

Inaugural-Dissertation

zur Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

Ruprecht – Karls - Universität Heidelberg

vorgelegt von

Diplom-Biologin

Michaela Schanné

aus:

Zweibrücken

Tag der mündlichen Prüfung: 21. Juli 2006

**Microarray-basierte
genotypische Resistenzbestimmung in HIV**

Gutachter: Prof. Dr. Hans-Georg Kräusslich

PD Dr. Stefan Wiemann

Inhaltsverzeichnis

1	Zusammenfassung	1
2	Abstract	2
3	Einleitung	3
3.1	Humanes Immundefizienzvirus (HIV)	3
3.1.1	Allgemeine Grundlagen zur HIV-Therapie	3
3.1.2	Wirkungsmechanismen der anti-retroviralen Medikamentenklassen	3
3.1.3	Resistenzentwicklung	5
3.1.4	Resistenzmechanismen	6
3.1.5	Methoden zur Resistenzbestimmung	7
3.1.6	Indikationen zur Bestimmung der Resistenz	9
3.2	Humane Gene mit Einfluss auf HIV / AIDS	10
3.3	Microarray-Technologie	13
3.3.1	Definition relevanter Begriffe	13
3.3.2	DNA-Microarrays	14
3.3.3	SNP-Nachweis mittels Microarrays	15
3.3.4	Direkte Hybridisierung	15
3.3.5	Sortierungsreaktion	16
3.3.6	Enzymatische Assays	17
3.3.7	Arrayed Primer Extension	17
4	Problemstellung und Zielsetzung	19
5	Material und Methoden	20
5.1	Material	20
5.1.1	Geräte und Laborausstattung	20
5.1.2	Chemikalien und Chemikaliensätze	21
5.1.3	Verbrauchsmaterialien	22
5.1.4	Enzyme	22
5.1.5	Verwendete Puffer und Lösungen	23
5.1.6	Oligodesoxynukleotide	24
5.1.7	Probenmaterial	25
5.2	Methoden	26
5.2.1	Isolierung genomischer DNA aus Blut	26
5.2.2	Isolierung viraler RNA	26
5.2.3	Reverse Transkription	27
5.2.4	Polymerase-Kettenreaktion	28
5.2.5	Agarosegel-Elektrophorese	29
5.2.6	Isolierung von DNA-Amplifikaten aus PCR-Ansätzen	30
5.2.7	Photometrische Konzentrationsbestimmung von Nukleinsäurelösungen	30
5.2.8	Enzymatischer Verdau der PCR-Produkte	31
5.2.9	Herstellung der Microarrays	32
5.2.10	Blocken der Oberfläche	32
5.2.11	Vorbereitung der Microarrays	33
5.2.12	Arrayed Primer Extension (APEX)	34
5.2.13	Detektion der elongierten Oligonukleotide	35
5.2.14	Quantifizierung der Signalintensitäten	36
5.2.15	Auswertung der Daten	36
6	Ergebnisse	37
6.1	Datenbankrecherche und Auswahl der Sequenzen	37
6.1.1	Auswahl der nachzuweisenden Positionen	37

6.1.2	Sequenzverarbeitung	38
6.2	Chiplayout	41
6.2.1	Festlegen der Sequenzgrundlage	41
6.2.2	Erweiterung um beschriebene Mutationen	43
6.3	Optimierung der APEX-Reaktion für HIV-Sequenzen	45
6.3.1	Temperaturprofil	46
6.3.2	Spottingpuffer	49
6.3.3	Oberflächen	50
6.3.4	Aminomodifikation und Sondenlänge	51
6.3.5	Nachbehandlung der Microarrays	53
6.4	Kontrollsonden	54
6.4.1	Positionskontrollen	54
6.4.2	Extensionkontrollen	55
6.5	Degenerierte Sonden	57
6.5.1	„Wobble“-Modifikationen	58
6.5.2	Inosin	61
6.6	Sondendesign	61
6.7	Protease-Microarray	62
6.7.1	Auswahl der Patientenproben	62
6.7.2	Vorbereitung der Proben	63
6.7.3	Evaluierung des Microarrays	64
6.7.4	Optimierung des Chiplayouts	68
6.7.5	Reproduzierbarkeit der Ergebnisse	69
6.7.6	Detektion von Minoritäten	71
6.7.7	Untersuchung von Proben mit komplexen Mutationsmustern	72
6.8	Reverse Transkriptase-Microarray	73
6.8.1	Auswahl der Patientenproben	73
6.8.2	Reaktionsbedingungen	74
6.8.3	Evaluierung	75
6.8.4	Überprüfen der Referenzsequenzen	77
6.9	Genotypisierung von Proben verschiedener Subtypen	78
6.10	Auswertung der Daten	79
6.11	HIV-assoziierte Gene (Pilotprojekt)	84
6.11.1	Multi Drug Receptor Gene 1 (MDR1)	84
6.11.2	CCR5 Δ 32	86
7	Diskussion	88
7.1	Optimierungsversuche	88
7.2	Kontrollsonden	89
7.3	Verwendung degenerierter Sonden	90
7.4	PR / RT-Microarray - mögliche Fehlerquellen	92
7.5	Detektion von Mischungen	95
7.6	Datenanalyse	95
7.7	Vergleich alternativer Methoden	98
7.8	Genotypisierung von nicht-Subtyp B-Proben	102
7.9	Kosten der Genotypisierung	104
7.10	Genetischer Hintergrund (Pilotprojekt)	105
8	Literaturverzeichnis	106
9	Eigene Publikationen	117
10	Abkürzungen	118
11	Anhang	120

