $1 - pM$ .

Zur Berechnung der Vaterschaftsausschlußwahrscheinlichkeit wurde der AVACH-Wert (*Allgemeine Vaterschafts-Ausschluß-Chance*) nach Garber [Garber u. Morris 1983, Hummel et al. 1971] ermittelt.



## 3 Ergebnisse

### 3.1 Vorversuche zur Optimierung der PCR-Bedingungen

Vor Beginn der populationsgenetischen Untersuchungen wurden in Vorversuchen für die einzelnen Systeme geeignete PCR-Bedingungen ausgetestet. Weiterhin wurden im Hinblick auf die Konstruktion eines Multiplex-Amplifikationssystems, auf das weiter unten eingegangen wird, vergleichbare Annealingtemperaturen für die unterschiedlichen Systeme in Vorversuchen ermittelt. Ausgangspunkt dieser Versuche stellte die in Kap. 2.4.2 beschriebene Berechnung der Primer spezifischen Schmelztemperatur dar.

#### **Vorversuche zur Fragmentanalyse**

PCR-Bedingungen, die von den Erstbeschreibern der untersuchten STR-Loci angegeben waren, wurden als Ausgangspunkt für die entsprechende Optimierungsstudie verwendet. Die PCR für D11S488 [Browne et al. 1993] lief zunächst unter der Touchdown-Methode mit einer Annealingtemperatur von 58 °C. Die Systeme D18S51, HUMFIBRA und HUMVWFA31/A wurden nach Urquhart [Urquhart et al. 1995] bei 60 °C amplifiziert und D19S246 nach eigenen Berechnungen bei 54 °C. Das Reaktionsvolumen betrug für alle Marker 50 µl.

Davon ausgehend wurden die PCR-Bedingungen der einzelnen Marker systematisch variiert. In gestarteten Versuchsreihen (Abb. 3.1) fanden Vervielfältigungen bei Annealingtemperaturen von 54 °C, 58 °C, 60 °C und 62 °C statt. Drei verschiedene Reihen setzten sich wie folgt zusammen, wobei  $x$  der Annealingtemperatur und  $y$  dem Ausgangswert beim Touchdownverfahren entsprach. Im ersten Zyklus nahm  $y$  den Wert  $(x + 10)$  an.

#### *Versuchsreihe 1:*

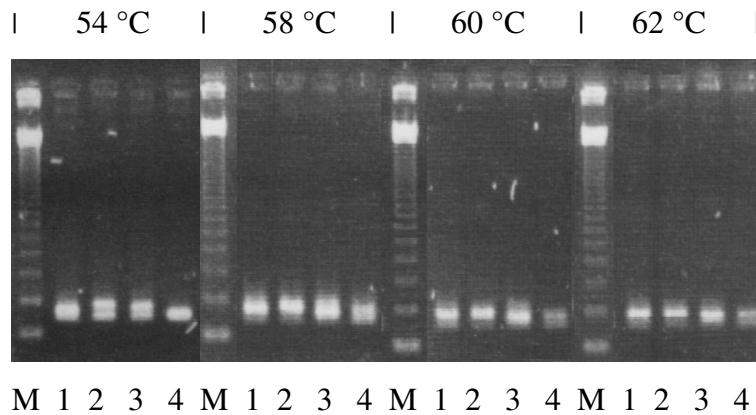
10 Zyklen mit der Folge 94 °C für 30 sec,  $y$  °C für 30 sec, 72 °C für 60 sec, wobei  $y$  bei jedem Zyklus um 1 °C verringert wird, weitere 20 Zyklen bei 94 °C für 30 sec,  $x$  °C für 30 sec, 72 °C für 60 sec und ein letzter Extensionsschritt bei 72 °C für 5 min. [Browne et al. 1993]

#### *Versuchsreihe 2:*

28 Zyklen mit der Folge 94 °C für 60 sec,  $x$  °C für 60 sec, 72 °C für 60 sec und ein letzter Extensionsschritt bei 72 °C für 10 min. [Urquhart et al. 1995]

*Versuchsreihe 3:*

30 Zyklen mit der Folge 95 °C für 30 sec, x °C für 40 sec, 72 °C für 30 sec und ein letzter Extensionsschritt bei 72 °C für 5 min



*Abb. 3.1: Ausschnitt aus der Versuchsreihe 3 für D19S246. Die Amplifikationsprodukte sind in einem 2 %igen Agarosegel bei Annealingtemperaturen von 54 °C, 58 °C, 60 °C und 62 °C dargestellt. Als DNA-Längenstandard (M) ist die 123 bp DNA-Leiter zu sehen. Das Gel wurde nach der Elektrophorese in Ethidiumbromid gefärbt und unter UV-Licht fotografiert. Die mit den Nummern 1 bis 4 versehenen Produkte entstammen der Populationsstichprobe.*

Die Vervielfältigung der untersuchten Loci mittels Touchdown-Methode erwies sich als ungeeignet. In den anderen Versuchsreihen wurde bei den unterschiedlichen Systemen im Vergleich zu den Ausgangsbedingungen keine auffällige Änderung in Qualität oder Quantität des Amplifikates beobachtet. Aufgrund der Resultate aus den Vorversuchen wurden die populationsgenetischen Untersuchungen der einzelnen STR-Loci wie im Methodenteil beschrieben durchgeführt.

Im Hinblick auf die Etablierung eines Multiplexsystems wurde in den Vorversuchen gleichzeitig nach Bedingungen gesucht, unter denen sich alle verwendeten Systeme amplifizieren ließen. Abbildung 3.2 zeigt einen Ausschnitt der fünf Marker unter drei verschiedenen Amplifikationsbedingungen. Aufgrund der Resultate aus den Vorversuchen und in Ergänzung von Ergebnissen aus der Fragmentanalyse erfolgte die Durchführung der Multiplexuntersuchungen bei 54 °C.

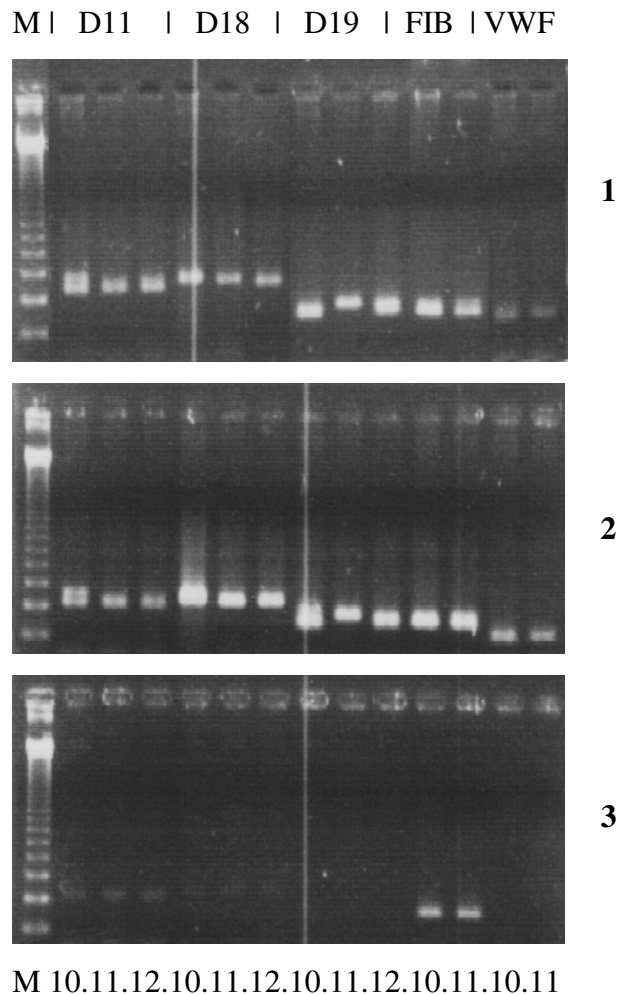


Abb. 3.2: Darstellung von Amplifikationsprodukten der fünf Marker unter drei Bedingungen im Agarosegel. D11S488, D18S51, D19S246, HUMFIBRA und HUMVWFA31/A wurden folgendermaßen vervielfältigt: **1**: 30 Zyklen bei 95 °C für 30 sec, 54 °C für 40 sec, 72 °C für 30 sec; **2**: 28 Zyklen bei 95 °C für 60 sec, 60 °C für 60 sec, 72 °C für 60 sec; **3**: Touchdown (TD) bei 58 °C für 30 sec. DNA-Längenstandard (M) ist die 123 bp DNA-Leiter.

### Vorversuche zur Sequenzierung

Zunächst wurde versucht, die oben ermittelten PCR-Bedingungen der Fragmentanalyse auch für die Sequenzierung zu übernehmen. Dies war bei D18S51, D19S246 und HUMFIBRA ohne weiteres möglich, nicht jedoch bei D11S488. Die Touchdown-PCR brachte mit den biotinylierten Primern für die Sequenase-Sequenzierung Amplifikate schlechter Qualität und Quantität hervor. Nacheinander wurden verschiedene Parameter ohne Erfolg verändert.

- Einsatz unterschiedlicher Mengen für den mit Biotin markierten Primer von 4 bis 10 pmol
- Zugabe genomischer DNA von 100 bis 200 ng pro Ansatz
- Änderung der Annealingtemperatur im Touchdownverfahren von 54 °C bis 60 °C
- Erhöhung der Zyklenzahl von 30 auf 35

Aufgrund der schlechten Ergebnisse wurde von der Biotinmarkierung und Touchdown-PCR abgesehen und D11S488 entsprechend den oben beschriebenen PCR-Ergebnissen bei 54 °C amplifiziert. Um ein reines Produkt für die Sequenzierreaktion zu erhalten, kam die Gelextraktion mit QIAEX II zur Anwendung. Das Eluat aus diesem Verfahren wurde noch einmal mit einer herkömmlichen PCR bei 54 °C amplifiziert und wegen der eher geringen Amplifikatmenge mittels Taq Cycle Sequencing sequenziert.

Mit der Möglichkeit bei der Gelextraktion Heterozygote durch das Ausschneiden der Banden aus dem Agarosegel voneinander zu trennen, kam das QIAEX II Kit mitunter auch im D19S246 System zur Anwendung. Dementsprechend wurden die Marker unter folgenden Methoden sequenziert.

<b>D11S488</b>	Taq Cycle Sequencing
<b>D18S51</b>	Sequenase-Sequenzierung
<b>D19S246</b>	Sequenase-Sequenzierung, Taq Cycle Sequencing bei Auftrennung von Heterozygoten
<b>HUMFIBRA</b>	Sequenase-Sequenzierung
<b>HUMVWFA31/A</b>	wurde nicht sequenziert

### **3.2 Etablierung eines Multiplex Amplifikationssystems**

Mit Hilfe der Polymerasekettenreaktion ist es möglich, mehrere Amplifikationsreaktionen in einem Reaktionsgefäß mit der gleichen Ziel-DNA parallel durchzuführen. Mit dem Einsatz verschieden farbmarkierter Primer können unterschiedliche Bereiche simultan amplifiziert und durch Auftragen in eine Spur des Polyacrylamidgels analysiert werden. Man spricht hierbei von Multiplex. Der Vorteil liegt darin, daß vergleichsweise geringe Mengen an genomischer DNA notwendig sind. Gerade in Bereichen wie der forensischen Medizin findet die Multiplex-Methode ihre Anwendung in der Identifikation von Individuen. Zahlreiche

Arbeitsschritte und Reagentien entfallen, wenn nicht für jedes System ein getrennter Ansatz zusammenpipettiert werden muß.

Bevor ein Multiplex PCR-Amplifikationssystem konstruiert werden kann, müssen einige Voraussetzungen gegeben sein. Die einzelnen Systeme in einer Multiplex-Gruppe sollen sich entweder in ihrer Fragmentlänge oder Farbmarkierung eindeutig voneinander unterscheiden. Amplifikationsbedingungen und Annealingtemperatur müssen übereinstimmen. Damit sich in der Auswertung die Peakhöhen entsprechen, ist auf eine Angleichung der Primermenge zu achten.

Bereits bei der Auswahl der Primer der in dieser Arbeit untersuchten Systeme wurde auf ähnliche Schmelztemperaturen geachtet, wie aus Tabelle 3.1 zu ersehen ist. Die Farbmarkierung wurde so gewählt, daß möglichst jeder Marker unterschiedlich gekennzeichnet war, wenn Überschneidungen in den Fragmentlängen zu erwarten waren.

*Tabelle 3.1: Amplifikationscharakteristika der untersuchten Systeme*

<b>System</b>	<b>Allellänge (bp)</b>	<b>Farbmarkierung</b>	<b>Annealingtemperatur</b>
D11S488	242-302	FAM	58-60 °C
D18S51	278-322	HEX	60 °C
D19S246	182-230	TAMRA	54 °C
HUMFIBRA	180-212	TAMRA	60 °C
HUMVWFA31/A	134-162	HEX	60 °C

In den Vorversuchen (Kap. 3.1) wurden für alle Systeme gemeinsame Amplifikationsbedingungen mit einer Annealingtemperatur von 54 °C ermittelt. In einer weiteren Versuchsreihe war die optimale Menge für die einzelnen Primer bestimmt worden, welche sehr gering gehalten werden sollte. Ausgehend von 0,5 pmol für kräftige Systeme bis hin zu 25 pmol für schwache Amplifikate wurde versucht, immer zwei STR-Marker einander anzugleichen. Bei Primermengen unter 8 pmol war die Allelausprägung unzureichend und mit 3 pmol die Nachweisgrenze in gemischten Proben erreicht. Die Magnesiumkonzentration lag bei 1,5 mM.

Die Untersuchungen ergaben für alle Systeme die Fähigkeit zum Multiplex. Entsprechende Primermengen sind in Abbildung 3.3 und 3.4 wiedergegeben. D19S246 und HUMFIBRA ließen sich aufgrund der Farbmarkierung und Allellänge nicht gemeinsam darstellen und wurden in zwei Gruppen zu einem Tri- und Quadruplex aufgeteilt.

Auf der Ordinate ist die Peakhöhe in Fluoreszenzeinheiten aufgetragen, die zwischen 100 und 4000 liegen können. Die Abszisse ist in Scan-Nummern aufgeteilt, durch Verrechnung mit dem internen Längenstandard läßt sich die Allelgröße in Basenpaaren (bp) ermitteln.

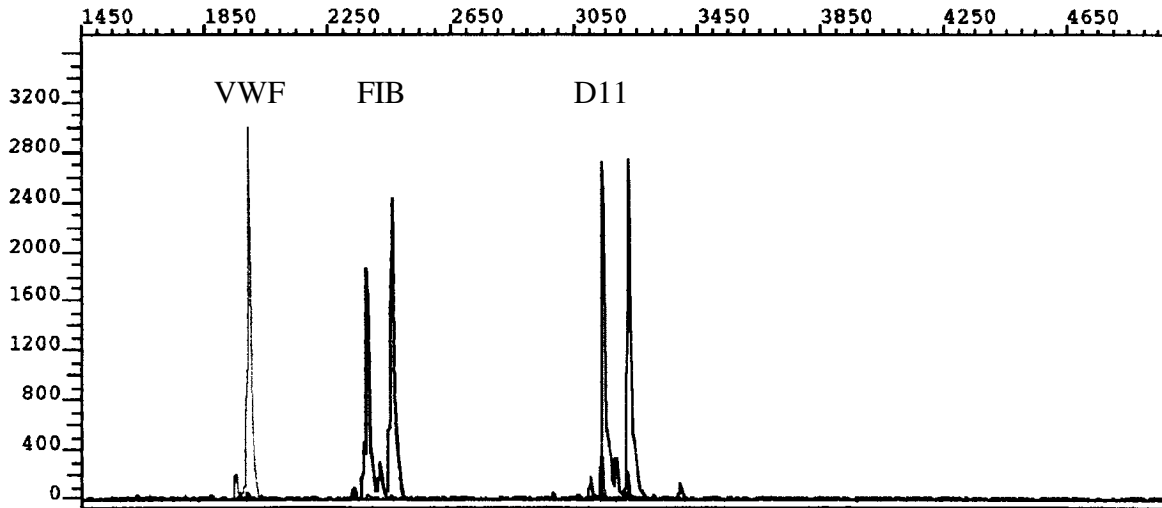


Abb. 3.3: Darstellung des Triplex als Elektropherogramm. D11S488 (10 pmol), HUMFIBRA (10 pmol) und HUMVWFA31/A (25 pmol)

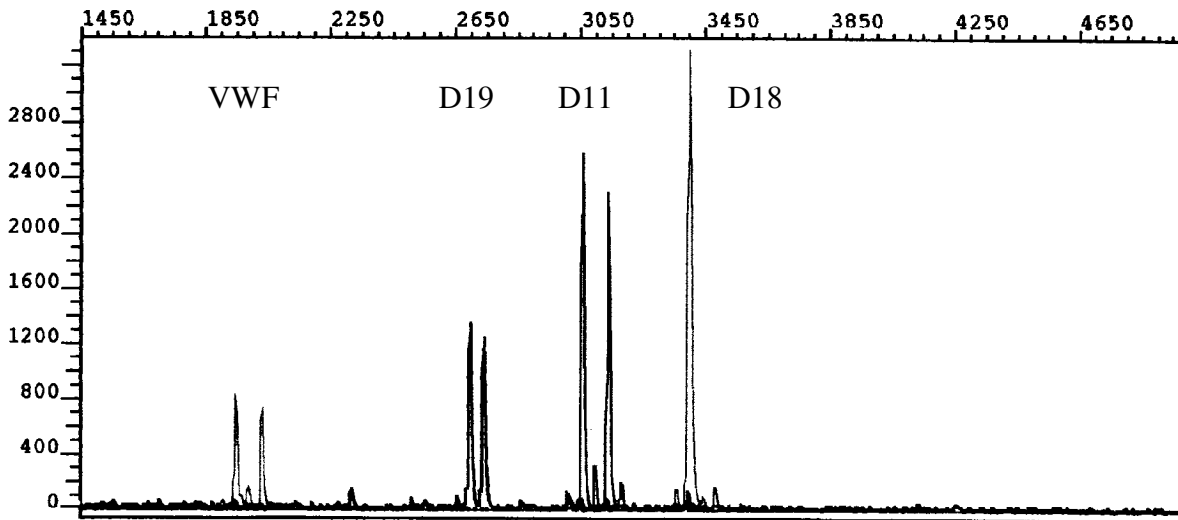


Abb. 3.4: Darstellung des Quadruplex als Elektropherogramm. D11S488 (10 pmol), D18S51 (10 pmol), D19S246 (10 pmol) und HUMVWFA31/A (25 pmol)



### 3.3 Populationsgenetische Untersuchungen

Fünf Short Tandem Repeat (STR) Loci wurden auf ihren Polymorphismus in einer Populationsstichprobe von 100 gesunden unverwandten hessischen Blutspendern untersucht. Die auf unterschiedlichen Chromosomen lokalisierten Marker ließen sich mit Hilfe der PCR-Technik amplifizieren.

Die Fluoreszenzfarbmarkierung jeweils eines Locus spezifischen Primerpaares ermöglichte es, die Amplifikate über Laserscanning in der Fragmentanalyse zu detektieren. Dies erfolgte mit dem automatischen 373A Sequenzer, in welchem die Fragmente elektrophoretisch auf einem 6 %igen denaturierenden Polyacrylamidgel aufgetrennt und anhand eines internen Längenstandards bestimmt worden waren.

Die Nomenklatur der Allele wurde in Anlehnung an die *DNA Commission of International Society of Forensic Haemogenetics* (ISFH) vorgenommen. Um die Sequenzstruktur der Fragmente kennenzulernen, die für den Polymorphismus der STR-Marker verantwortlich ist, erfolgte die Sequenzierung ausgewählter Proben.

#### 3.3.1 HUMD11S488

In der Fragmentanalyse wurden für jedes Individuum die entsprechenden Allele bestimmt. Abbildung 3.5 zeigt das Elektropherogramm einer Probe für den Locus D11S488. Die beiden Spitzen (Peaks) stellen die Fragmente eines Heterozygoten dar, repräsentiert durch die Fluoreszenzintensität des farbmarkierten DNA-Produktes.

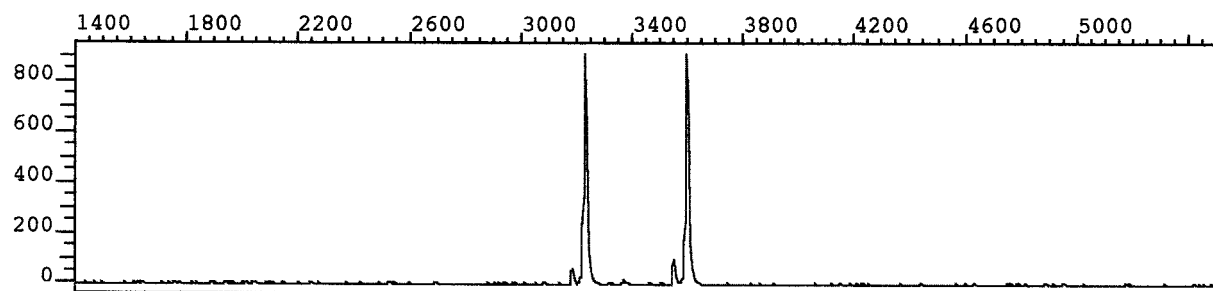


Abb. 3.5: Darstellung der DNA-Fragmente von D11S488 als Elektropherogramm. Die beiden kleinen Peaks jeweils links des großen sind „stutter peaks“ und 4 bp kürzer als das zugehörige Hauptallel.

### 3.3.1.1 Allelfrequenzen und Genotypen

Die Analyse des Locus D11S488 ergab einen Polymorphismus von 15 Allelen, die zwischen 242 bp und 302 bp lagen. In Abbildung 3.6 sind die beobachteten Allelfrequenzen graphisch aufzeigt. Der x-Achse ist die Größe der Allele in Basenpaaren (bp) zu entnehmen. Die Höhe der Säulen entspricht der Frequenzhäufigkeit. Es ist eine bimodale Verteilung zu erkennen, wobei die häufigsten Allele unter dem zweiten Hügel auftraten. Mit einer Frequenz von 0,185 war Allel 33 (270 bp) am stärksten vertreten.

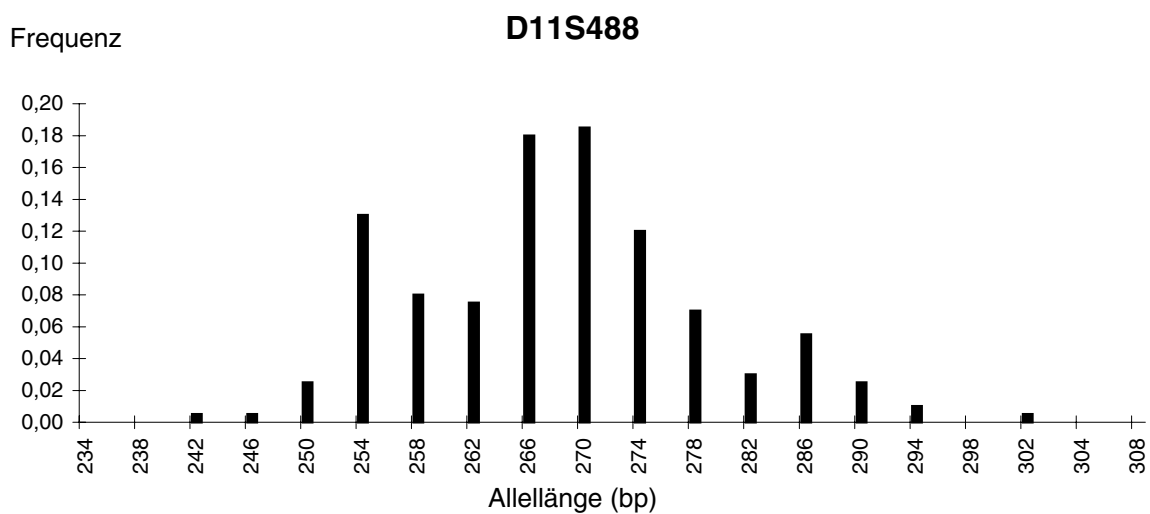


Abb. 3.6: Verteilung der Allelfrequenzen von D11S488

Von 120 möglichen Genotypen wurden 46 in den Untersuchungen beobachtet. Sie sind in Tabelle 3.2 zusammen mit der Anzahl an Beobachtungen (beob. (n)) aufgelistet. Die Namensgebung der Allele beruhte auf der Anzahl gefundener Repeats, wie sie aus der Sequenzierung (Kap. 3.3.1.2) zu ersehen war.

Tabelle 3.2: Allel- und Genotypfrequenzen von *D11S488*

Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)
<b>26-35</b>	1	<b>29-35</b>	4	<b>31-34</b>	1	<b>33-35</b>	2
<b>27-35</b>	1	<b>29-36</b>	1	<b>31-35</b>	2	<b>33-36</b>	1
<b>28-29</b>	1	<b>29-37</b>	3	<b>31-36</b>	3	<b>33-37</b>	2
<b>28-31</b>	1	<b>30-30</b>	1	<b>31-37</b>	1	<b>33-38</b>	3
<b>28-33</b>	2	<b>30-32</b>	5	<b>32-32</b>	6	<b>33-39</b>	1
<b>28-34</b>	1	<b>30-33</b>	1	<b>32-33</b>	4	<b>34-34</b>	1
<b>29-29</b>	2	<b>30-34</b>	3	<b>32-34</b>	5	<b>34-37</b>	2
<b>29-30</b>	2	<b>30-35</b>	1	<b>32-35</b>	2	<b>34-38</b>	1
<b>29-31</b>	2	<b>30-36</b>	1	<b>32-37</b>	2	<b>34-41</b>	1
<b>29-32</b>	4	<b>30-38</b>	1	<b>32-39</b>	1	<b>35-37</b>	1
<b>29-33</b>	2	<b>31-32</b>	1	<b>33-33</b>	5		
<b>29-34</b>	3	<b>31-33</b>	4	<b>33-34</b>	5		
Allel	Größe (bp)	beob. (n)	Frequenz	Allel	Größe (bp)	beob. (n)	Frequenz
<b>26</b>	242	1	0,005	<b>34</b>	274	24	0,120
<b>27</b>	246	1	0,005	<b>35</b>	278	14	0,070
<b>28</b>	250	5	0,025	<b>36</b>	282	6	0,030
<b>29</b>	254	26	0,130	<b>37</b>	286	11	0,055
<b>30</b>	258	16	0,080	<b>38</b>	290	5	0,025
<b>31</b>	262	15	0,075	<b>39</b>	294	2	0,010
<b>32</b>	266	36	0,180	<b>41</b>	302	1	0,005
<b>33</b>	270	37	0,185				

Grundvoraussetzung für den Einsatz genetischer Marker in gerichtlichen und vaterschaftlichen Gutachten ist die Testung auf Abweichung vom Hardy-Weinberg Gleichgewicht. Aufgrund der relativ geringen Populationsstichprobe und der im Gegensatz dazu großen Anzahl von Genotypen kam es vor, daß viele Genotypfrequenzen den Wert Null besaßen. Deshalb wurde das Hardy-Weinberg Gleichgewicht durch Formen von Allelgruppen getestet, die wie in Tabelle 3.3 ersichtlich, zusammengesetzt wurden. Eine signifikante Abweichung konnte nicht nachgewiesen werden.

Weitere statistische Parameter von vaterschafts- und gerichtsmedizinischem Interesse sind in Tabelle 3.4 wiedergegeben.

Tabelle 3.3: Chi-Quadratstest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D11S488

Allelgruppe	Allele
I	26-30
II	31-32
III	33
IV	34-41
Chi <sup>2</sup> = 6,279	
df = 9	
P = 0,71	

Tabelle 3.4: weitere statistische Parameter für D11S488

HR	GeneD	PIC	DI	pM	AVACH
0,85	0,880	0,868	0,969	0,031	0,759

HR: Heterozygotenrate

DI: discrimination index

GeneD: gene diversity

pM: probability of match

PIC: polymorphic information content

AVACH: Allgemeine-Vaterschafts-Ausschluß-Chance

### 3.3.1.2 Sequenzstruktur der Allele

Um die Sequenzstruktur der Allele zu bestimmen, wurden 36 Fragmente sequenziert. Nach Möglichkeit war jede gefundene Allellänge mit mindestens einer Sequenzierungsprobe vertreten. In den Untersuchungen wurden weitestgehend homozygote Proben verwendet. Als Ergänzung dienten heterozygote Amplifikate, deren Allellängen um mindestens 12 bp differierten. Nach dem Auftragen auf ein 3 %iges Agarosegel wurden die Fragmente elektrophoretisch aufgetrennt, aus dem Gel herausgeschnitten und für jedes Allel getrennt weiterverarbeitet. Die Sequenzierung erfolgte in beide Richtungen. In Abbildung 3.7 sind die gefundenen Sequenzvariationen zusammengefaßt.

Die Primer sind unterstrichen dargestellt, der Repeatbereich durch Fettdruck hervorgehoben.

Die tiefgestellten Zahlen geben das Intervall an, in dem die Anzahl der Sequenzen variierte.

Der D11S488 Locus bestand aus Blöcken von **AAAG** und **GAAG** Repeats, deren Häufigkeit sich in den einzelnen Allelen unterschied. Diese Variationen traten nur in kompletten Tetrarepeats auf. Mit dem IUB-Code (siehe Tabelle im Anhang) ließ sich der *compound*

*repeat* vereinfacht als YAAG schreiben. Die Summe dieser Repeats ergab die jeweilige Allelbezeichnung. Allel 33 beispielsweise besaß folgende Repeateinheit: (AAAG)<sub>12</sub> (GAAG)<sub>10</sub> GAAAG (GAAG)<sub>5</sub> GAAAGAAAGGAAAG (GAAG)<sub>6</sub>.

Der Repeatbereich ließ sich in zwei variable Regionen VR I und VR II aufteilen (Abb. 3.7), die durch eine konstante GAAAG Sequenz voneinander getrennt wurden. Zahlreiche Variationen in der Häufigkeit des (AAAG)<sub>m</sub> (GAAG)<sub>n</sub> Repeats der 5´Region (VR I) waren vornehmlich für die Komplexität des Systems verantwortlich. Im Gegensatz dazu wies das 3´Ende (VR II) nur eine geringe Variabilität mit zwei GAAG-Einheiten auf, die sich zwischen 4- und 6mal wiederholten.

AGGCAATAGAGACCCTGTG AAAGAAGA (AAAG)<sub>5</sub> GTAAGAA **(AAAG)**<sub>7-17</sub>  
**(GAAG)**<sub>4-15</sub> GAAAG **(GAAG)**<sub>4-5</sub> **GAAAGAAAGG** AAAG **(GAAG)**<sub>4-6</sub> GAAAG  
 (AAAG)<sub>2</sub> AGAGAGGAAG (AAAG)<sub>2</sub> GAAGGAAACTAA CTCTCAGTAACAAT  
TCATCATC



Abb. 3.7: D11S488 mit seinen variablen Regionen VR I und VR II (P1 und P2 sind die entsprechenden Primer)

### 3.3.1.3 Neue Allele

Im Vergleich mit bisherigen Publikationen zu D11S488 [Browne et al. 1993] waren innerhalb der vorliegenden Untersuchung zwei Allele beobachtet worden, die noch nicht von anderen Autoren beschrieben wurden. Es handelt sich hierbei um die Allele 38 und 41, die insgesamt bei sechs untersuchten Individuen auftraten. Sie befinden sich am Ende des Allelbereichs und gehören zu den größten Allelen des Locus.

Beim Ausrichten der untersuchten Sequenz mit der GenBank [accession no. L04732] wurde ein DNA-Bereich gefunden, der nicht zuzuordnen ist. Allerdings entspricht er exakt einer vorangegangenen Sequenz. Im Anhang ist der entsprechende Allel-Ausschnitt wiedergegeben. Aufgrund der „Verdopplung“ enthält die GenBank ein zu langes Allel, das in seiner Größe nicht den Untersuchungsergebnissen entspricht.

### 3.3.1.4 Varianten

Von 36 sequenzierten Fragmenten des Locus D11S488 wurden 14 Allelgruppen mit insgesamt 29 verschiedenen Allelen gefunden. In Tabelle 3.5 werden die untersuchten DNA-Sequenzen für die beiden variablen Regionen VR I und VR II zusammengestellt. Zu jedem Fragment ist angegeben, wie häufig es sequenziert wurde. Bei der Sequenzanalyse der Proben waren Beobachtungen von Allelen gemacht worden, die bei gleicher Länge an Basenpaaren eine unterschiedliche Sequenz aufwiesen. Mit der Fragmentanalyse war eine Unterscheidung dieser Allele nicht möglich, sie wurden daher in einer Gruppe zusammengefaßt.

### 3.3.2 HUMD18S51

In Abbildung 3.8 ist das Elektropherogramm für eine heterozygote Probe des Locus D18S51 dargestellt, wie es Grundlage für die Auswertung bei der Fragmentanalyse war. Jeder Peak entspricht einer Allellänge. Die Höhe des Ausschlags korreliert mit der Fluoreszenzintensität des markierten DNA-Fragments. Aus den Scan-Nummern der Abszisse wurde die Basenpaarlänge ermittelt.