1 Zusammenfassung

Eines der Hauptprobleme der antiviralen Therapie ist das Entstehen resistenzassoziierter Mutationen. Der Nachweis solcher Mutationen ist für eine erfolgreiche Therapie eine entscheidende Voraussetzung. In dieser Arbeit wurde ein Microarray zur Detektion von relevanten Mutationen innerhalb der Protease (PR) und –Reversen Transkriptase (RT) des humanen Immundefizienzvirus (HIV) entwickelt und validiert. Als zeit- und kostengünstigere Alternative zur routinemäßig eingesetzten Sanger-Sequenzierung wurde die Arrayed Primer Extension (APEX)-Methode zur genotypischen Resistenzbestimmung etabliert. Zur Abfrage von 26 resistenzassozierten Codons innerhalb der PR- und 33 Codons innerhalb der RT-kodierenden Region wurden Oligonukleotide erstellt. Aufgrund der enormen Variabilität des HIV-Genoms berücksichtigen diese alle Sequenzvarianten, die mit einer Frequenz von >5% (PR) bzw. >10% (RT) in der Los Alamos-Datenbank verzeichnet sind. Um diese Sequenzabdeckung technisch und finanziell zu realisieren, wurde die Verwendung degenerierter Oligonukleotidsonden für APEX etabliert. Für die Validierung des PR- und RT-Microarrays wurden 94 bzw. 48 PCR-Produkte aus Patientenproben mit bekannter Sequenz analysiert. Dafür wurden die APEX-Reaktionsbedingungen erfolgreich angepasst. Die Vergleichssequenzen wurden durch Sanger-Sequenzierung ermittelt. 93% (PR) bzw. 88% (RT) aller untersuchten Positionen wurden mittels APEX übereinstimmend bestimmt. Der Anteil an Positionen, für die aufgrund schwacher Signalintensitäten keine Sequenzzuweisung vorgenommen werden konnten, liegt bei 3% (PR) und 5% (RT). Es wurde gezeigt, dass dieses Ergebnis durch Optimierung der Sonden deutlich verbessert wird. Die zufällig ausgewählten Proben umfassten sowohl HIV-Subtyp B- als auch nicht-Subtyp B-Sequenzen. Aufgrund der verwendeten Sonden ist die Übereinstimmung für die untersuchten Subtyp B-Proben im Vergleich mit den nicht-Subtyp B-Proben höher (94% / 88% für PR; 88% / 84% für RT). Sequenzmischungen wurden qualitativ gut detektiert, eine Quantifizierung ist aber Sonden-abhängig. Der Fehlerbereich ist vergleichbar mit dem anderer Genotypisierungsmethoden. Das entwickelte Testverfahren kann kostengünstig mit hohem Probendurchsatz durchgeführt werden. Parallel zur HIV-Genotypisierung wurde anhand zweier Beispiele gezeigt, dass auch Mutationen im Wirtsgenom, die mit der Krankheitsprogression assoziiert sind, sehr gut mittels APEX nachgewiesen werden können. Dies wird künftig eine sinnvolle Ergänzung des Microarrays sein und eine individuellere Therapie ermöglichen.

2 Abstract

Mutations that confer drug resistance are a major obstacle in antiviral therapy. The detection of such mutations in the protease (PR) and reverse transcriptase (RT) of human immunodeficiency virus (HIV) is an important step in the successful treatment of HIV infection. A standard method involved in mutation detection, Sanger sequencing, has proven to be costly and temporally disadvantageous. In order to overcome such inadequacies, a genotypic resistance assay based on APEX has been developed and established.

Oligonucleotides were designed for the detection of 26 resistance related codons in PR and 33 codons in RT coding region. Due to the immense sequence heterogeneity of the viral genome, primers were selected to map sequence variants that representing >5% PR and >10% RT in the Los Alamos Sequence Databank. The application of degenerated primers to APEX was successfully established to cover all these sequence combinations. For microarray validation 94 PR and 48 RT patient derived PCR products were analysed. As reference sequences, information of these samples were previously obtained by Sanger sequencing. Concordance between the reference method and APEX was high, at 93% and 88% for PR and RT, respectively. Due to weak signal intensities it was impossible to determine the sequence via primer extension for 3% PR and 5% RT of positions analyzed. Efficiency of detection was improved by modification of the chip layout by optimization of oligonucleotides, demonstrated by means of the protease microarray. The analyzed patient set was randomly chosen and contained HIV-subtype B as well as non-subtype B samples. Currently, non-subtype B samples could be investigated by PR and RT microarrays. Concordance with standard sequencing is lower than for subtype B sequences (88%/94% for PR; 84%/88% for RT). Furthermore the microarrays provided a confident detection of sequence mixtures in a qualitative manner; however quantification is dependent on site-specific oligonucleotides. The developed APEX assay achieves similar results compared with other methods for genotypic resistance testing in HIV. The developed microarrays allow a cost- and time-effective genotypic resistance testing in a high-throughput manner.