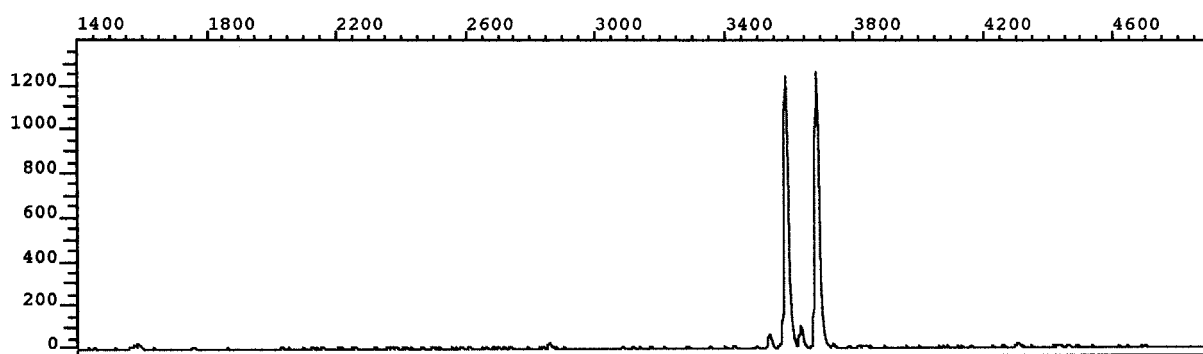


Abb. 3.8: Darstellung der DNA-Fragmente von D18S51 als Elektropherogramm

Tabelle 3.5: Allele mit DNA-Sequenzen der Repeatregion von *D11S488*

Allel	bp	VR I		VR II	seq. Allele
26	242	(AAAG) <sub>14</sub> (GAAG) <sub>4</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>4</sub>	1
27	246	(AAAG) <sub>15</sub> (GAAG) <sub>4</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>4</sub>	1
28	250	(AAAG) <sub>12</sub> (GAAG) <sub>7</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>5</sub>	1
29	254	(AAAG) <sub>7</sub> (GAAG) <sub>11</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	2
		(AAAG) <sub>8</sub> (GAAG) <sub>10</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
30	258	(AAAG) <sub>7</sub> (GAAG) <sub>12</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>14</sub> (GAAG) <sub>7</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>5</sub>	1
31	262	(AAAG) <sub>7</sub> (GAAG) <sub>13</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>10</sub> (GAAG) <sub>11</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>5</sub>	1
32	266	(AAAG) <sub>7</sub> (GAAG) <sub>14</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>10</sub> (GAAG) <sub>11</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	2
		(AAAG) <sub>11</sub> (GAAG) <sub>10</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>12</sub> (GAAG) <sub>9</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>13</sub> (GAAG) <sub>8</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>14</sub> (GAAG) <sub>7</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>15</sub> (GAAG) <sub>6</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>15</sub> (GAAG) <sub>7</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
33	270	(AAAG) <sub>16</sub> (GAAG) <sub>6</sub>	GAAAG	(GAAG) <sub>4</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>11</sub> (GAAG) <sub>11</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	2
		(AAAG) <sub>12</sub> (GAAG) <sub>10</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	2
		(AAAG) <sub>13</sub> (GAAG) <sub>9</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	4
34	274	(AAAG) <sub>14</sub> (GAAG) <sub>8</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
		(AAAG) <sub>13</sub> (GAAG) <sub>10</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
35	278	(AAAG) <sub>14</sub> (GAAG) <sub>9</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
36	282	(AAAG) <sub>15</sub> (GAAG) <sub>9</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
38	290	(AAAG) <sub>15</sub> (GAAG) <sub>10</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
39	294	(AAAG) <sub>16</sub> (GAAG) <sub>11</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
41	302	(AAAG) <sub>13</sub> (GAAG) <sub>15</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAAGGAAAG (GAAG) <sub>6</sub>	1
41	302	(AAAG) <sub>17</sub> (GAAG) <sub>13</sub>	GAAAG	(GAAG) <sub>5</sub> GAAAGAAA GGAAAG (GAAG) <sub>6</sub>	1

### 3.3.2.1 Allelfrequenzen und Genotypen

Wie aus Abbildung 3.9 zu ersehen ist, ergab die Untersuchung des Locus D18S51 eine regelmäßige Verteilung von 12 Allelen über einen Bereich von 278 bp bis 322 bp Länge. Die aufeinanderfolgenden Allele 12 (282 bp) bis 16 (298 bp) wiesen durch ihre ähnlich hohe Frequenzhäufigkeit ein Plateau auf. Sie bestimmten weitestgehend das Bild in der Population. Die Fragmentlängen unterschieden sich durch einen Größenzuwachs von jeweils vier Basenpaaren.

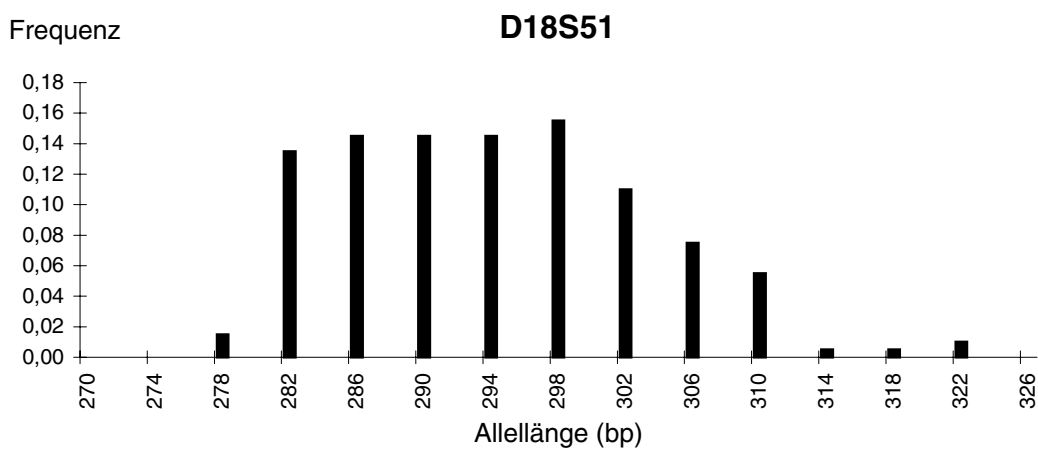


Abb. 3.9: Verteilung der Allelfrequenzen von D18S51

In Tabelle 3.6 sind die gefundenen Genotypen und die Anzahl ihrer Beobachtungen zusammengefaßt. Von 78 denkbaren Genotypkombinationen traten 38 bei den Untersuchungen auf. Die Namensgebung der Allele erfolgte nach der Anzahl gefundener Repeat-einheiten, auf sie wird in Kap. 3.3.2.2 näher eingegangen.



Tabelle 3.6: Allel- und Genotypfrequenzen von D18S51

Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)
11-13	1	13-14	3	14-17	4	16-17	2
11-15	1	13-15	2	14-18	2	16-18	3
11-17	1	13-16	9	14-19	1	16-19	2
12-12	2	13-17	1	15-15	2	17-17	1
12-13	4	13-18	3	15-16	4	17-18	3
12-14	12	13-19	1	15-17	4	17-19	2
12-15	2	13-21	1	15-18	3	17-20	1
12-16	4	14-14	1	15-19	5	18-22	1
12-17	2	14-15	2	15-22	1		
13-13	2	14-16	3	16-16	2		
Allel	Größe (bp)	beob. (n)	Frequenz	Allel	Größe (bp)	beob. (n)	Frequenz
<b>11</b>	278	3	0,015	<b>17</b>	302	22	0,110
<b>12</b>	282	27	0,135	<b>18</b>	306	15	0,075
<b>13</b>	286	29	0,145	<b>19</b>	310	11	0,055
<b>14</b>	290	29	0,145	<b>20</b>	314	1	0,005
<b>15</b>	294	29	0,145	<b>21</b>	318	1	0,005
<b>16</b>	298	31	0,155	<b>22</b>	322	2	0,010

Bei der Testung des Hardy-Weinberg Gleichgewichts für dieses System wurde keine Abweichung festgestellt. Wie bereits unter Kap. 3.3.1.1 beschrieben sind zur Berechnung Allelgruppen gebildet worden. Das 3-Gruppen-Modell setzte sich für den D18S51 Locus, wie aus Tabelle 3.7 ersichtlich, zusammen.

Weitere Ergebnisse aus statistischen Berechnungen sind in Tabelle 3.8 zusammengefaßt.

Tabelle 3.7: Chi-Quadratstest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D18S51

Allelgruppe	Allele
I	11-13
II	14-16
III	17-22
Chi <sup>2</sup> = 3,203	
df = 5	
P = 0,67	

Tabelle 3.8: weitere statistische Parameter für D18S51

HR	GeneD	PIC	DI	pM	AVACH
0,91	0,874	0,860	0,956	0,044	0,743

HR: Heterozygotenrate

DI: discrimination index

GeneD: gene diversity

pM: probability of match

PIC: polymorphic information content

AVACH: Allgemeine-Vaterschafts-Ausschluß-Chance

### 3.3.2.2 Sequenzstruktur der Allele

Zur Sequenzierung des Locus D18S51 wurden 12 homozygote Proben analysiert. Unter Verwendung von biotinylierten Primern konnten mit Hilfe der Sequenase die entsprechenden Bereiche in beide Richtungen sequenziert werden. Wie aus dem Sequenzabschnitt (siehe unten) zu ersehen ist, setzte sich der Locus aus einem einfachen Repeat (*simple repeat*), welcher durch Fettdruck hervorgehoben ist, zusammen. Unterstrichen sind die Primersequenzen gekennzeichnet. Die einzelnen Allele unterschieden sich lediglich in der Anzahl des **AGAA** Repeatblocks. Im kleinsten der beobachteten Fragmente, dem Allel 11, wiederholte sich der Tetrarepeat 11mal (Abb. 3.10).

Größtes gefundenes Fragment war Allel 22 mit 322 bp. Die Untersuchungen hatten ergeben, daß der Repeatbereich, bestehend aus dem 4 bp langen Block, die einzige variable Region darstellte (Abb. 3.10). Das AGAA-Muster kam in den einzelnen Allelen zwischen 11- und 22mal vor.

```

GAGCCATGTT CATGCCACTG CCCTTCACTC TGAGTGACAA ATTGAGACCT
TGTCTC      (AGAA)11-22   AAAGAGAG AGGAAAGAAA GAGAAAAGA
AAAGAAATAG TAGCAACTGT TATTGTAAGA CATCTCCACA CACCAGAGAA
GTTAATTTTA ATTTAACAT  GTTAAGAACA GAGAGAAGCC AACATGTCCA
CCTTAGGCTG ACGGTTTGTT TATTTGTGTC GTTGCTGGTA GTCGGGTTTG
    
```

VR

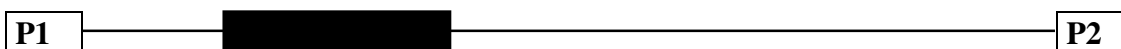


Abb. 3.10: D18S51 mit seiner variablen Region VR (P1 und P2 sind die Primerbereiche)

### 3.3.2.3 Neue Allele

Die Ergebnisse der vorliegenden Arbeit stimmen weitgehend mit anderen Publikationen über D18S51 [Straub et al. 1993, Urquhart et al. 1995] überein. Die mitunter beschriebenen 2er Repeats wurden bei keinem der 100 untersuchten Individuen beobachtet.

### 3.3.2.4 Varianten

Der Locus D18S51 setzte sich aus einem einfachen Tetrarepeat (AGAA)<sub>n</sub> zusammen. In der flankierenden Region waren keinerlei Varianten oder Mutationen beobachtet worden. Isoallele wurden ebenfalls nicht beobachtet.

### 3.3.3 HUMD19S246

In Abbildung 3.11 sind im Elektropherogramm die beiden Allellängen eines Heterozygoten für das System D19S246 dargestellt.

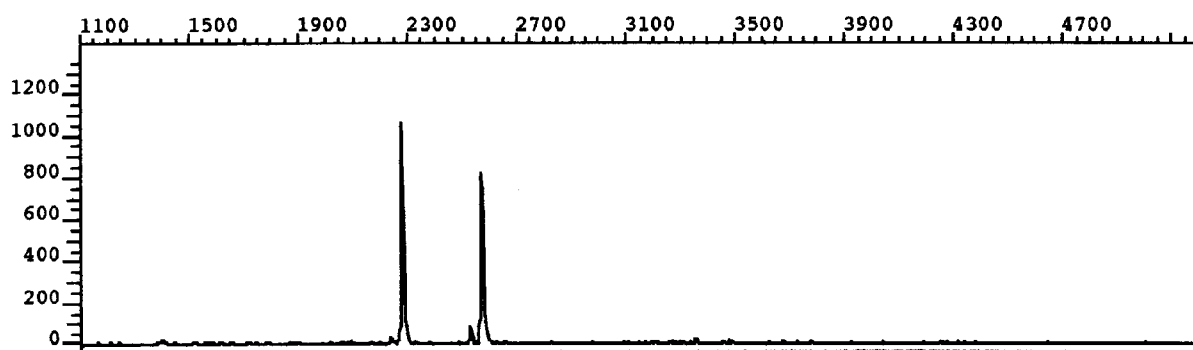


Abb. 3.11: Darstellung der DNA-Fragmente von D19S246 als Elektropherogramm. Die Peakhöhe ergibt sich durch die Fluoreszenzintensität. Auf der Abszisse sind Zahlenwerte für den Scanbereich aufgetragen, sie lassen sich durch Verrechnung mit dem Längenstandard in bp ausdrücken.

### 3.3.3.1 Allelfrequenzen und Genotypen

Bei der Analyse des Locus D19S246 wurden 11 verschiedene Allele gefunden. Diese lagen in einem Bereich zwischen 182 bp und 230 bp. Eine graphische Darstellung, in der die Frequenzen gegen die Allelverteilung in Basenpaaren (bp) aufgetragen sind, ist in der Abbildung 3.12 ersichtlich. Allel 29 mit 182 bp fiel dabei besonders auf. Es war das kleinste beobachtete Allel und zugleich weitaus am häufigsten mit einer Frequenz von 0,325 vertreten.

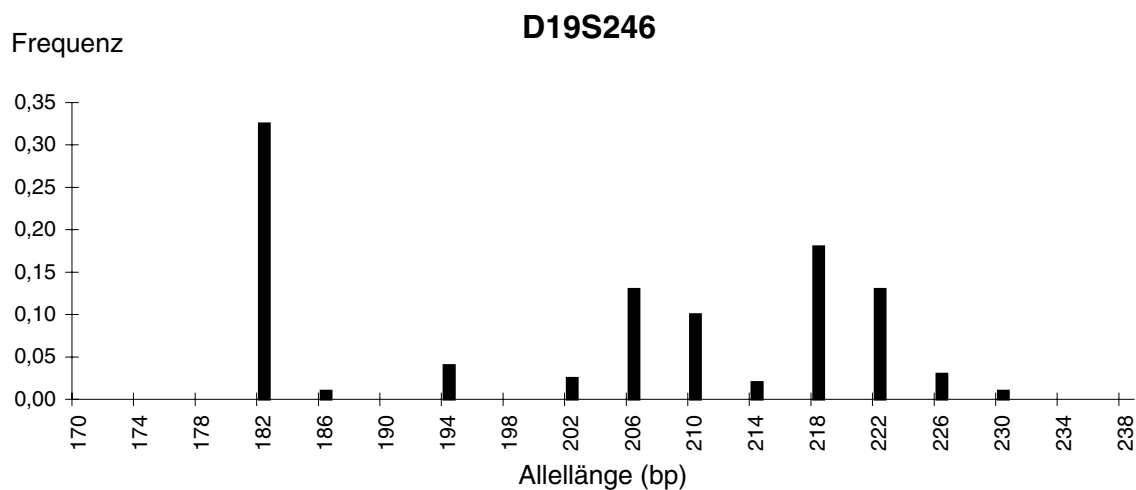


Abb. 3.12: Verteilung der Allelfrequenzen von D19S246

Die Auswertung ergab 31 unterschiedliche Genotypen bei 66 möglichen Allelkombinationen. Zusammen mit der Anzahl ihrer Beobachtungen (beob. (n)) sind sie in Tabelle 3.9 aufgelistet. Auf die Benennung der Allele wird in Kap. 3.3.3.2 näher eingegangen.

Wie bereits unter Kap. 3.3.1.1 beschrieben wurden zur Testung des Hardy-Weinberg Gleichgewichts Allelgruppen gebildet. Aus Tabelle 3.10 ist ersichtlich, daß keine signifikante Abweichung nachgewiesen werden konnte.

Weitere statistische Parameter sind in Tabelle 3.11 für den Locus D19S246 zusammengestellt.