In addition to HIV resistance testing it is useful to acquire information about several host genes which are involved in the progression to AIDS. For two examples it was shown that relevant mutations can be very well detected by APEX. The parallel analysis of such mutations will enable a more individualized therapy.

3 Einleitung

3.1 Humanes Immundefizienzvirus (HIV)

3.1.1 Allgemeine Grundlagen zur HIV-Therapie

Die Behandlung der HIV-Infektion wird unter dem Begriff Antiretrovirale Therapie (ART) zusammengefasst, deren Ziel eine möglichst vollständige Hemmung der Virusreplikation ist, das heißt, eine Senkung der Viruslast unter eine Nachweisgrenze von etwa 20 HIV-Genomkopien pro Milliliter Blut (Raboud *et al.*, 1998). Damit einhergehend sollen die Krankheitsprogression aufgehalten, die HIV-assoziierten Symptome reduziert und eine klinisch relevante Immunrestitution erreicht werden. Die wichtigsten Laborparameter für die Verlaufsbeurteilung einer HIV-Infektion sind die quantitative Bestimmung der CD4+-Lymphozyten und der HIV-RNA (Mellors *et al.*, 1997).

Im Rahmen der antiviralen Behandlung können derzeit Medikamente aus vier verschiedenen Wirkstoffklassen eingesetzt werden: Nukleosid- und Nukleotidanaloga (NRTIs) und Nicht-nukleosidische Reverse Transkriptase-Inhibitoren (NNRTIs), Inhibitoren der Protease (PIs) und seit Mai 2003 Fusionsinhibitoren mit dem Prototyp T-20. Damit beläuft sich die Zahl der zugelassenen Präparate auf über 20 (Chang & Tyring, 2004). Durch die Anwendung dieser Substanzen in einer Kombinationstherapie gegenüber einer Monotherapie kann die Behandlung und Prognose HIV-infizierter Patienten verbessert werden (Palella *et al.*, 1998, Egger *et al.*, 1997, Egger *et al.*, 2002).

3.1.2 Wirkungsmechanismen der anti-retroviralen Medikamentenklassen

NRTIs

Ziel aller RT-Inhibitoren (NRTIs wie NNRTIs) ist es, die Umschreibung des viralen RNA-Genoms in doppelsträngige DNA zu hemmen. Die reverse Transkription ist die Voraussetzung der Integration in das Wirtsgenom.

Die Wirkungsweise der so genannten Nukleosidischen Reverse Transkriptase-Inhibitoren (NRTIs) liegt in ihrer Fähigkeit, als Nukleosidanaloga mit den natürlichen Nukleosiden in Konkurrenz zu treten. Ihr Einbau während der Synthese verursacht einen Abbruch der

Nukleinsäure, bedingt durch das Fehlen einer freien OH-Gruppe am 3'-Ende des Ribosemoleküls. Nukleosidanaloga werden erst durch die intrazelluläre Phosphorylierung zu Triphosphatderivaten wirksam (Arts & Wainberg, 1996). Alternativsubstrate zu allen vier physiologischen Nukleotiden kommen zum Einsatz: AZT und D4T sind Thymidin-Analoga, DDC, FTC und 3TC imitieren Cytidin. Bei DDI handelt es sich um ein Inosin-Analogon, das in Didesoxyadenosin umgewandelt wird, und Abacavir wird anstelle von Guanosin eingebaut. Ein Nachteil dieser Substanzklasse liegt in der mitochondrialen Toxizität als Langzeitnebenwirkung (Brinkman *et al.*, 1999). Durch den Einbau der alternativen Substanzen wird der Zellstoffwechsel der Organellen gestört, es kommt zur Degeneration.

NNRTIs

Wie die NRTIs greifen die NNRTIs ebenfalls die virale Reverse Transkriptase an. Der Funktionsmechanismus der NNRTIs unterscheidet sich allerdings dahingehend, dass diese nicht als Alternativsubstrate eingebaut werden. Sie inhibieren das Enzym durch direkte, nicht-kompetitive Bindung an einen Bereich nahe der Substratbindungsstelle. In der Folge kommt es zur Bildung eines Komplexes, durch den das katalytisch aktive Zentrum blockiert wird und einer deutlichen Verlangsamung der Polymersisation. Derzeit sind drei Substanzen erhältlich (Nevirapin, Delavirdin und Efavirenz), die in den Jahren zwischen 1996 und 1998 ihre Marktzulassung erhielten (Übersichtsartikel Menendez-Arias, 2002).

Protease-Inhibitoren

Die aktive Protease ist ein Dimer aus identischen Proteinuntereinheiten, die aus 99 Aminosäuren bestehen. Sie ist für die Prozessierung viraler Makromoleküle, der so genannten Gag- und Gag/Pol-Vorläuferproteine, verantwortlich. Wird das Enzym gehemmt, unterbleibt die proteolytische Spaltung, in der Folge entstehen Viruspartikel, die ihre Infektiosität verloren haben. In dem Wissen um die Molekülstruktur wurden Anfang der 90er Jahre Substanzen entwickelt, die durch Bindung an das aktive Zentrum in der Lage sind, das Enzym kompetitiv zu hemmen.

Protease-Inhibitoren haben seit 1995 die Behandlung der HIV-Infektion revolutioniert. In mehreren großen Studien wurde die Wirksamkeit der ersten Substanzen erfolgreich nachgewiesen (z.B. Hammer *et al.*, 1997; Cameron *et al.*, 1998).