Tabelle 3.9: Allel- und Genotypfrequenzen von D19S246

Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)
29-29	9	29-40	1	35-38	3	37-38	1
29-30	2	32-38	2	35-39	5	37-39	1
29-32	3	32-39	2	35-40	1	38-38	3
29-34	1	32-40	1	35-41	1	38-39	5
29-35	11	34-35	1	36-36	1	39-40	1
29-36	5	34-36	2	36-37	1	39-41	1
29-38	16	34-37	1	36-38	3	40-40	1
29-39	8	35-36	4	36-39	3		
Allel	Größe (bp)	beob. (n)	Frequenz	Allel	Größe (bp)	beob. (n)	Frequenz
29	182	65	0,325	37	214	4	0,020
30	186	2	0,010	38	218	36	0,180
32	194	8	0,040	39	222	26	0,130
34	202	5	0,025	40	226	6	0,030
35	206	26	0,130	41	230	2	0,010
36	210	20	0,100				

Tabelle 3.10: Chi-Quadratstest zur Abweichung vom Hardy-Weinberg Gleichgewicht für D19S246

Allelgruppe	Allele
I	29-30
II	32-36
III	37-41
Chi <sup>2</sup> = 0,165	
df	= 5
P	= 1,0

Tabelle 3.11: weitere statistische Parameter für D19S246

HR	GeneD	PIC	DI	pM	AVACH
0,86	0,185	0,793	0,931	0,069	0,645

HR: Heterozygotenrate

DI: discrimination index

GeneD: gene diversity

pM: probability of match

PIC: polymorphic information content

AVACH: Allgemeine-Vaterschafts-Ausschluß-Chance

### 3.3.3.2 Sequenzstruktur der Allele

Zur Untersuchung der Sequenzstruktur des Locus D19S246 wurden 37 Allele sequenziert. Wie bereits unter Kap. 3.3.1.2 beschrieben, waren auch hierbei Fragmente im Agarosegel aufgetrennt, herausgeschnitten und mittels Taq Cycle Sequencing einzeln bearbeitet worden. Verwendet wurden heterozygote Proben, deren Allele möglichst weit auseinander lagen, damit die Fragmente einfacher separiert werden konnten. Zur Sequenzanalyse sollte möglichst jede beobachtete Allellänge zur näheren Bestimmung herangezogen werden.

In Abbildung 3.13 sind die Primer unterstrichen dargestellt, mit Fettdruck der Repeatbereich hervorgehoben. Die tiefgestellten Zahlen geben die Anzahl der einzelnen Sequenzen an. Anhand der Intervalle läßt sich die Vielfalt möglicher Allelkombinationen erkennen.

Der Locus D19S246 bestand aus zahlreichen 4er Repeat Blöcken, von denen nur die **TGTA** und **TCTA** Sequenzen in bestimmten Abschnitten in ihrer Anzahl variierten. Mit dem IUB-Code (siehe Anhang) wurden die kompletten Tetranukleotide vereinfacht als TSTA beschrieben. Die Allelbenennung ergab sich aus der Summe aller TSTA Einheiten, die zwischen den Primern lagen.

Während innerhalb des Locus zahlreiche TGTA und TCTA Blöcke beobachtet wurden, so schienen nur drei von ihnen in unterschiedlicher Häufigkeit aufzutreten. Es ließ sich eine Einteilung in zwei variable Regionen VR I und VR II vornehmen (Abb. 3.13), wobei VR I gegenüber VR II eine deutlich geringere Variabilität aufwies.

GAAACACATCATTACCCAC TTCAA (TCTA)<sub>2</sub> TATAA (**TCTA**)<sub>2-3</sub> A (TCTA)<sub>2</sub>  
TCTG (TCTA)<sub>3</sub> (TGTA)<sub>1</sub> TCTG (**TGTA**)<sub>7-12</sub> (**TCTA**)<sub>5-13</sub> (TGTA)<sub>1</sub> (TCTA)<sub>4</sub>  
(TGTA)<sub>1</sub> (TCTA)<sub>1</sub> TATTTTTT GAAAGGTGGAATCTCACTCT



Abb. 3.13: D19S246 mit seinen variablen Regionen VR I und VR II (P1 und P2 sind die Primerbereiche)

### 3.3.3.3 Neue Allele

Bei der Untersuchung des Locus D19S246 wurde ein Allel beobachtet, das bisher in der Literatur noch keine Erwähnung fand. Mit einer Frequenz von 0,010 war das Allel 30 sehr selten und stellte in seiner Größe das zweitkleinste dar. Von der Struktur reihte es sich in das Tetrarepeatmuster ein.

### 3.3.3.4 Varianten

Von 37 sequenzierten Fragmenten wurden 10 Allelgruppen mit insgesamt 17 verschiedenen Allelen beobachtet. In Tabelle 3.12 sind diese unter ihren beiden variablen Regionen aufgelistet. Gleichzeitig erfolgt die Angabe, wie häufig die entsprechende Sequenz bei den Untersuchungen auftrat.

*Tabelle 3.12: Allele mit DNA-Sequenzen der Repeatregion von D19S246*

Allel	bp	VR I	VR II	seq. Allele
<b>29</b>	182	(CTAT) <sub>2</sub>	(GTAT) <sub>7</sub> (CTAT) <sub>5</sub>	13
<b>30</b>	186	(CTAT) <sub>2</sub>	(GTAT) <sub>7</sub> (CTAT) <sub>6</sub>	1
<b>32</b>	194	(CTAT) <sub>2</sub>	(GTAT) <sub>8</sub> (CTAT) <sub>7</sub>	2
<b>34</b>	202	(CTAT) <sub>2</sub>	(GTAT) <sub>8</sub> (CTAT) <sub>9</sub>	1
		(CTAT) <sub>2</sub>	(GTAT) <sub>10</sub> (CTAT) <sub>7</sub>	1
<b>35</b>	206	(CTAT) <sub>2</sub>	(GTAT) <sub>11</sub> (CTAT) <sub>7</sub>	4
		(CTAT) <sub>2</sub>	(GTAT) <sub>12</sub> (CTAT) <sub>6</sub>	1
<b>36</b>	210	(CTAT) <sub>2</sub>	(GTAT) <sub>9</sub> (CTAT) <sub>10</sub>	1
		(CTAT) <sub>2</sub>	(GTAT) <sub>12</sub> (CTAT) <sub>7</sub>	2
<b>37</b>	214	(CTAT) <sub>2</sub>	(GTAT) <sub>10</sub> (CTAT) <sub>10</sub>	1
<b>38</b>	218	(CTAT) <sub>3</sub>	(GTAT) <sub>8</sub> (CTAT) <sub>12</sub>	1
		(CTAT) <sub>2</sub>	(GTAT) <sub>8</sub> (CTAT) <sub>13</sub>	1
		(CTAT) <sub>2</sub>	(GTAT) <sub>9</sub> (CTAT) <sub>12</sub>	2
		(CTAT) <sub>2</sub>	(GTAT) <sub>10</sub> (CTAT) <sub>11</sub>	1
<b>39</b>	222	(CTAT) <sub>3</sub>	(GTAT) <sub>9</sub> (CTAT) <sub>12</sub>	2
<b>40</b>	226	(CTAT) <sub>3</sub>	(GTAT) <sub>9</sub> (CTAT) <sub>13</sub>	1
		(CTAT) <sub>3</sub>	(GTAT) <sub>10</sub> (CTAT) <sub>12</sub>	2

Es wurden Allele von gleicher Länge aber unterschiedlicher Sequenz beobachtet. Solche „Isoallele“ fanden sich bei Allel 34, 35, 36, 38 und 40. Sie ließen sich in der Fragmentanalyse nicht voneinander unterscheiden und wurden dort in einer Gruppe zusammengefaßt. Eine nähere Betrachtung der variablen Region VR I zeigte ein zwei- oder dreimaliges Vorkommen des TCTA Repeats, wobei die höhere Repeatzahl vorwiegend bei größeren Allelen beobachtet wurde.

### 3.3.4 HUMFIBRA

In Abbildung 3.14 ist das Elektropherogramm eines Individuums für den Locus HUMFIBRA (Fibrinogen alpha) dargestellt. Mit Hilfe der Fragmentanalyse konnten auf diese Weise die Allele einer DNA-Probe bestimmt werden. Die Abszisse zeigt den Scanbereich und die Ordinate Fluoreszenzeinheiten, in welchen die Peakhöhe angegeben wurde.

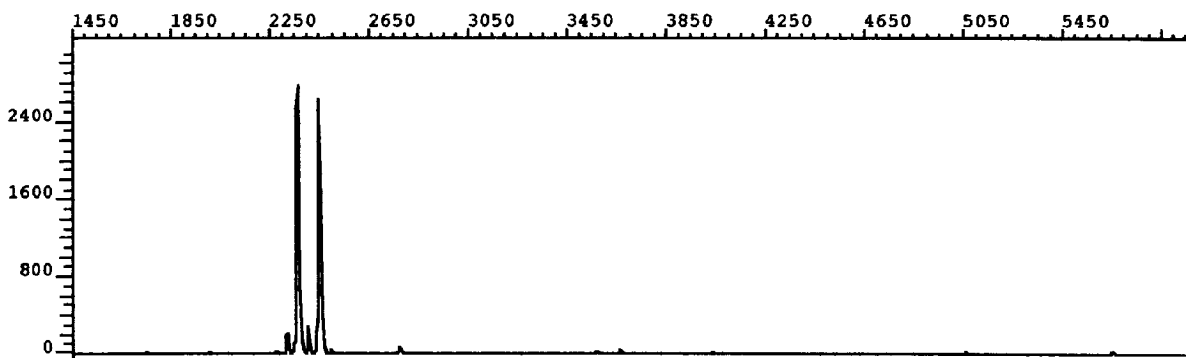


Abb. 3.14: Darstellung der DNA-Fragmente von HUMFIBRA als Elektropherogramm

#### 3.3.4.1 Allelfrequenzen und Genotypen

Bei der Bestimmung der Allele für den Locus HUMFIBRA wurden 10 Fragmentlängen von 180 bp bis 212 bp beobachtet. Abbildung 3.15 veranschaulicht die Allelverteilung im Zusammenhang mit der Frequenz. Es fällt auf, daß neben den regelmäßigen Abständen von 4 bp Differenz ein Allel mit nur 2 bp Abstand vorkam und so den Tetrarepeat unterbrach. Häufigstes Allel war Allel 16 (192 bp) mit einer Frequenz von 0,245.



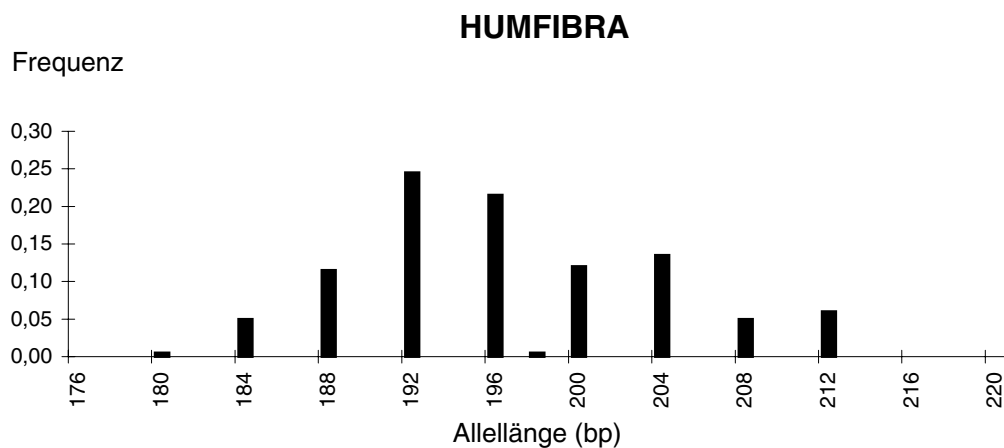


Abb. 3.15: Verteilung der Allelfrequenzen von HUMFIBRA

Von 55 möglichen Genotypen wurden 29 beobachtet, die in Tabelle 3.13 gemeinsam mit der Beobachtungsanzahl (beob. (n)) aufgelistet sind. Die Benennung der Allele richtete sich nach der Anzahl gefundener Repeats, wobei Allele, die außerhalb des 4er Repeats lagen, mit einem Dezimalpunkt geschrieben wurden, dem die Anzahl der Basenpaare folgte, um welche das Allel länger war.

Tabelle 3.13: Allel- und Genotypfrequenzen von HUMFIBRA

Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)
<b>13-20</b>	1	<b>15-17</b>	9	<b>16-20</b>	2	<b>18-20</b>	1
<b>14-16</b>	3	<b>15-18</b>	2	<b>16-21</b>	5	<b>18-21</b>	1
<b>14-18</b>	1	<b>15-19</b>	5	<b>17-17</b>	5	<b>19-20</b>	1
<b>14-19</b>	2	<b>15-20</b>	1	<b>17-18</b>	6	<b>19-21</b>	2
<b>14-20</b>	3	<b>16-16</b>	5	<b>17-19</b>	6	<b>20-21</b>	1
<b>14-21</b>	1	<b>16-17</b>	10	<b>17-21</b>	2		
<b>15-15</b>	2	<b>16-18</b>	9	<b>17.2-19</b>	1		
<b>15-16</b>	3	<b>16-19</b>	7	<b>18-19</b>	3		
Allel	Größe (bp)	beob. (n)	Frequenz	Allel	Größe (bp)	beob. (n)	Frequenz
<b>13</b>	180	1	0,005	<b>17.2</b>	198	1	0,005
<b>14</b>	184	10	0,050	<b>18</b>	200	24	0,120
<b>15</b>	188	23	0,115	<b>19</b>	204	27	0,135
<b>16</b>	192	49	0,245	<b>20</b>	208	10	0,050
<b>17</b>	196	43	0,215	<b>21</b>	212	12	0,060

Bei der Testung auf Abweichung vom Hardy-Weinberg Gleichgewicht wurden die Allele zu Gruppen zusammengefaßt (Tabelle 3.14). Dies war nötig, da einige Genotypfrequenzen durch die relativ geringe Populationsstichprobe den Wert Null annahmen. Eine Berechnung nach dieser Methode ist zwar nicht exakt, stellte aber eine gute Näherung für den Chi-Quadratstest dar. Eine signifikante Abweichung konnte für HUMFIBRA nicht nachgewiesen werden.

Weitere statistische Parameter sind in Tabelle 3.15 aufgeführt.

*Tabelle 3.14: Chi-Quadratstest zur Abweichung vom Hardy-Weinberg Gleichgewicht für HUMFIBRA*

Allelgruppe	Allele
I	13-16
II	17
III	17.2-21
Chi <sup>2</sup> = 5,717	
df	= 5
P	= 0,33

*Tabelle 3.15: weitere statistische Parameter für HUMFIBRA*

HR	GeneD	PIC	DI	pM	AVACH
0,89	0,839	0,820	0,945	0,055	0,681

HR: Heterozygotenrate

DI: discrimination index

GeneD: gene diversity

pM: probability of match

PIC: polymorphic information content

AVACH: Allgemeine-Vaterschafts-Ausschluß-Chance

### 3.3.4.2 Sequenzstruktur der Allele

Um die Struktur der Allele zu bestimmen, wurden 19 Fragmente sequenziert. Jede Allellänge wurde mindestens einmal, in der Regel zwei- bis dreimal, untersucht. Sowohl homozygote als auch heterozygote Proben wurden für die Analyse herangezogen.

Der Repeat bestand beim HUMFIBRA Locus aus **TCTT** Blöcken, die sich in einem bestimmten Bereich zwischen 10- und 18mal wiederholten, und weiteren TC reichen Sequenzen wie sie in Abb. 3.16 ersichtlich sind. Für die variable Region (VR) war der TCTT Tetrarepeat verantwortlich. Verlängerungen oder Verkürzungen der Allele waren auf das

Hinzutreten oder Verschwinden dieses 4er Blocks zurückzuführen. Diese Eigenschaft kennzeichnet einen einfachen oder sogenannten *simple repeat*. Entsprechend erfolgte die Namensgebung der Allele, indem die Anzahl der TCTT Repeats in der variablen Region addiert wurde. Davon abweichende 2er Repeats wurden mit einem Dezimalpunkt und der Anzahl Basenpaare, um die das Allel verlängert war, beziffert wie das Allel 17.2. Aufgrund der geringen Basenpaardifferenz zwischen den beiden Allelen (196 bp und 204 bp) des heterozygoten Trägers mit dem Allel 17.2 war eine Trennung mittels QIAEX II nicht möglich.

```

GCCCCATAGG TTTTGAACTC ACAGATTAAA CTGTAACCAA AATAAAATTA
GGCATATTTA CAAGCTAGTT (TCTT)3 TTTTC (TCTT)10-18
TC(CCTT)3 TCTTCCTTTC TTTTTTGCTG GCAATTACAG ACAAATCA

```



Abb. 3.16: HUMFIBRA mit seiner variablen Region VR (P1 und P2 sind die Primerbereiche). In der Sequenz ist die Primerregion durch Unterstreichen und der Repeatbereich mittels Fettdruck kenntlich gemacht.