Fusionsinhibitoren

Der Eintritt des Virus in die CD4-Zelle bietet theoretisch drei verschiedene Angriffsstellen: die Bindung an den CD4-Rezeptor, an die Korezeptoren CCR5 bzw. CXCR4 und die Fusion des Virus mit der Wirtszelle. Substanzen, die die Wechselwirkung zwischen CD4-Rezeptor und Virus hemmen, bezeichnet man als *Attachment*-Inhibitoren, Korezeptor-Antagonisten verhindern eine Anlagerung an die Korezeptoren. Fusionsinhibitoren binden reversibel oder irreversibel an eine Domäne des viralen Transmembranproteins gp41 und hemmen so die für den Eintritt in die CD4-Zelle notwendige Konformationsänderung. Diese Wirkstoffklassen werden unter dem Begriff *Entry*-Inhibitoren zusammengefasst. Allein T-20 (Enfuvirtid, Fuzeon®) als Prototyp des Fusionsinhibitors hat derzeit eine Zulassung auf dem europäischen Markt. Es handelt sich dabei um ein Peptid mit einer Größe von 36 Aminosäuren und muss subkutan injiziert werden. Sein Wirkungsmechanismus begründet sich auf der Fähigkeit, an eine Intermediärstruktur des gp41 zu binden, die bei der Fusion des Virus mit der Zelle entsteht (Lalezari *et al.*, 2003, Lazzarin, 2005).

3.1.3 Resistenzentwicklung

Entstehung von Mutationen

Das Auftreten resistenter Virusvarianten wurde erstmals 1989 beschrieben (Larder *et al.*, 1989) und erschwert seitdem zunehmend den erfolgreichen Einsatz der antiretroviralen Substanzen. Resistenzen gegen eingesetzte Wirkstoffe sind der Hauptgrund für ein Therapieversagen (D'Aquila *et al.*, 1995; Erali *et al.*, 2001; Hirsch *et al.*, 1998). Triebfeder der Resistenzentwicklung sind Mutationen im HIV-Genom, die durch verschiedene Mechanismen verursacht werden. Zentrale Stellungen in der Virusreplikation nehmen zwei Enzyme ein: die virale Reverse Transkriptase, die für die Umschreibung der RNA-Genome in cDNA verantwortlich ist, und die zelluläre RNA-Polymerase II, deren Aufgabe darin besteht, anhand der DNA-Matrize die mRNA-ähnlichen Virusgenome zu synthetisieren. Beiden Enzymen fehlt es an einer 3'-Exonukleaseaktivität, daher kommt es zu einer hohen Fehlerrate (1/1000) während des Einbaus von Nukleotiden (Preston *et al.*, 1988, Roberts *et al.*, 1988). In der Folge manifestieren sich Mutationen im HIV-Genom. Hinzu kommt eine hohe Replikationsrate von $10,3 \times 10^9$ Viren pro Tag in Abwesenheit von Inhibitorsubstanzen (Perelson, *et al.*, 1996). Allein die Kombination von Fehler- und Replikationsrate erzeugt bei einer Genomgröße von 10 500 Basen jede mögliche Mutation

mehrfach pro Tag. Weiterhin besteht die Möglichkeit der genetischen Rekombination zwischen verschiedenen Varianten (Moutouh *et al.*, 1996), die als Grundlage für die Entstehung von Mutationen angesehen werden.

Resistenzen

Aufgrund dieser Mutationen kann es zu einer verminderten Empfindlichkeit gegenüber Inhibitorsubstanzen kommen (ausführlicher Übersichtsartikel Shafer, 2002). Schon bald nach dem Einführen der ersten Medikamente in der anti-retroviralen Behandlung wurden die ersten Resistenzbildungen beobachtet (Larder *et al.*, 1989, Larder & Kemp, 1989, Kaplan *et al.*, 1994; Condra *et al.*, 1995).

Resistenzen gegen RT-Inhibitoren können hervorgerufen werden durch das Auftreten von einer oder weniger Mutationen, die die Replikationsfähigkeit der Viren nicht zwangsweise beeinträchtigen (Boucher *et al.*, 1993, Goudsmit *et al.*, 1996, Clavel *et al.*, 2000). Eine PI-Resistenz unter weitgehender Aufrechterhaltung der viralen Fitness entwickelt sich erst durch eine Akkumulation mehrerer Mutationen, man spricht deshalb auch von einer so genannten genetischen Barriere (Condra *et al.*, 1996; Condra *et al.*, 1998; Nijhuis *et al.*, 1999, Mammano *et al.*, 2000). Die Entstehung resistenter Virusvarianten ist in der Regel ein schrittweiser Prozess. Um dem vorzubeugen ist es wichtig, die Virusreplikation möglichst vollständig zu unterdrücken (Hertogs *et al.*, 1998; Mammano *et al.*, 2000).