### 3.3.4.3 Neue Allele

Der Vergleich mit bisherigen Publikationen bestätigte, daß neben einem regelmäßigen Tetrarepeat bei diesem Locus auch 2er Repeats gefunden wurden, welche sich zwischen die 4er Repeats plazieren. Während andere Autoren von mehreren 2er Repeats berichteten, so war in der vorliegenden Untersuchung nur ein Dinukleotid beobachtet worden. Es handelte sich um das Allel 17.2, das allerdings von anderen Autoren noch nicht beschrieben wurde. Auffällig ist, daß Dinukleotide meist nur mit geringen Frequenzen zwischen den Tetrarepeats auftreten.

Bei einem Vergleich der eigenen Sequenz mit der GenBank [accession no. M64982] wurde keinerlei Abweichung festgestellt.

### 3.3.4.4 Varianten

Keines der sequenzierten Allele hatte eine Abweichung von der Grundstruktur gezeigt. Ein Vorkommen von Isoallelen konnte daher nicht beobachtet werden.

### 3.3.5 HUMVWFA31/A

In der Fragmentanalyse wurden für jedes Individuum der Populationsstichprobe die zugehörigen Allele bestimmt. Abbildung 3.17 zeigt das Elektropherogramm einer Probe für den Locus HUMVWFA31/A. Die beiden Spitzen (Peaks) stellen die Fragmente eines Heterozygoten mit zwei unterschiedlichen Allellängen dar, welche durch die beiden Peaks repräsentiert werden. Den Werten auf der Abszisse für den Scan-Bereich konnten entsprechende Basenpaarlängen zugeordnet werden. Die Peakhöhe (Ordinate) richtete sich nach der Fluoreszenzintensität der Fragmente, die durch den Primer farbmarkiert waren.

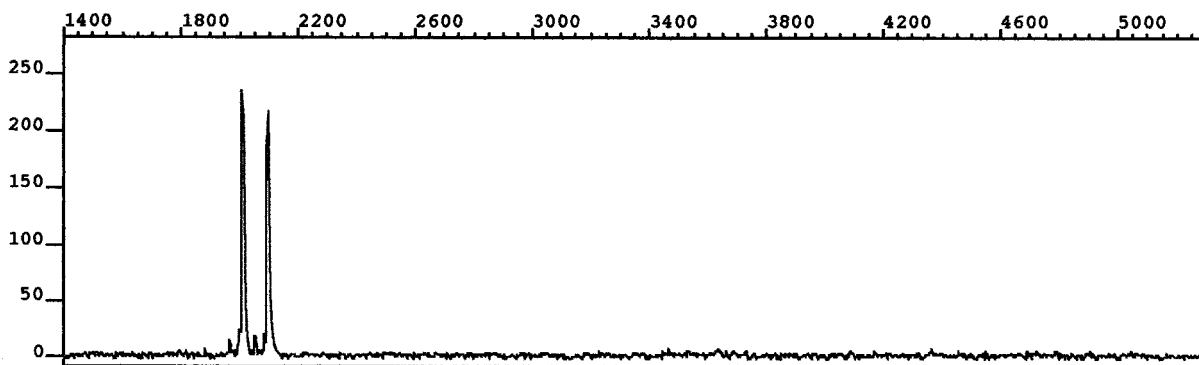


Abb. 3.17: Darstellung der DNA-Fragmente von HUMVWFA31/A als Elektropherogramm

#### 3.3.5.1 Allelfrequenzen und Genotypen

Die Analyse des Locus HUMVWFA31/A ergab eine nahezu symmetrische Verteilung von 8 Allelen, die zwischen 134 bp und 162 bp lagen. In Abbildung 3.18 sind die beobachteten Allelfrequenzen graphisch dargestellt. Ihre Auftragung auf der Abszisse erfolgte entsprechend der Größe in Basenpaaren. Durch die Höhe der Säulen wurde die Frequenzhäufigkeit wiedergegeben. Die höchste Frequenz von 0,270 wies Allel 17 mit 150 bp auf.

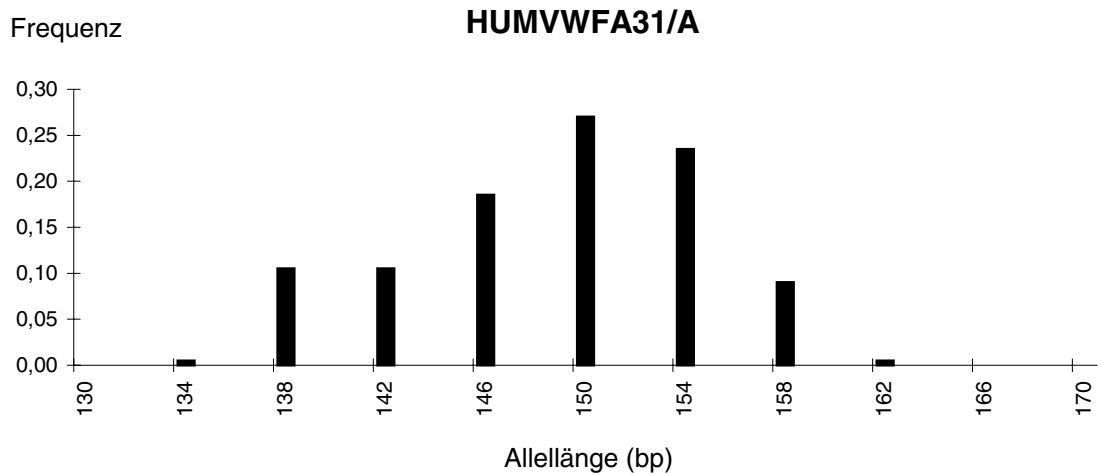


Abb. 3.18: Verteilung der Allelfrequenzen von HUMVWFA31/A

Von 36 möglichen Genotypen wurden 20 beobachtet, die in Tabelle 3.16 zusammen mit den gezählten Beobachtungen (beob. (n)) aufgeführt sind. Die Benennung der Allele erfolgte nach deren Repeatanzahl.

Tabelle 3.16: Allel- und Genotypfrequenzen von HUMVWFA31/A

Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)	Genotyp	beob. (n)
<b>13-17</b>	1	<b>14-19</b>	2	<b>15-20</b>	1	<b>17-17</b>	11
<b>14-14</b>	1	<b>15-16</b>	4	<b>16-16</b>	5	<b>17-18</b>	8
<b>14-16</b>	5	<b>15-17</b>	9	<b>16-17</b>	7	<b>17-19</b>	4
<b>14-17</b>	3	<b>15-18</b>	4	<b>16-18</b>	8	<b>18-18</b>	6
<b>14-18</b>	9	<b>15-19</b>	3	<b>16-19</b>	3	<b>18-19</b>	6
Allel	Größe (bp)	beob. (n)	Frequenz	Allel	Größe (bp)	beob. (n)	Frequenz
<b>13</b>	134	1	0,005	<b>17</b>	150	54	0,270
<b>14</b>	138	21	0,105	<b>18</b>	154	47	0,235
<b>15</b>	142	21	0,105	<b>19</b>	158	18	0,090
<b>16</b>	146	37	0,185	<b>20</b>	162	1	0,050

Für den Einsatz genetischer Marker in Vaterschaftsgutachten und in der forensischen Medizin ist es nötig, deren statistischen Werte zu berechnen und zu prüfen. Zur Testung des Hardy-Weinberg Gleichgewichts wurden Allelgruppen gebildet, wie sie in Tabelle 3.17 aufgeführt sind. Eine signifikante Abweichung für dieses System konnte nicht nachgewiesen werden.

Weitere statistische Parameter sind in Tabelle 3.18 wiedergegeben.

*Tabelle 3.17: Chi-Quadratstest zur Abweichung vom Hardy-Weinberg Gleichgewicht für HUMVWFA31/A*

Allelgruppe	Allele
I	13-16
II	17
III	18-20
Chi <sup>2</sup> = 4,548	
df = 5	
P = 0,47	

*Tabelle 3.18: weitere statistische Parameter für HUMVWFA31/A*

HR	GeneD	PIC	DI	pM	AVACH
0,77	0,808	0,780	0,934	0,066	0,620

HR: Heterozygotenrate

DI: discrimination index

GeneD: gene diversity

pM: probability of match

PIC: polymorphic information content

AVACH: Allgemeine-Vaterschafts-Ausschluß-Chance

### 3.3.5.2 Sequenzstruktur der Allele

Die Sequenzstruktur wurde nicht untersucht, da zum HUMVWFA31/A Locus bereits eingehende Ergebnisse anderer Autoren vorlagen. Der *compound repeat* [Urquhart et al. 1994] setzt sich nach der Beschreibung von Urquhart folgendermaßen zusammen: (ATCT)<sub>2</sub> (GTCT)<sub>3-4</sub> (ATCT)<sub>9-13</sub>. Unter Verwendung des IUB-Codes ergibt sich daraus ein RTCT-Repeat. Auszüge aus der GenBank [accession no. M25858] sind im Anhang wiedergegeben.

### **3.3.5.3 Neue Allele**

Im Vergleich mit bisherigen Publikationen anderer Autoren wurden in der untersuchten Populationsstichprobe des HUMVWFA31/A Locus keine neuen Allele beobachtet. Das Auftreten von 2er Repeats, wie sie für diesen Marker in der Literatur beschrieben wurden, konnte in der vorliegenden Studie nicht bestätigt werden.





## 4 Diskussion

### PCR-Bedingungen und stutter peaks

Die Analyse von *short tandem repeat* (STR) Sequenzen mittels Polymerasekettenreaktion (PCR) [Weber u. May 1989, Tautz 1989, Litt u. Luty 1989, Edwards et al. 1991] gehört derzeit zur am meisten verwendeten Methode in der forensischen Personenidentifikation [Edwards et al. 1992, Kimpton et al. 1992, Brinkmann 1992]. Da von den STRs die Dinukleotidrepeats am weitesten im menschlichen Genom verbreitet sind, werden sie vornehmlich zur Erstellung von Genkarten eingesetzt. Untersuchungen haben ergeben, daß Dinukleotide während der Amplifikation häufig zu einem enzymatischen Slippage neigen [Kimpton et al. 1993], welcher sich in der Fragmentanalyse als sogenannte artifizielle *stutter peaks* darstellt. Aufgrund dieser Eigenschaft sind Dinukleotide für einen Einsatz in der Forensic nur bedingt geeignet [Hauge u. Litt 1993]. Tetranukleotidrepeats dagegen zeigen eine bedeutend geringere Anfälligkeit für den Slippage und werden daher bevorzugt in der Personenidentifikation eingesetzt [Edwards et al. 1991].

Untersuchungsgegenstand der vorliegenden Arbeit waren die fünf Tetranukleotid DNA-Marker D11S488, D18S51, D19S246, HUMFIBRA und HUMVWFA31/A. Für sie sollte die Eignung im Einsatz zur Identifikation von Individuen und Vaterschaftbegutachtung geprüft werden. PCR-Amplifikate, die von Tetranukleotidrepeat-Loci resultieren, produzieren typischerweise einen einzelnen kleinen (<10 %) *stutter peak*, der in der Regel um 1 - 4 bp kürzer als das entsprechende Hauptallel ist. Seine Interpretation ist sehr viel einfacher als die der Dinukleotidrepeats [Urquhart et al. 1995]. Das Auftreten von *stutter peaks* wurde bei etwa 5 % der untersuchten Proben beobachtet. Als wahrscheinlichste Ursache wird die inkomplette Addition einer Extrabase durch die Taq Polymerase am Ende der Extension gesehen [Clark 1988, Kimpton et al. 1993]. Sequenzuntersuchungen von Walsh et al. [1996] haben gezeigt, daß den *stutter peaks* im Vergleich zum Hauptallel auch eine komplette Repeateinheit fehlen kann. Welche Repeateinheit dabei verloren geht, kann von Allel zu Allel unterschiedlich sein. Die Größe des *stutter peaks* in Relation zum Hauptallel steigt mit der Anzahl ununterbrochener Hauptrepeateinheiten. Bei der Fragmentanalyse der in dieser Arbeit untersuchten Systeme wurden ebenfalls *stutter peaks* beobachtet. Mit einer Peakhöhe, die deutlich unter 10 % im Vergleich zum Hauptallel lag, kam es zu keinerlei Schwierigkeiten bei der Allelbestimmung.

In Vorversuchen wurden für jedes System unter Austestung geeigneter Annealing-temperaturen und Primerkonzentrationen PCR-Reaktionsbedingungen und Amplifikationsbedingungen optimiert.

### **Vorteile einer Multiplexanalyse**

Mit Hilfe von PCR-Amplifikationen können DNA-Profile erstellt werden. Dabei besitzen STRs eine größere Sensitivität als andere Systeme. Selbst bei schlecht gelagerten Proben oder bereits degradierter DNA ist eine PCR dieser Loci aufgrund ihrer geringen Allelgrößen (meist <300 bp) noch erfolgreich [Hagelberg et al. 1991, Gill et al. 1992, Jeffreys et al. 1992, Gill et al. 1993, Schlotterer u. Tautz 1992]. Unter einfachen Reaktionsbedingungen können mehrere Marker gemeinsam in einem Reaktionsgefäß als sogenannter Multiplex amplifiziert werden [Edwards et al. 1991, Morral u. Estivill 1992, Schwartz et al. 1992]. Durch geschickte Auswahl von Systemen, die unter gleichen Reaktionsbedingungen laufen und deren Primer bei übereinstimmenden Allelgrößen unterschiedlich farbmarkiert sind, können mehrere Genorte gleichzeitig in einem Multiplex erfaßt werden. Dies spart nicht nur zahlreiche Arbeitsschritte sondern auch DNA, denn gerade bei wenig Probenmaterial kann eine Multiplexamplifikation hochgradig diskriminierende Einzelamplifikationen ersetzen. Besonders im forensischen Bereich ist diese Methode, mit der vornehmlich eine Beschleunigung des Analyseprozesses erreicht werden soll, sehr etabliert [Gill et al. 1996]. In Verbindung mit der direkten Bestimmung amplifizierter Produkte über das Polyacrylamidgel wird die DNA-Profilierung der Automatisierung zugänglich gemacht [Kimpton et al. 1993]. Um mehrere Systeme in einem Multiplex zu vereinen, müssen leichte Veränderungen in den individuellen PCR-Bedingungen hingenommen werden, was zwangsläufig die Leistungsfähigkeit in der Amplifikation geringgradig herabsetzt. Dies stellt bei ausreichender DNA-Menge und Qualität jedoch keine Schwierigkeit dar. Allerdings sind zur Identifizierung schwieriger Proben in der Bearbeitung gerichtlicher Fälle Single-Locus Amplifikationen erforderlich. Damit sind *singleplexing* und *multiplexing* sich nicht gegenseitig ausschließende Techniken [Gill et al. 1996].

In der vorliegenden Arbeit wurden zwei Multiplex Amplifikationssysteme (Triplex und Quadruplex) entwickelt, die in Verbindung mit dem automatischen fluoreszenzmarkierten Detektionssystem geeignet sind, gemeinsam fünf hoch polymorphe Tetranukleotid-STR-Loci zu amplifizieren. Da die Systeme D19S246 und HUMFIBRA überschneidende Allellängenbereiche besitzen konnten sie nicht in einem Multiplex zusammengefaßt werden. Durch

Angleichung der Annealingtemperaturen (54 °C) und der individuellen Primerkonzentrationen (D11S488, D18S51, D19S246, HUMFIBRA je 10 pmol; HUMVWFA31/A mit 25 pmol) konnten leistungsfähige Amplifikationen von allen Loci im Multiplexsystem erreicht werden. Bei geringeren Primerkonzentrationen wurde eine Reduktion der Signalstärke beobachtet, was bei Veränderungen in den Reaktionsbedingungen mitunter zum Allelverlust führte. Im Gegensatz zu den STRs benötigen VNTR-Loci (bzw. AmpFLP's; Amplifikations-Fragment-Längen-Polymorphismen) für eine erfolgreiche Amplifikation häufig spezielle Reaktionsbedingungen, die sich von Marker zu Marker stark unterscheiden. Daher gestaltet sich eine gemeinsame Amplifikation von VNTRs (AmpFLP's) als sehr schwierig [Sullivan et al. 1992, Tully et al. 1993]. Zudem kann eine Amplifikation mittels PCR lediglich bei VNTR-Loci erfolgen, die eine relativ kurze Allellänge aufweisen, sog. AmpFLP's.

Die nahe beieinander liegenden Allelbereiche der einzelnen STR-Loci erlauben es, eine größere Anzahl von Markern gemeinsam in einer Spur bzw. auf einem Gel laufen zu lassen. Eine weitere Erleichterung bietet der Gebrauch unterschiedlicher Fluoreszenzfarben für Loci mit überlappenden Allelgrößen. Hier wäre in weiteren Untersuchungen zu prüfen, ob alle fünf Systeme zu einem Pentaplex zusammengefaßt werden können, wenn der Primer von HUMFIBRA anstatt von TAMRA mit dem Farbstoff FAM gekoppelt wird. Schließlich haben sich alle fünf Systeme bei den Untersuchungen als multiplexfähig erwiesen. Die Möglichkeit, vier Fluoreszenzfarben voneinander zu unterscheiden, erlaubt es, in jeder Spur einen internen Längenstandard mitlaufen zu lassen; dies beseitigt Interpretationsfehler, die durch unterschiedliche Laufgeschwindigkeiten innerhalb des Gels auftreten können.

Bevor die Loci standardmäßig im Labor zur Identifikation von Individuen angewandt werden können, muß gesichert sein, daß die Größenbestimmung der Allele in Genauigkeit, Wiederholbarkeit und Eindeutigkeit erwiesen ist. Erste Untersuchungen wurden anhand der 100 gesunden unverwandten Blutspender in dieser Studie durchgeführt. Die Systeme zeigten sich robust und erbrachten reproduzierbare Ergebnisse. Derzeit laufen im Institut für Transfusionsmedizin und Immunhämatologie des Blutspendedienstes Hessen des DRK Frankfurt am Main weitere Analysen mit den Markern, um die entsprechenden Parameter zu kontrollieren. Zusätzlich können aufgrund der durchgeführten Sequenzuntersuchungen Allelleitern als interne Standards etabliert werden.

Die Sensitivität von Einzelproben gegenüber gemischten Proben ist unterschiedlich zu beurteilen und abhängig von dem Zweck, den sie erfüllen sollen. Wenn in einer gerichtsmedizinischen Untersuchung zahlreiche Loci in ein DNA-Profil einfließen, kann mit wenig

Probenmaterial oft eine große Ausbeute erzielt werden. Gibt es jedoch Schwierigkeiten in der Beurteilung mit fehlender Eindeutigkeit, sind auch hier Single-Locus Untersuchungen angebracht. Beim Einsatz von STR-Markern zur Verlaufsbeobachtung von Knochenmarkstransplantationen, bei der unter anderem das Wiederauftreten eines Rezidivs frühzeitig erkannt werden soll, ist die Verwendung von Einzelproben notwendig. Bereits geringe qualitative und quantitative Unterschiede in der Allelausprägung sind hierbei von enormer Wichtigkeit.

### **Allelverteilung in der Population**

Die populationsgenetischen Untersuchungen wurden mit fünf STR-Loci (D11S488, D18S51, D19S246, HUMFIBRA, HUMVWFA31/A) an 100 gesunden unverwandten kaukasischen Blutspendern aus der Region Hessen durchgeführt, denen zur Gewinnung genomischer DNA peripheres Blut entnommen worden war. Allel- und Genotypfrequenzen sind über Einzellocus PCR-Amplifikationen mittels Fragmentanalyse am 373A DNA-Sequencer von Applied Biosystems ermittelt worden. Die einzelnen Discriminationsindices (DI) lagen über 93 % und die kombinierte Vaterschaftsausschlußwahrscheinlichkeit bei 99,73 %. Bevor ein genetischer Marker für gerichtliche Untersuchungen eingesetzt werden kann, muß er auf Abweichung vom Hardy-Weinberg Gleichgewicht getestet werden. Zur Durchführung dieses Tests stehen unterschiedliche Methoden zur Verfügung. Der goodness-of-fit test [Li 1955] ist besonders für große Stichproben und der Exakttest [Levene 1949, Haldane 1954, Chapco 1976] für eine große Anzahl von Allelen geeignet. Bei relativ kleinen Stichproben, in denen viele Genotypfrequenzen Null werden, bietet sich die Bildung von Allelgruppen an [Rand et al. 1992, Pestoni et al. 1995]. Letztere Methode ist zwar weniger sensitiv, aber dennoch informativ. In der vorliegenden Arbeit wurde das Hardy-Weinberg Gleichgewicht mittels Bildung von Allelgruppen getestet und zeigte für keines der Systeme eine signifikante Abweichung. Da die fünf Loci auf unterschiedlichen Chromosomen liegen, ist eine gekoppelte Vererbung ausgeschlossen. In Übereinstimmung mit dem Empfehlungen der *DNA Commission of the International Society for Forensic Haemogenetics* wurden die Allele nach der Anzahl ihrer Repeateinheiten benannt [DNA Commission of the ISFH 1992, 1994, Bär et al. 1997]. Ergebnisse von STR-Typisierungen unterschiedlicher Laboratorien, in denen verschiedene Primersequenzen verwendet werden, können somit einfach verglichen werden. Wenn ein Allel nicht mit dem Standardmotiv übereinstimmt, wird die Anzahl von zusätzlichen Basenpaaren hinter einem Dezimalpunkt vermerkt.

Bei den hier vorgestellten Untersuchungen zum STR-Locus **D11S488**, erstmals beschrieben von Browne et al. [1993], wurden 15 Allele beobachtet, deren Längen für die Allele 26 bis 41 zwischen 242 bp und 302 bp lagen. Am häufigsten war Allel 33 (270 bp) mit einer Frequenz von 0,185 vertreten. Ausgehend von den ermittelten Sequenzdaten wurde die Nomenklatur vom Erstbeschreiber modifiziert. Die beiden Allele 38 und 41 waren bisher noch nicht in Populationsstudien erwähnt worden. Sie befinden sich am Rand des Allelbereichs und gehören zu den größten Allelen des Locus. Das Allel 38 (290 bp) entspricht in der Nomenklatur von Browne et al. [1993], welcher CEPH Familien (sie entstammen einem Untersuchungsfeld aus dem Centre d'Étude du Polymorphisme Humain in Paris) untersucht hat, dem Allel A2; das neue Allel 41 (302 bp) ist um 8 bp länger als sein beschriebenes Allel A1. Durch das Füllen der bisher bestandenen „Lücke“ mit dem Allel 38 steht zu vermuten, daß möglicherweise auch ein Fragment zwischen Allel 39 und 41 existiert. Wenn davon auszugehen ist, daß sich die Allele durch einen regelmäßigen Anstieg um vier Basenpaare auszeichnen, könnten weiterführende Studien mit größeren Stichproben darüber Aufschluß geben.

Erstbeschreiber für das System **D18S51** war Straub et al. [1993], der den Locus als reinen Tetrarepeat beschrieb und in einer kaukasischen Population (CEPH) 13 Allele beobachtete, die zwischen 267 bp und 319 bp lagen. Zwei Jahre später berichtete Urquhart et al. [1995] in seinem Artikel ebenfalls über dieses System und beschrieb 15 Allele, von denen 14 (279 bp-323 bp) in der britischen Population beobachtet wurden. Zwei dieser Allele fielen aus dem Tetrarepeatmuster heraus und wiesen einem Abstand von 2 bp auf. Beide Autoren konnten zu diesem Zeitpunkt noch keine sequenzierten Daten vorweisen. Aufgrund der in dieser Arbeit durchgeführten Sequenzanalyse konnte eine Modifikation der ursprünglich beschriebenen Allelnomenklatur entsprechend den beobachteten Repeateinheiten vorgenommen werden. In der vorliegenden Populationsstudie fanden sich 12 Allele (Allel 11 bis 22), mit einer Länge von 278 bp bis 322 bp. Die Allele zeichneten sich durch einen regelmäßigen Zuwachs von 4 bp aus. Im Gegensatz zu den Untersuchungen an britischen Kaukasiern [Barber u. Parkin 1996] fanden sich in dem hier untersuchten Kollektiv keine Allele mit einer 2 bp Größendifferenz. Durch zusätzliche Untersuchungen lassen sich möglicherweise noch weitere Allele finden, die durch Verdopplung oder Verlust von Repeateinheiten größer oder kleiner als die bisherigen Allele sind. Weiterhin wäre zu prüfen, ob auch in dieser Population Repeatdifferenzen von 2 bp auftreten [Barber u. Parkin 1996].

Weber et al. hatte 1993 erstmals Untersuchungen zur Allelverteilung an CEPH Familien für den Locus **D19S246** unternommen, wobei er 10 Allele beobachtete. In der vorliegenden Studie wurden 11 Allele (Allel 29 bis Allel 41) mit Längen von 182 bp bis 230 bp gefunden. Am häufigsten war Allel 29 mit einer Frequenz von 0,325 vertreten. Allel 30 gehört zu den neuen Allelen mit einer niedrigen Frequenz von 0,010. Es ersetzt damit die „Lücke“, die Weber et al. zwischen dem kleinsten und dem erst 8 bp später auftretenden Allel beschrieben hatte und reiht sich damit in den regelmäßigen Tetrarepeat ein. „Lücken“, in denen aufeinanderfolgende Allele einen Abstand von 8 bp aufweisen, finden sich zwischen den Allelen 30 und 32 sowie zwischen Allel 32 und 34. Durch weiterführende Untersuchungen könnte möglicherweise geklärt werden, ob noch weitere Allele in diesen Bereichen zu beobachten sind. Eine Vermutung in der Art liegt nahe, wenn davon auszugehen ist, daß der Polymorphismus, der den Allelen zugrunde liegt, durch einfache Duplikation oder Deletion von Repeateinheiten entsteht.

Während Mills et al. [1992] den STR-Marker **HUMFIBRA** (FGA) mit neun Allelen bei CEPH-Familien beschrieb, der mit Ausnahme eines 2er Repeats als Tetranukleotidrepeat auftrat, zeigte die Arbeitsgruppe Urquhart et al. [1995] bereits 20 verschiedene Fragmentlängen für den Locus in einer britischen Populationsstichprobe. Neben den üblichen Allelschritten von 4 bp beobachtete Urquhart zahlreiche Allele für den FGA-Locus, die um 2 bp differierten. Obwohl Mills die Primerbereiche anders gewählt hat, lassen sich die Ergebnisse anhand der GenBank miteinander vergleichen. Beide Autoren hatten zu diesem Zeitpunkt noch keine sequenzierten Daten vorliegen. Die hier vorliegende Untersuchung ergab für den STR-Marker HUMFIBRA 10 Allele mit einem regelmäßigen Tetrarepeat, wobei das Allel 17.2 um nur 2 bp differierte. In letzter Zeit wiesen mehrere Studien im europäischen Raum zwischen 19 und 21 häufige Allele auf [Barber et al. 1996, Gill et al. 1996, Gill et al. 1997]. Dementsprechend empfiehlt es sich, für den Locus HUMFIBRA zusätzliche Untersuchungen an einer größeren Populationsstichprobe durchzuführen, um möglicherweise erweiterte Ergebnisse in Bezug auf die Allelverteilung zu erhalten.

Kimpton et al. beschrieb 1992 als erster den STR-Locus **HUMVWFA31/A** als Tetrarepeat. Er war damit einer von vielen, der sich mit diesem Marker beschäftigte [Urquhart et al. 1994, Urquhart et al. 1995, Pestoni et al. 1995, Gill et al. 1996, Möller et al. 1994a, 1994b, Walsh et al. 1996, Rousselet et al. 1996, Nagai et al. 1996, Martin et al. 1995]. Zwischen 7 und 13 Allele wurden in den einzelnen Studien beschrieben, wobei ein Allel mit 2 bp Differenz beschrieben wurde [Urquhart et al. 1995, Gill et al. 1996]. Da dieser Locus in der Literatur mit

sequenzierten Daten bereits ausreichend beschrieben wurde, ist in der vorliegenden Arbeit auf die Sequenzierung verzichtet worden. Die 8 beobachteten Allele 13 bis 20 wurden durch Fragmentanalyse ermittelt, wobei die allgemeingültige Allelnomenklatur übernommen wurde. Allele, die vom Tetranukleotidmuster abweichen sind nicht beobachtet worden.

### Polymorpher Informationsgehalt

Mit den in dieser Arbeit etablierten Multiplex-Amplifikationssystemen wurde ein hochgradiger Unterscheidungsapparat geschaffen, der auf dem STR-PCR-Identifikationssystem beruhend für den Triplex einen kombinierten polymorphen Informationsgehalt (kombinierte pM) von  $1,13 \times 10^{-4}$  und für den Quadruplex von  $6,2 \times 10^{-6}$  aufweist. Die einzelnen Werte der untersuchten Marker für den polymorphen Informationsgehalt (pM) sind in Tabelle 4.1 zusammenfassend dargestellt.

*Tabelle 4.1: pM-Werte für die fünf individuellen Loci und pM-Werte für Triplex, Quadruplex und möglichen Pentaplex (Quadruplex plus HUMFIBRA)*

	pM		PM
D11S488	0,031	D11S488	0,031
HUMFIBRA	0,055	D18S51	0,044
HUMVWFA31/A	0,066	D19S246	0,069
<b>Tetraplex</b>	<b><math>1,13 \times 10^{-4}</math></b>	HUMVWFA31/A	0,066
		<b>Quadruplex</b>	<b><math>6,20 \times 10^{-6}</math></b>
		<b>Quadruplex + FIB</b>	<b><math>3,00 \times 10^{-7}</math></b>

Da die Systeme D19S246 und HUMFIBRA überlappende Allelbereiche aufwiesen und zudem noch mit der gleichen Fluoreszenzfarbe markiert waren, sind sie bei den bisherigen Untersuchungen noch nicht gemeinsam in einem Multiplex getestet worden. Durch einen Austausch der Primerfarbe könnten jedoch in weiterführenden Studien alle fünf Systeme als Pentaplex zusammengefaßt werden. Der kombinierte pM-Wert würde bei  $3 \times 10^{-7}$  liegen. Damit läge die Wahrscheinlichkeit, daß zwei Proben in den fünf Systemen völlig identisch sind, bei 1 zu 30 Millionen.

### Sequenzstruktur

STRs zeichnen sich dadurch aus, daß sie aus einfachen sich wiederholenden Sequenzstrukturen bestehen, den Repeats. Variationen in der Anzahl der Repeateinheiten spiegeln den hohen Grad an Längenpolymorphismus wieder. Um die Struktur der Repeateinheiten, die Anzahl der vorliegenden Repeats für die Allelbestimmung und die genaue Größe der Allelfragmente kennenzulernen, wurden die Allele der einzelnen Systeme mit Ausnahme von HUMVWFA31/A sequenziert. Repräsentative Allele jeder Größe sind in die Sequenzierung miteinbezogen worden. Die Allellängen, die anhand der Fragmentanalyse bestimmt wurden, waren abhängig von der Markierungsmethode, dem markierten Primer und von der Taq Polymerasemarke [Knowles et al. 1992, Straub et al. 1993]. Dabei gab es Loci, deren Sequenzstruktur einfach mit einem Motiv aufgebaut war, wie z.B. D18S51 und HUMFIBRA, und andere, bei denen ein größerer Komplex den Polymorphismus hervorrief.

Für den Locus **D18S51** zeigt die Sequenzierung eine einfache Repeatregion (*simple repeat*), die aus dem Tetranukleotid AGAA besteht. Ein Beispiel für einen ähnlich einfachen Repeat ist der Locus HUMFES/FPS, der eine einzige Repeatsequenz ATTT aufweist [Gill et al. 1995, Möller et al. 1994a]. Der Allelpolymorphismus des STR-Systems D18S51 ist auf die regelmäßigen Insertionen von kompletten AGAA Tetranukleotideinheiten zurückzuführen. Innerhalb der flankierenden Regionen wurden keine Sequenzvariationen festgestellt. Im Vergleich zu Barber u. Parkin [1996] wurde in der vorliegenden Untersuchung kein Allel mit einer Differenz von 2 bp beobachtet. Gill et al. beschrieb 1996 diesen Locus aufgrund von inkompletten 2er Repeats als hochgradig diskriminierenden komplexen STR, der auf einem *compound repeat* beruht.

Ebenfalls durch eine einfache Repeatstruktur ist der STR-Locus **HUMFIBRA** mit dem Tetranukleotidblock TCTT charakterisiert. Aufgrund des Auftretens von Allelen, die um 2 bp vom regulären Tetrarepeat abweichen, besitzt dieser Marker entsprechend anderen Autoren [Barber et al. 1996, Gill et al. 1996, Urquhart et al. 1995, Gill et al. 1995] den Aufbau eines *compound repeat*.

Der YAAG Repeat des Locus **D11S488** zählt ebenfalls zu den *compound repeats*. Er setzt sich aus mehreren Blöcken von AAAG und GAAG Motiven zusammen, von denen nur einige in ihrer Anzahl variieren: (AAAG)<sub>7-17</sub> (GAAG)<sub>4-15</sub> GAAAG (GAAG)<sub>4-5</sub> GAAAGAAAG GAAAG (GAAG)<sub>4-6</sub>. Mit Variationsmöglichkeiten an unterschiedlichen Stellen im Repeatbereich wurden Allele gleicher Länge aber unterschiedlicher Sequenz beobachtet. Damit besitzt dieser Marker nicht nur einen Längenpolymorphismus, sondern auch einige



strukturelle Variationen und kann ähnlich dem D12S391 System [Lareu et al. 1996] als STR mit eingeschobenen Mikrovariationen [Brinkmann 1996], Mikroheterogenitäten [Möller et al. 1994a] oder als compound STR [Urquhart et al. 1994] bezeichnet werden.