3.1.4 Resistenzmechanismen

NNRTIs wirken durch ihre direkte Bindung an das Enzym (3.1.2). Mutationen, die eine Affinitätsverringering zwischen Bindungsstelle und Inhibitor hervorrufen, beeinträchtigen dessen Wirkungsweise. Die Resistenzentwicklung gegenüber NRTIs hingegen begründet sich auf zwei unterschiedliche biochemische Mechanismen (De Mendoza, 2002). Die so genannte **sterische Inhibition** wird durch Mutationen vermittelt, die es dem Enzym ermöglichen, zwischen physiologischen Nukleotiden und NRTIs zu unterscheiden. Der Einbau verschiebt sich zugunsten der dNTPs. Durch die ATP- oder Pyrophosphat-vermittelte **Phosphorylyse** hingegen können bereits eingebaute NRTIs aus dem DNA-Strang wieder freigesetzt werden. Sie wird verursacht durch die Mutationen M41L, D67N, K70R, L210W, T215Y und K219Q innerhalb der Reversen Transkriptase (Meyer, 2000).

Bei Veränderungen, die im Zusammenhang mit Resistenzen gegenüber Protease-Inhibitoren stehen, unterscheidet man so genannte Primär- und Sekundär-Mutationen. **Primärmutaionen** vermitteln den Phänotyp einer Protease-Resistenz, zu ihnen zählen Mutationen, die unter dem Selektionsdruck einer Inhibitors auftreten und solche, die im katalytisch-aktiven Zentrum des Enzyms liegen und die Affinität zwischen PI und Protease herabsetzen. **Sekundärmutationen** kompensieren einen Verlust der viralen Fitness, der durch das Auftreten primärer Mutationen verursacht wird (Shafer, 2002).

Der Fusioninhibitor T-20 blockiert eine für die Aufnahme in die Wirtszelle notwendige Konformationsänderung des viralen Transmembranproteins gp41 (3.1.2). Die genetische Barriere zur Ausbildung einer T-20-Resistenz ist niedrig, das Zentrum der Resistenzbildung umfasst ein Motiv von zehn Aminosäuren (zwischen 36 und 45) im gp41 (Greenberg & Cammack, 2004). Schon ein einzelner Aminosäureaustausch allerdings erlaubt es dem Virus, diesen Mechanismus zu umgehen und die Wirksamkeit des Fusionsinhibitors signifikant zu senken.

3.1.5 Methoden zur Resistenzbestimmung

Generell lassen sich zwei Methoden zur Resistenzmessung gegenüber anti-retroviralen Substanzen unterscheiden: die Bestimmung des Phänotyps oder eine genotypische Untersuchung auf Ebene der Nukleinsäure (Shafer, 2002; Wilson, 2003, Gallant, 2005). Voraussetzung für beide Methoden ist eine Viruslast von mindestens 500-1000 Genomkopien/ml.

Phänotypische Resistenzbestimmung

Ein phänotypischer Resistenztest quantifiziert direkt die Wirkung einer Inhibitorsubstanz auf die Replikation von HIV in Zellkultur oder die *in vitro* Aktivität isolierter viraler Enzyme. Unter dem Selektionsdruck steigender Inhibitorkonzentrationen wird die Replikationskinetik bzw. die Enzymaktivität gemessen und mit der des Wildtyp-Virus verglichen. Dabei wird meist der IC₅₀-Wert bestimmt (50% inhibitory concentration), das heißt die Konzentration, bei der die Virusreplikation um 50% gehemmt wird (Hirsch, 1998). Die Resistenzfaktoren des untersuchten Virus werden mittels Division von IC₅₀ durch die entsprechenden Werte des Referenzstammes ermittelt (Beerenwinkel *et al.*, 2003; Garcia-Lerma & Heneine, 2001). Im Gegensatz zur genotypischen Resistenzbestimmung handelt es sich nicht um eine indirekte

Dateninterpretation, die Resistenzfaktoren leiten sich unmittelbar aus den quantitativen Messergebnissen ab. Das bringt die Notwendigkeit einer Bestimmung analytischer und wenn möglich auch klinischer Grenzwerte mit sich. Die Festlegung dieses so genannten *Cut-offs* ist entscheidend für die Interpretation der Ergebnisse (Demeter & Haubrich, 2001, Haubrich, 2004). Die Nachteile einer Phänotypisierung sind der zeitliche wie technische Aufwand und die hohen Kosten des Verfahrens. Der Prozess unterliegt zudem relativ vielen Störgrößen und ist zum Teil nur schwer reproduzierbar. Kommerziell erhältliche Testsysteme zur phänotypischen Resistenzbestimmung sind zum Beispiel *PhenoSense* (ViroLogic, USA) oder *Antivirogram* (Virco, Belgien). Indikationen für eine phänotypische Resistenzbestimmung sind die Testung von neuen Inhibitoren oder das Überprüfen schwer interpretierbarer, komplexer oder neuer Genotypen (Hertogs *et al.*, 1998; Petropoulos *et al.*, 2000).

Genotypische Resistenzbestimmung

Genotypische Verfahren basieren auf der Analyse Resistenz-assoziiierter Mutationen. Generell können für genotypische Resistenzbestimmungen zwei unterschiedliche experimentelle Ansätze verwendet werden.

Die Sequenzierung nach Sanger (Sanger *et al.*, 1977) ist ein stabiles Verfahren, die Ergebnisse lassen sich im Vergleich zu phänotypischen Test gut reproduzieren (Demeter *et al.*, 1998, Korn *et al.*, 2003). Sie liefert die vollständige Sequenzinformation der untersuchten Abschnitte, in der Regel zumeist Protease und die für die Resistenzentstehung notwendigen Bereiche der Reversen Transkriptase (bis Aminosäure 340). Ein Vergleich mit einer Wildtyp-Referenzsequenz identifiziert alle vorhandenen Mutationen, unabhängig davon, ob sie Resistenz-assoziiert sind. Punkt-Mutations-Assays hingegen suchen gezielt nach Veränderungen bestimmter Positionen, die bekanntermaßen mit einer Resistenzentwicklung in Verbindung stehen (Gonzalez *et al.*, 2004, Lipshutz *et al.*, 1995; Sugiura *et al.*, 2003; Vahey *et al.*, 1999).

Genotypisierungen liefern aber in jedem Fall nur den indirekten Nachweis einer Resistenz gegenüber anti-retroviralen Wirkstoffen. Für die meisten dieser Substanzen sind Mutationen beschrieben, die zu einer verminderten, oder aber auch erhöhten Sensibilität führen können. Komplexe Resistenzmuster, die auch kompensierende, sekundäre Mutationen enthalten können, erschweren oftmals die Beurteilung des Resistenzgrades (Demeter & Haubrich, 2001; Sarmati *et al.*, 2002). Zudem stellt sich bei Einführung neuer Wirkstoffe das Problem, dass zunächst noch keine Daten über bestehende Zusammenhänge zwischen Mutationsmustern

und Resistenzen verfügbar sind. Grundlage für die Interpretation der Mutationsmuster ist die Korrelation zwischen Geno- und Phänotyp, wobei die Interpretationsrichtlinien nur den jeweiligen Wissensstand über die Interaktion zwischen Mutation und Resistenz widerspiegeln (Haubrich, 2004). Um ein individuelles, genotypisches Mutationsmuster einem Phänotyp zuordnen zu können, kann mittels entsprechender Computerprogramme (Beerenwinkel *et al.*, 2003) oder in Datenbanken mit vergleichbaren Resistenzbefunden verglichen werden. Man spricht dabei dann von einem so genannten *virtuellen Phänotyp*. Eine zu diesem Zweck häufig genutzte Datenbank ist z.B. die Stanford HIV-Datenbank (<http://www.hivdb.stanford.edu/>).

Die Interpretationsrichtlinien basieren auf Ergebnissen paarweise durchgeführter Geno- und Phänotypisierungen, sind allerdings nicht standardisiert. So werden unter Umständen gleichen Genotypen unterschiedliche Resistenzprofile zugeordnet (Erali *et al.*, 2001; Korn *et al.*, 2003).

Der Grad der Resistenz wird meist in verschiedene Kategorien eingestuft und reicht von sensibel über reduzierte Empfindlichkeit und intermediäre Resistenz bis zur hochgradigen Resistenz.

3.1.6 Indikationen zur Bestimmung der Resistenz

Als Indikation für einen Resistenztest gilt derzeit (neben anderen Gründen) eine Umstellung der Behandlung infolge Therapieversagen. Das gilt besonders dann, wenn eine umfangreiche Vorbehandlung des Patienten gegeben ist. (Hirsch *et al.*, 2000; Vandamme *et al.*, 2004; Deutsch-Österreichische Leitlinien zur antiretroviralen Therapie der HIV-Infektion, 2005).

Im Verdacht einer Infektion mit einer bereits resistenten Virusvariante, sollte ebenfalls eine Resistenzbestimmung durchgeführt werden. Die Übertragung resistenter Viren gilt inzwischen als bewiesen, ebenso wie die Stabilität der Resistenz auch ohne den Selektionsdruck einer antiviralen Behandlung. Bei ca. 10% der Neuinfektionen zeigen sich bei den übertragenen Viren Resistenzmutationen. Deshalb ist eine prinzipielle Resistenzbestimmung vor Einleitung der ersten Therapie zu empfehlen (Boden *et al.*, 1999; Cane, 2005, Duwe *et al.*, 2001; Little *et al.*, 1999; Yerly *et al.*, 1999). Dies trifft besonders für schwangere HIV-Patientinnen zu. Vor Einleitung einer ersten Therapie in der Schwangerschaft sollte grundsätzlich das Resistenzprofil bestimmt werden um die

Wirksamkeit der Transmissionsprophylaxe nicht zu vermindern (Deutsch-Österreichische Empfehlungen zur HIV-Therapie in der Schwangerschaft, 2003).

Die klinische Relevanz der Resistenzbestimmung vor Therapieumstellung konnte in mehreren Studien gezeigt werden. Die Behandlung von Patienten, deren Therapie auf bekannten Resistenzdaten basiert, konnte deutlich verbessert werden (Hirsch *et al.*, 1998; Wegner *et al.*, 2004; Weinstein *et al.*, 2001). Dies gilt sowohl für genotypische (Durant *et al.*, 1999), wie auch für phänotypische Resistenztests.

Obwohl die Kosten für eine genotypische Resistenzbestimmung etwa nur die Hälfte einer Phänotypisierung betragen, belaufen auch sie sich je nach Labor auf 300-500 Euro (Richman, 2004). Sequenzierungen nach Sanger, wie sie zu Genotypisierungszwecken angewendet werden, sind im Vergleich zu anderen routinediagnostischen Verfahren relativ zeit- und kostenintensiv. Eine Vereinfachung und Kostensenkung könnten möglicherweise Microarray-basierte Analysen bieten, deren besondere Attraktivität in der Durchführung von Hochdurchsatz-Untersuchungen liegt (Stürmer *et al.*, 2004).