Charakteristisch ist der GAAAG Block, der in jeder Sequenz zu finden ist, und die beiden variablen Regionen voneinander trennt. Die Entstehung der beiden Motive AAAG und GAAG ist möglicherweise auf eine Transition von A nach G zurückzuführen. Weiterhin kommen Insertionen von Guanin oder Adenin in Frage, was zahlreiche GAAAG Sequenzen erklären könnte. Nahezu der komplette Bereich zwischen den Primern besteht aus Sequenzblöcken von Guanin und Adenin. Sie bilden den Repeatanteil und kleine Variationen davon. Gerade die erste Base im Tetrarepeat, ob Adenin oder Guanin, entscheidet über das Vorkommen von Mikroheterogenitäten bzw. Isoallelen in der stark variablen Region I (VR I).

Ebenso wie der D11S488 Locus kann auch der **D19S246** Locus den *compound repeats* zugeordnet werden. Der TSTA-Repeatbereich besteht aus den beiden variablen Regionen (TCTA)<sub>2-3</sub> und (TGTA)<sub>7-12</sub> (TCTA)<sub>5-13</sub>. Neben einem Längenpolymorphismus traten auch hier innerhalb gleicher Allellängen strukturelle Variationen auf, die sich im denaturierenden Polyacrylamidgel nicht unterscheiden ließen [Möller et al. 1994a]. Die kürzeren Allele 29-37 zeigten nur Variationen in der variablen Region II (VR II). Möglicherweise ist eine Duplikation eines Repeats in VR I ein späteres Ereignis in der Evolution, das nur bei längeren Allelen zu finden ist. Es kann angenommen werden, daß der compound repeat TCTA und TGTA ursprünglich durch Punktmutation (von Cytosin zu Guanin oder umgekehrt) aus einem simple repeat entstanden ist. In den flankierenden Regionen sind noch weitere von diesen oder ähnlich strukturierten Repeatblöcken zu beobachten. So ließe sich die Entstehung von TCTG Blöcken durch Punktmutation aus dem TCTA Muster oder durch AT Deletion erklären. Diese Sequenzzusammensetzung verspricht noch weitere potentielle Repeateinheiten, die möglicherweise variablen Charakter besitzen. Daher sollten zusätzliche Sequenzuntersuchungen prüfen, ob dieses System nicht mehr als ein einfacher compound repeat ist.

### **Vor- und Nachteile von simple und complex repeat Markern**

Um eine Einheit in die Allelbezeichnung von short tandem repeat (STR) Markern zu bringen, machte Urquhart et al. [1994] den Vorschlag einer universalen Allelnomenklatur für einfache und komplexe Repeats. Er unterscheidet dabei simple repeats, compound repeats und complex repeats. Simple repeats enthalten Einheiten identischer Länge und Sequenz, wie beispielsweise HUMFES/FPS [Polymeropoulos et al. 1991] oder das hier beschriebene

D18S51. Bestehen sie dabei aus zwei oder mehreren simple repeats, spricht man von compound repeats (D11S488, D19S246). Unter complex repeats werden Blöcke aus mehreren verschiedenen Repeateinheiten unterschiedlicher Länge verstanden, die von mehr oder weniger variablen Sequenzen unterbrochen werden.

Die Verwendung von simple oder complex repeat Markern für Vaterschaftsgutachten oder gerichtliche Belange bringt Vor- und Nachteile mit sich. Simple Repeats haben den Vorteil, daß sie einfacher in der Anwendung sind. Ihre Allele unterscheiden sich in der Regel durch einen regelmäßigen Anstieg um jeweils eine Repeateinheit. Allerdings sind auch hier Ausnahmen vorhanden [Urquhart et al. 1993, Möller u. Brinkmann 1994]. Im TH01 System z.B. besitzt das Allel 4.3 eine Längendifferenz um 1 bp [Puers et al. 1993]. Seidl et al. beobachtete 1996 beim CYP19 Marker eine 3 bp Deletion in der 5' flankierenden Region. Compound repeat Marker dagegen sind mitunter schwieriger zu handhaben. Durch das Auftreten von Mikroheterogenitäten sind Allele gleicher Länge nicht unbedingt Allele gleicher Sequenz. Soll diese Trennung miteinbezogen werden, ist eine Sequenzierung unumgänglich und der Arbeitsgang entsprechend aufwendig. Im nicht denaturierenden Polyacrylamidgel sind signifikante strukturelle Variationen der STRs nicht zu erkennen [Lareu et al. 1996]. Complex repeats, bei denen häufiger das Repeatmuster verlassen wird, indem Allele mit 2 bp oder 1 bp Differenz auftreten, sind schwieriger in der Fragmentanalyse nachzuweisen. Die Bestimmung von simple oder compound bzw. complex STR-Polymorphismen sollte daher grundsätzlich mittels Allelleitern durchgeführt werden [DNA Commission of the ISFH 1992, 1994]. Wegen ihres größeren Polymorphismus erleichtern dagegen hypervariable Loci häufig die Lösung schwieriger Vaterschaftsfragen, was mit konventionellen Markern aufwendiger ist und weitaus mehr Zeit in Anspruch nimmt [Tamaki et al. 1996].

Zusammenfassend enthalten die fünf in dieser Arbeit untersuchten Marker zahlreiche Charakteristika, die sie für einen Einsatz im Hinblick auf forensische Zwecke interessant machen. Eine PCR-Amplifikation der Systeme ist sicher durchzuführen. Die Loci besitzen eine hohe Heterozygotenrate und lassen sich auch mit anderen STRs im Multiplex analysieren. Aufgrund der günstigen pM-Werte und der Übereinstimmung mit dem Hardy-Weinberg Gleichgewicht stellen diese STR-Marker eine sinnvolle Ergänzung in der Durchführung abstammungsgenetischer DNA-Untersuchungen dar.

# Zusammenfassung

Short tandem repeat (STR) Loci sind ideale Marker für gerichtliche und abstammungsgenetische DNA-Untersuchungen. Sie bestehen aus sich wiederholenden 2-6 bp langen Einheiten und sind über das gesamte menschliche Genom verteilt. Aufgrund ihrer geringen Allellängen (100-600 bp) lassen sich STRs leicht mit Hilfe der Polymerasekettenreaktion (PCR) amplifizieren.

In der vorliegenden Arbeit sind fünf STR-Polymorphismen der Loci D11S488, D18S51, D19S246, HUMFIBRA (FGA) und HUMVWFA31/A auf ihre Populationsgenetik und Sequenzstruktur hin untersucht worden. Die Daten wurden anhand genomischer DNA von 100 gesunden, unverwandten kaukasischen Blutspendern aus der Region Hessen gewonnen.

Über ein 6 %iges denaturierendes Polyacrylamidgel wurden die PCR-Produkte aufgetrennt und unter Gebrauch fluoreszenzmarkierter Primer mit einem 373A DNA-Sequenzier analysiert.

Der Locus D11S488 ist durch eine zusammengesetzte Repeatregion (compound repeat) von AAAG und GAAG Blöcken gekennzeichnet. Mit einer Variationsmöglichkeit an vier unterschiedlichen Positionen kam es zum Auftreten von Mikroheterogenitäten in Allelen gleicher Länge. 29 verschiedene Allele wurden gefunden, die basierend auf ihrer Gesamtrepeatanzahl  $(YAAG)_{26-41}$  (242 bp-302 bp), in 15 Allelklassen gruppiert wurden. Bei D19S246 (TGTA und TCTA) liegt ebenfalls ein compound repeat vor. Mikroheterogenitäten führten in 11 Allelgruppen (182-230 bp) zu 17 unterschiedlichen Allelen. Der Locus D18S51 (AGAA) ist ein STR-Polymorphismus mit einer einfachen Repeatstruktur (simple repeat). 12 unterschiedliche Allele wurden beobachtet, die alle einen regelmäßigen Tetranukleotidrepeat aufwiesen. Die Allelspanne reichte von Allel 11 mit 278 bp bis Allel 22 mit 322 bp. Ebenfalls ein einfacher Repeat bestimmt das HUMFIBRA System (TCTT). Eines der beobachteten 10 Allele differierte allerdings um nur 2 bp.

Zusammen mit dem compound repeat HUMVWFA31/A (8 Allele) sind alle untersuchten Marker multiplexfähig. Ein ausgeprägter Polymorphismus der individuellen Loci, die dem Hardy-Weinberg Equilibrium folgen, sowie eine Übereinstimmungswahrscheinlichkeit (pM) der Merkmale zwischen unverwandten Personen von  $3 \times 10^{-7}$  machen eine Analyse dieser fünf Tetranukleotid-Marker zu einer sinnvollen Ergänzung in der Bearbeitung abstammungsgenetischer Fragestellungen.



# Publikationsnachweis

Teilergebnisse dieser Arbeit wurden vorgestellt oder veröffentlicht in:

Seidl C, Müller S, Jäger O, Seifried E: Populationsgenetische Untersuchungen der Short Tandem Repeat Polymorphismen D19S246 und D21S11. Vortrag wissenschaftl. Symposium der Sachverständigen für Abstammungsgutachten 1995

Seifried E, Müller S, Jäger O, Seidl C: Anwendung von Short Tandem Repeat Polymorphismen bei abstammungsgenetischen Fragestellungen. Vortrag - Tagung der deutschsprachigen Arbeitsgruppe der ISFH; Dortmund; (1996)

Seidl C, Müller S, Kilp M, Seifried E: Fluorescence based multiplex analysis of the STR polymorphism FIBRA (FGA), VWFA31 and D18S51 in German caucasoid individuals. In: Olaisen B, Brinkmann B, Lincoln PJ (Hrsg.): Progress in Forensic Genetics 7; Elsevier; Amsterdam, Lausanne, New York; (1998) 353-355

Seidl C, Müller S, Seifried E: Sequence analysis and allele frequencies of STR loci D19S246 and D11S488 in German Caucasians. In: Olaisen B, Brinkmann B, Lincoln PJ (Hrsg.): Progress in Forensic Genetics 7; Elsevier; Amsterdam, Lausanne, New York; (1998) 356-359

Seidl C, Müller S, Jäger O, Seifried E: Sequence analysis and population data of short tandem repeat polymorphisms at locus D8S639 and D11S488. Int J Leg Med (1999), **112**: 355-359



---

# Literaturverzeichnis

Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B: DNA recommendations Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *Int J Leg Med* (1997), **110**: 175-176

Barber MD, McKeown BJ, Parkin B: Structural variation in the alleles of a short tandem repeat system at the human alpha fibrinogen locus. *Int J Legal Med* (1996), **108**: 180-185

Barber MD, Parkin BH: Sequence analysis and allelic designation of the two short tandem repeat loci D18S51 and D8S1179. *Int J Legal Med* (1996), **109**: 62-65

Beckmann JS, Weber JL: Survey of human and rat microsatellites. *Genomics* (1992), **12**: 627-631

Boerwinkle E, Xiong W, Fourest E, Chan L: Rapid typing of a tandemly repeated hypervariable locus by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. *Proc Natl Acad Sci USA* (1989), **17**: 212-216

Botstein D, White RL, Skolnick M, Davis RW: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Gen* (1980), **32**: 182-190

Brinkmann B: The STR approach. In: Carracedo A, Brinkmann B, Bär W (Hrsg.): *Advances in forensic haemogenetics 6*; Springer; Berlin, Heidelberg; (1996)

Brinkmann B: The use of STRs in stain analysis. In: *Proceedings from the third international Symposium on Human Identification*; Promega Co.; Madison, USA; (1992) 357-373

Browne D, Gen M, Evans GA, Clark SP, Litt M: Tetranucleotide repeat polymorphism at the D11S488 locus. *Hum Mol Genet* (1993), **2**: 89

Budowle B, Monson KL, Chakraborty R: Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Leg Med* (1996), **108**: 173-176

Burke T, Dolf G, Jeffreys J, Wolff R: *DNA fingerprinting: approaches and applications*. Birkhauser; Basel (1991)

Chapco W: An exact test of the Hardy-Weinberg equilibrium in small samples. *Biometrics* (1976), **33**: 536-542

Clark JM: Novel non-template nucleotide addition reactions catalysed by procaryotic and eucaryotic DNA polymerase. *Nucleic Acids Res* (1988), **16**: 9677-9686

DNA Commission of the ISFH: DNA recommendations – 1992 report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. *Int J Leg Med* (1992), **105**: 63-64

DNA Commission of the ISFH: DNA recommendations - 1994 report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. *Int J Leg Med* (1994), **107**: 159-160

Edwards A, Civitello A, Hammond HA, Caskey CT: DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* (1991), **49**: 746-756

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R: Genetic variation of five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* (1992), **12**: 241-253

Engelstein M, Hudson TJ, Lane JM, Lee MK, Leverone B, Landes GM, Peltonen L, Weber JL, Dracopoli NC: A PCR-based linkage map of human chromosome 1. *Genomics* (1993), **15**: 251-258

Evett IW, Gill P: A discussion of the robustness of methods for assess in the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis* (1991), **12**: 226-230

Fornage M, Chan L, Siest G, Boerwinkle E: Allele frequency distribution of the (TG)<sub>n</sub>(AG)<sub>m</sub> microsatellite in the apolipoprotein C-II gene. *Genomics* (1992), **12**: 63-68

Fregeau CJ, Fourney RM: DNA typing with fluorescently tagged short tandem repeats: A sensitive and accurate approach to human identification. *Biotechniques* (1993), **15**: 100-119

Garber RA, Morris JW: General equations for the average power of exclusion for genetic systems of n codominant alleles in one-parent and no-parent cases of disputed parentage. *American Association of Blood Banks* (1983); 277-280

Gill P, d'Aloja E, Andersen J, Dupuy B, Jangblad M, Johnsson V, Kloosterman AD, Kratzer A, Lareu MV, Meldegaard M, Phillips C, Pfitzinger H, Rand S, Sabatier M, Scheithauer R, Schmitter H, Schneider P, Vide MC: Report of the European DNA profiling group (EDNAP): an investigation of the complex STR loci D21S11 and HUMFIBRA (FGA). *Forensic Sci Int* (1997), **86**: 25-33

Gill P, Evett I: Population genetics of short tandem repeat (STR) loci. *Genetica* (1995), **96**: 69-87



Gill P, Ivanov P, Kimpton CP, Piercy R, Benson NJ, Tully G, Evett I, Hagelberg E, Sullivan KM: Identification of the Romanov family by DNA analysis. *Nature Genet* (1993), **6**: 130-135

Gill P, Kimpton CP, Sullivan KM: A rapid polymerase chain reaction method for identifying fixed specimens. *Electrophoresis* (1992), **13**: 173-175

Gill P, Kimpton CP, Urquhart A, Oldroyd N, Millican ES, Watson SK, Downes TJ: Automated short tandem repeat (STR) analysis in forensic casework - a strategy for the future. *Electrophoresis* (1995), **16**: 1543-1552

Gill P, Sullivan KM, Werrett DJ: The analysis of hypervariable DNA profiles: problems associated with the objective determination of the probability of a match. *Hum Genet* (1990), **85**: 75-79

Gill P, Urquhart A, Millican E, Oldroyd N, Watson S, Sparkes R, Kimpton CP: A new method of STR interpretation using inferential logic - development of a criminal intelligence database. *Int J Legal Med* (1996), **109**: 14-22

Hagelberg E, Gray IC, Jeffreys AJ: Identification of the skeletal remains of a murder victim by DNA analysis. *Nature* (1991), **352**: 427-429

Haldane JBS: An exact test for randomness of mating. *J Genet* (1954), **52**: 631-635

Hauge XY, Litt M: A study of the origin of „shadow“ bands seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum Mol Genet* (1993), **2**: 411-415

Hultman T, Bergh S, Moks T, Uhlen M: Bidirectional solid phase sequencing of in vitro amplified plasmid DNA. *BioTechniques* (1991), **10**: 84-93

Hummel K, Ihm P, Schmidt V: Biostatistische Abstammungsbegutachtung mit Blutgruppenbefunden. Gustav Fischer Verlag; Stuttgart; (1971)

Innis M, Myambo KB, Gelfand DH, Brow MAD: DNA sequencing with *Thermus aquaticus* DNA polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci USA* (1988), **85**: 9436-9440

Jeffreys AJ, Allen MJ, Hagelberg E, Sonnberg A: Identification of the skeletal remains of Josef Mengele by DNA analysis. *Forensic Sci Int* (1992), **56**: 65-76

Jeffreys AJ, Royal NT, Wilson V, Wong Z: Spontaneous mutation rates to new length alleles at tandem-repetitive loci in human DNA. *Nature* (1988), **332**: 278-281

Jones DA: Blood samples: Probability of discrimination. *J Forens Sci Soc* (1972), **12**: 355-359

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M: Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods and Applications* (1993), **3**: 13-22

Kimpton CP, Walton A, Gill P: A further tetranucleotide repeat polymorphism in the vWF gene. *Hum Mol Genet* (1992), **1**: 287