3.2 Humane Gene mit Einfluss auf HIV / AIDS

Neben Resistenz-assoziierten Mutationen im HIV-Genom spielt auch der genetische Hintergrund des Patienten eine wichtige Rolle hinsichtlich der Wirkungsweise antiviraler Substanzen sowie der Krankheitsprogression. Die Literatur beschreibt unterschiedliche Gene, deren Ausprägungen den Eintritt des Virus in die Wirtszelle, die Antwort des Immunsystems oder den zellulären Transport antiviraler Medikamente beeinflusst (Nelson, 1999, O'Brien & Nelson, 2004).

Chemokin-Rezeptoren sind Oberflächen-Proteine, deren Peptid-Liganden, die Chemokine, die die Migration der Rezeptor-tragenden Zellen vermitteln. CCR5 wird zusammen mit anderen Chemokin-Rezeptoren auf dem kurzen Arm von Chromosom 3 kodiert. So genannte R5-Isolate von HIV-1 nutzen CCR5 als Ko-Rezeptor zur Infektion von Makrophagen und primären T-Zellen. Aufgrund ihres Zelltropismus davon zu unterscheiden sind X4-Isolate, denen der Ko-Rezeptor CXCR4 zum Eintritt in primäre T-Zellen und T-Zell-Linien dient. Eine Deletion von 32 Basenpaaren im CCR5-Gen führt zur Bildung eines nicht funktionellen Proteins, das auf der Zelloberfläche nicht nachzuweisen ist (Liu *et al.*, 1996). Homozygoten Trägern fehlt der für die Infektion notwendige Ko-Rezeptor, Heterozygote exprimieren im

Vergleich zum Wildtyp weniger als die Hälfte an Rezeptoren auf der Oberfläche der lymphoiden Zellen. In Abhängigkeit von homo- oder heterozygoter Ausprägung vermittelt diese Mutation eine fast vollständige Resistenz gegen die R5-HIV-1-Infektion oder doch zumindest die Verzögerung des Krankheitsverlaufs.

CCR5 Δ 32 geht auf ein einzelnes, natürliches Mutationsereignis zurück, das vermutlich vor einigen Jahrtausenden in Nordwesteuropa stattgefunden hat (Libert *et al.*, 1998). Der Genotyp CCR5 Δ 32/+ kommt mit einer erstaunlich hohen Prävalenz von ~20% und homozygot mit ~1% in der kaukasischen Bevölkerung vor (Carrington, *et al.*, 1999, Dean *et al.*, 1996, Liu *et al.*, 1996, O'Brien *et al.*, 1999).

Das Multi Drug Resistance Transporter 1 (MDR1)-Gen kodiert für ein Glykoprotein, das verantwortlich ist für den Membrantransport verschiedenster Substrate auf Kompartiment- und Zell-Ebene und somit auch Einfluss nimmt auf die Aufnahme antiviraler Medikamente aus dem Verdauungssystem.

Fellay *et al.* (2002) beschreiben einen Zusammenhang zwischen dem MDR1-Polymorphismus (MDR1 3435 C/T), dem Expressionsgrad und der CD4Zellzahl bei Patienten, die mit den Protease-Inhibitoren Nelfinavir und Efavirenz behandelt wurden. Die Überexpression von MDR1 führt zu einer Verminderung des intrazellulären Medikamentenspiegels. Ursache dafür ist vermutlich der Export des Inhibitors aus dem Zytoplasma durch das Glykoprotein. Träger der Mutation 3435 TT zeigen eine niedrigere MDR1-Expression, höhere Plasmaspiegel und CD4-Zellzahlen. Auch Patienten, die heterozygot MDR1 3435 C/T sind, zeigten im Vergleich mit Wildtyp-Trägern (3435 CC) unter Therapie mit den genannten PIs höhere CD4-Zahlen. Der Polymorphismus hat eine unmittelbare Auswirkung auf den Expressionsgrad des Genproduktes und somit direkten pharmakogenetischen Einfluss auf die Therapie. Es handelt sich dabei um eine Punktmutation, die sich für eine Detektion mittels APEX eignet.

Eine wichtige Rolle spielen auch die HLA-Gene (*Human Leucocyte Antigen*), die im Haupthistokompatibilitätskomplex (MHC) kodiert sind. Diese Moleküle präsentieren gegenüber CD8-T-Zellen Peptide, die im Cytosol aus Antigenen abgespalten (Klasse I) oder CD4-T-Zellen Peptide, die in zellulären Vesikeln abgebaut wurden (Klasse II) und spielen somit eine wichtige Rollen in der Immunabwehr. Der MHC enthält an verschiedenen Loci verschiedene Allele, er ist das am meisten polymorphe Gencluster im menschlichen Genom. Im Zusammenhang mit AIDS sind für eine Vielzahl verschiedener Ausprägungen dieser Gene sowohl eine verlangsamte als auch beschleunigte Progression im Krankheitsverlauf beschrieben (z.B. Al Jabri, 2001, Carrington & O'Brien, 2003, Uyl *et al.*, 2004). Auch die

protektive Rolle bestimmter HLA-Polymorphismen gegenüber einer HIV-Infektion wird diskutiert (Carrington & O'Brien, 2003). Generell führt eine vollständig heterozygote Ausprägung der HLA-Klasse I-Gene zu einer Verlangsamung der Progression, während Träger homozygoter Allele den Nachteil eines schnelleren Fortschreitens der Krankheit aufweisen (Carrington *et al.*, 1999).