Kirby LT: Probability and Statistical Analysis. In: *DNA fingerprinting: an introduction*. Stockton Press; New York; (1992) 149-177

Knowles JA, Vieland VJ, Gilliam TC: Perils of gene mapping with microsatellite markers. *Am J Hum Genet* (1992), **51**: 905-909

Lareu MV, Pestoni C, Schürenkamp M, Rand S, Brinkmann B, Carracedo A: A highly variable STR at the D12S391 locus. *Int J Leg Med* (1996), **109**: 134-138

Levene H: On a matching problem arising in genetics. *Ann Math Stat* (1949), **20**: 91-94

Li CC: *Population genetics*. Chicago (1955); University of Chicago Press

Litt M, Hauge X, Sharma V: Shadow bands seen when typing polymorphic repeats. *BioTechniques* (1993), **15**: 280-284

Litt M, Luty JA: A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am Hum Genet* (1989), **44**: 397-401

Martin P, Alonso A, Budowle B, Albarrán C, Gracia O, Sancho M: Spanish population data on 7 tetrameric short tandem repeat loci. *Int J Leg Med* (1995), **108**: 145-149

Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Ac Res* (1988), **16**: 1215

Mills KA, Even DA, Murray JC: Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). *Hum Mol Genet* (1992), **1**: 779

Möller A, Brinkmann B: Locus ACTBP2 (SE33). Sequencing data reveals considerable polymorphism. *Int J Leg Med* (1994), **106**: 262-267

Möller A, Meyer E, Brinkmann B: Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. *Int J Leg Med* (1994a), **106**: 319-323

- Möller A, Wiegand P, Grüschow C, Seuchter SA, Baur MP, Brinkmann B: Population data and forensic efficiency values for the STR system Hum VWA, HumMBP and HumFABP. *Int J Leg Med* (1994b), **106**: 183-189
- Morral N, Estivill X: Multiplex PCR amplification of three microsatellites within the CFTR gene. *Genomics* (1992), **13**: 1362-1364
- Mullis KB, Faloona F, Scharf S, Saiki R, Horn G, Erlich H: Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbour Symp Quant Biol* (1986), **51**: 263-273
- Nagai A, Yamada S, Watanabe Y, Bunai Y, Ohya I: Analysis of the STR loci HUMF13A01, HUMLIPOL, HUMZH01, HUMTPOX and HUMVWFA31 in a Japanese population. *Int J Leg Med* (1996), **109**: 34-36
- Newton CR, Graham A: PCR. Spektrum, Akad. Verl.; Heidelberg, Berlin, Oxford; (1994)
- Nomenclature Committee of International Union of Biochemistry (1985) Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. *Eur J Biochem* (1985), **150**: 1-5
- Oudet C, Heilig R, Hanauer A, Mandel JL: Nonradioactive assay for new microsatellite polymorphisms at the 5' end of the dystrophin gene, and estimation of intragenetic recombination. *J Hum Genet* (1991), **49**: 311-319
- Pena SDJ, Chakraborty R, Eppelen JT, Jeffreys AJ: DNA fingerprinting: state of the science. Birkhauser; Basel; (1993)
- Pestoni C, Lareu MV, Rodriguez MS, Muñoz I, Barros F, Carracedo A: The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data for Galicia (NW Spain). *Int J Leg Med* (1995), **107**: 283-290
- Polymeropoulos MH, Rath DS, Xiao H, Merrill RM: Tetranucleotide repeat polymorphism at the c-fes/fps proto-oncogene (FES). *Nucleic Acids Res* (1991), **19**: 4018
- Polymeropoulos MH, Xiao H, Rath DS, Merrill CR: Tetranucleotide repeat polymorphism at the human intestinal fatty acid binding protein gene (FABP2). *Nucleic Acids Res* (1990), **18**: 7198
- Puers C, Hammond HA, Jin L, Caskey CT, Schumm JW: Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01 [AATG]<sub>n</sub> and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *Am J Hum Genet* (1993), **53**: 953-958
- Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B: Population genetic and forensic efficiency data of 4 AmpFLPs. *Int J Leg Med* (1992), **104**: 329-333

Robert Koch-Institut, Berlin: Richtlinien für die Erstattung von Abstammungsgutachten - Novellierung 1996. Bundesgesundheitsbl (1996), **39**: 312-316

Rousselet F, Pfitzinger H, Mangin P: French Caucasian population data obtained from fluorescently detected HumvWFA31/A and HumF13A01 short tandem repeat loci: Int J Leg Med (1996), **109**: 5-9

Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA (1977), **74**: 5463-5467

Saiki RK, Scharf SJ, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N: Enzymatic amplification of  $\beta$ -Globulin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science (1985), **230**: 1350-1354

Schlotterer C, Tautz D: Slippage synthesis of simple sequence DNA. Nucleic Acids Res (1992), **20**: 211-215

Schwartz LS, Tarleton J, Popovich B, Seltzer WK, Hoffman EP: Fluorescent multiplex linkage analysis and carrier detection for Duchenne/Becker muscular dystrophy. Am J Hum Genet (1992), **51**: 721-729

Seidl C, Jäger O, Kilp M, Seifried E: Automated fluorescence PCR based analysis of the STR polymorphism at locus D8S639 and the CYP19 gene. Adv Forensic Haemogenet (1996), **6**: 216-218

Shriver MD, Jin L, Boerwinkle E, Deka R, Ferrell RE, Chakraborty R: A novel measure of genetic distance for high polymorphic tandem repeat loci. Mol Biol Evol (1995), **12**: 914-920

Sprecher CJ, Puers C, Lins AM, Schumm JW: General approach to analysis of polymorphic short tandem repeat loci. BioTechniques (1996), **20**: 266-276

Stallings RL, Ford AF, Nelson D, Torney DC, Hildebrand CE, Moyziz RK: Evolution and distribution of (GT)<sub>n</sub> repetitive sequences in mammalian genomes. Genomics (1991), **10**: 807-815

Straub RE, Speer MC, Luo Y, Rojas K, Overhauser J, Ott J, Gilliam C: A microsatellite genetic linkage map of human chromosome 18. Genomics (1993), **15**: 48-56

Sullivan KM, Pope S, Gill P, Robertson JM: Automated DNA profiling by fluorescent labeling of PCR products. PCR Methods Applic (1992), **2**: 34-40

Tabor S, Richardson CC: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc Natl Acad Sci USA (1987), **84**: 4767-4771

- Tamaki Y, Fukuda M, Wang W, Kishida T: Microsatellite typing in a paternity case against a deceased man whose two brothers were available for testing. *Nippon-Hoigaku-Zasshi* (1996), **50**: 82-86
- Tautz D: Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* (1989), **17**: 6463-6471
- Tomfohrde J, Wood S, Scgertzer M, Wagner MJ, Wells DE, Parrish J, Sadler LA, Blanton SH, Daiger SP, Wang Z, Wilkie PJ, Weber JL: Human chromosome 8 linkage map based on short tandem repeat polymorphisms: Effect of genotyping errors. *Genomics* (1992), **14**: 144-152
- Tully G, Sullivan KM, Gill P: Analysis of 6 VNTR loci by „multiplex“ PCR and automated fluorescent detection. *Hum Genet* (1993), **92**: 554-562
- Urquhart A, Kimpton CP, Downes TJ, Gill P: Variation in Short Tandem Repeat sequences - a survey of twelve microsatellite loci for use as forensic identification markers. *Int J Leg Med* (1994), **107**: 13-20
- Urquhart A, Kimpton CP, Gill P: Sequence variability of the tetranucleotide repeat of the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus. *Hum Genet* (1993), **92**: 637-638
- Urquhart A, Oldroyd NJ, Kimpton CP, Gill P: Highly discrimination heptaplex short tandem repeat PCR system for forensic identification. *BioTechniques* (1995), **18**: 116-121
- Walsh PS, Fildes NJ, Reynolds R: Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res* (1996), **24**: 2807-2812
- Weber J, May P: Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* (1989), **44**: 388-396
- Weber JL, Polymeropoulos MH, May PE, Kwitek AE, Xiao H, McPherson JD, Wasmuth JJ: Mapping of human chromosome 5 microsatellite DNA polymorphisms. *Genomics* (1991), **11**: 695-700
- Weber JL, Wang Z, Hanseen K, Stephenson M, Kappel C, Salzman S, Wilkie PJ, Keats B, Dracopoli NC, Brandriff BF, Olsen AS: Evidence for human meiotic recombination interference obtained through construction of short tandem repeat-polymorphism linkage map of chromosome 19. *Am J Hum Genet* (1993), **53**: 1079-1095
- Wiegand P, Budowle B, Rand S, Brinkmann B: Forensic validation of the STR systems SE33 and TC11. *Int J Leg Med* (1993), **105**: 315-320
- Ziegle JS, Su Y, Kevin P, Corcoran LN, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Diehl SR: Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics* (1992), **14**: 1026-1031



# Anhang

## IUB-Code

In Übereinstimmung mit den Empfehlungen des Nomenklaturkomitees der Internationalen Union für Biochemie [Nomenclatur Committee of International Union of Biochemistry (1985)] wurde folgender Code vereinbart.

IUB-Code		
<b>A</b> = Adenosin	<b>R</b> = A oder G	(puRine)
<b>C</b> = Cytosin	<b>Y</b> = C oder T	(pYrimidine)
<b>G</b> = Guanin	<b>K</b> = G oder T	(Keto)
<b>T</b> = Thymin	<b>M</b> = A oder C	(aMino)
<b>B</b> = C, G oder T	<b>S</b> = G oder C	(Strong-3H bonds)
<b>D</b> = A, G oder T	<b>W</b> = A oder T	(Weak-2H bonds)
<b>H</b> = A, C oder T	<b>N</b> = aNy base	
<b>V</b> = A, C oder G		

**Sequenz von D11S488**

Ausschnitt aus der GenBank-Sequenz [accession no. L04732]. Zum Vergleich ist die in dieser Arbeit beobachtete Sequenz unter die GenBank-Sequenz gelegt. Die Primerbereiche sind unterstrichen dargestellt.

```

      3       9       15       21       27       33       39
      |       |       |       |       |       |       |
1  GGT GAG GTT GCA GTG AGC CAT GAT CGC ACC ACT GCA CTC CAG CCT
46 AGG CAA TAG AGA CCC TGT GAA AGA AGA AAA GAA AGA AAG AAA GAA
   AGG CAA TAG AGA CCC TGT GAA AGA AGA AAA GAA AGA AAG AAA GAA
91 AGG TAA GAA AAA GAA AGA AAG AAA GAA AGA AAG AAA GAA AGA AAG
   AGG TAA GAA AAA GAA AGA AAG AAA GAA AGA AAG AAA GAA AGA AAG
136 AAA GAA AGA AAG AAA GAA AGA AAG GAA GGA AGG AAG GAA GGA AGG
     AAA GAA AGA AAG AAA GAA AGA AAG GAA GGA AGG AAG GAA GGA AGG
181 AAG GAA GGA AGG AAG GAA GAA AGG AAA GGA AGG AAG GAA GGA AGG
     AAG GAA GGA AGG AAG GAA GGA AGG AAA GGA AGG AAG GAA GGA AGG
226 AAG GAA GGA AGG AAA GGA AGG AAG GAA GGA AGG AAG GAA AGA AAG
     AAG                                     GAA AGA AAG
271 GAA AGG AAG GAA GGA AGG AAG GAA GGA AGG AAA GAA AGA AAG AGA
     GAA AGG AAG GAA GGA AGG AAG GAA GGA AGG AAA GAA AGA AAG AGA
316 GAG GGA GAA AGA AAG GAA GGA AAC TAA CTC TCA GTA ACA ATT CAT
     GAG GGA GAA AGA AAG GAA GGA AAC TAA CTC TCA GTA ACA ATT CAT
361 CAT CTA CAT ATT TTA ACA CAG TCA GGT CAG GCT ATC AGA TGG CCT
     CAT C

```

**Sequenz von HUMVWFA31/A**

Auszug aus der GenBank-Sequenz [accession no. M25858] von HUMVWFA31/A. Die Primer sind unterstrichen dargestellt.

```

1621 CTTGGCTGAG ATGTGAAAGC CCTAGTGGAT GATAAGAATA ATCAGTATGT
1671 GACTTGGATT GATCTATCTG TCTGTCTGTC TGTCTATCTA TCTATCTATC
1721 TATCTATCTA TCTATCTATC TATCTATCTA TCTATCCATC TATCCATCCA
1771 TCCTATGTAT TTATCATCTG TCCTATCTCT

```



# Danksagung

An dieser Stelle möchte ich allen, die zum Gelingen dieser Arbeit beigetragen haben, meinen Dank aussprechen.

Herrn Prof. Dr. med. E. Seifried danke ich für die Themenstellung zu dieser Dissertation sowie für die Möglichkeit, alle praktischen Arbeiten bei ihm im Institut auszuführen.

Sehr herzlich möchte ich mich auch bei Herrn Priv. Doz. Dr. med. C. Seidl für die gute Betreuung meiner Arbeit bedanken und für seine Bereitschaft, sich immer Zeit für die Diskussion von Problemen und Ergebnissen zu nehmen.

Gerade in „schweren Zeiten“ ist es schön zu wissen, daß man Freunde im Labor hat, die immer ein offenes Ohr haben. Deshalb gilt auch allen Laborkollegen, mein Dank für die freundliche Zusammenarbeit und Hilfsbereitschaft. Besonders Frau Kilp, Ute und Marianne, die mich in zahlreiche Labortechniken eingeführt und mir mit Rat und Tat zur Seite gestanden haben, gilt mein herzliches Dankeschön. Für die moralische Unterstützung möchte ich mich bei meinen Mitdoktoranden Michael und Barbara bedanken.

Nicht zuletzt möchte ich auch meiner Familie Edith und Miriam danken, die mir Kraft, Ausdauer und Zuversicht geschenkt haben. Bei Gerd möchte ich mich für seine Hilfe in Computerangelegenheiten und seine Geduld, sich mit ihm nicht geläufigen Fragestellungen der Wissenschaft auseinanderzusetzen, bedanken.



# Lebenslauf

**Name:** Sandra Müller  
**Geburtsdatum/ -ort:** 30. August 1971, Dieburg  
**Adresse:** Donnersbergstr. 7c, 64665 Alsbach-Hähnlein  
**Familienstand:** ledig

**Schulbildung:**  
 1978-1980 Schule auf der Aue (Grundschule) in Münster  
 1980-1982 Hahner Schule (Grundschule)  
 1982-1984 Friedrich-Ebert-Schule (Förderstufe) in Pfungstadt  
 1984-1991 Gymnasium Gernsheim  
 1991 Abitur

**Praktikum:**  
 September 1991-  
 Februar 1992 Praktikum im Pflegedienst am Kreiskrankenhaus Bergstraße  
 in Heppenheim

**Studium:**  
 SS 1992-  
 WS 1998/99 Studium der Humanmedizin an der Johann Wolfgang Goethe-  
 Universität in Frankfurt am Main

**Prüfungen:**  
 April 1994 Ärztliche Vorprüfung  
 März 1995 Erster Abschnitt der Ärztlichen Prüfung  
 September 1997 Zweiter Abschnitt der Ärztlichen Prüfung  
 April 1999 Dritter Abschnitt der Ärztlichen Prüfung

**Praktisches Jahr:**  
 April 1998-  
 März 1999 Augenklinik, Städtische Kliniken Frankfurt am Main - Höchst  
 Chirurgische Klinik, Kantonsspital Winterthur, Schweiz  
 Innere Medizin, Städtische Kliniken Frankfurt am Main - Höchst

**AiP:**  
 seit 1999 Ärztin im Praktikum an der Augenklinik der  
 Städtischen Kliniken Frankfurt am Main – Höchst

Frankfurt am Main, 20. Juni 2000

Sandra Müller



# Ehrenwörtliche Erklärung

Ich erkläre ehrenwörtlich, daß ich die dem Fachbereich Humanmedizin der Johann Wolfgang Goethe-Universität in Frankfurt am Main zur Promotionsprüfung eingereichte Arbeit mit dem Titel:

Populationsgenetische Untersuchungen der Tetranukleotid Tandem Repeat Polymorphismen  
D11S488, D18S51, D19S246, HUMFIBRA und HUMVWFA31/A

im Institut für Transfusionsmedizin und Immunhämatologie des Blutspendedienstes Hessen des Deutschen Roten Kreuzes, Frankfurt am Main unter der Leitung von Prof. Dr. med. E. Seifried mit der Unterstützung durch Priv. Doz. Dr. med. C. Seidl ohne sonstige Hilfe selbst durchgeführt und bei der Abfassung der Arbeit keine anderen als die in der Dissertation angeführten Hilfsmittel benutzt habe.

Ich habe bisher an keiner in- und ausländischen Medizinischen Fakultät ein Gesuch um Zulassung zur Promotion eingereicht, noch die vorliegende Arbeit als Dissertation vorgelegt.

Frankfurt am Main, 20. Juni 2000

Sandra Müller