Zu nennen sind in diesem Zusammenhang im Weiteren die Gene, die den *Killer-Immunoglobulin-like-Rezeptor* (KIR) auf natürlichen Killerzellen (NK) kodieren. Die Aufgabe dieser Rezeptoren liegt darin, MHC-1-negative Zellen zu eliminieren. Mit der Reduktion von MHC-1-Molekülen auf der Oberfläche infizierter Zellen versuchen Viren, dem Immunsystem zu entkommen. Die Aktivierung der NK-Zelle erfolgt durch eine fehlende Bindung von KIR an den MHC-1-Komplex. Für bestimmte KIR-Allele wurde eine verlangsamte Progression zu AIDS beschrieben (Martin *et al.*, 2002).

Neben diesen Beispielen existieren noch zahlreiche andere Faktoren, die Einfluss auf den Verlauf der HIV-Infektion haben können (Übersichtsartikel: z. B. Carrington *et al.*, 2001, O'Brien & Nelson, 2004).

Generell ist es im Rahmen einer antiviralen Therapie auch von Interesse, neben Resistenz-assoziierten Mutationen im HIV-Genom auch den genetischen Hintergrund des Patienten hinsichtlich der beschriebenen Gene zu bestimmen. Für beide Fragestellungen bieten sich neben der Sequenzierung nach Sanger auch Microarray-basierte Analysemethoden an, die gegebenenfalls eine parallele Untersuchung erlauben.

3.3 Microarray-Technologie

3.3.1 Definition relevanter Begriffe

Miniaturisierte Träger, auf denen Moleküle in hoher Anzahl und Dichte immobilisiert sind, bezeichnet man als **Biochips**. Der Begriff hat sich aufgrund der angestrebten Miniaturisierung in Anlehnung an den Mikrochip aus der Halbleiter- und Mikrosystemtechnik eingebürgert (Ermantraut, 1999). Die rasterartige Anordnung der Moleküle wird als **Array** oder, aufgrund des kleinen Formats, als **Microarray** bezeichnet.

Bei den immobilisierten Molekülen (**Sonden**) handelt es sich meist um DNA, DNA-Analoga wie PNA oder LNA, Peptide, Proteine oder Antikörper. Kommen Nukleinsäuren zum Einsatz, unterscheidet man aufgrund von Art und Länge der Sequenz zwischen **Oligonukleotid-Arrays** und **cDNA-Arrays**.

Das Aufbringen von DNA-Sonden auf das Trägermaterial kann in unterschiedlicher Weise erfolgen. Generell stehen dafür zwei Möglichkeiten zur Verfügung. Zum einen kann die DNA *in situ* unter Verwendung photolithographischer Masken synthetisiert werden (Lipshutz *et al.*, 1995). Aufgrund der Ausbeute und Qualität, die die Länge der Sonden limitieren, werden mit diesem Verfahren ausschließlich Oligonukleotid-Arrays generiert. Der Vorteil der Methode liegt darin, dass eine hohe Sondendichte erreicht werden kann. Die zweite Möglichkeit besteht im Transferieren vorsynthetisierter Moleküle mit Mikrodispensierungsnadeln durch Kontakt- oder kontaktfreie Verfahren (**Spotting**). Auf der Trägeroberfläche liegen die in Puffer gelösten Sonden als kleinste Tropfen vor, die als **Spots** bezeichnet werden. Häufig sind die Sonden mit einem so genannten **Spacer** versehen. Er dient einerseits der Kopplung an die Oberfläche, andererseits soll er aus sterischer Sicht die Zugänglichkeit der Sonden für eine Hybridisierung oder enzymatische Prozessierung erleichtern.

Aufgrund der hohen Form- und Thermostabilität und wegen der geringen Eigenfluoreszenz wird in der Herstellung von Microarrays derzeit bevorzugt Glas als Trägermaterial eingesetzt. Aber auch andere Substanzen wie Silizium oder Gold finden ihre Anwendung. Das Immobilisieren der Sonden, man spricht dabei auch von **Kopplung**, erfordert in der Regel eine chemische Modifikation der Oberfläche. Je nach Fragestellung muss die Beschichtung unterschiedliche Eigenschaften aufweisen. In jedem Fall ist aber eine Abstimmung zwischen der Oberfläche, den zu koppelnden Sonden und dem Puffer, in dem sie gelöst sind,

erforderlich. Zur Beschichtung eignen sich beispielsweise Silane, Aldehyde oder Poly-L-Lysin.

3.3.2 DNA-Microarrays

Die Microarray-Technologie macht sich das Grundprinzip des nach seinem Erfinder benannten Southern-Blots zunutze (Southern, 1975). In abgewandelter Form basiert sie auf der Hybridisierung einer unbekanntes Ziel-DNA mit Nukleinsäuremolekülen, die in bekannter Anordnung im Raster auf der Chipoberfläche immobilisiert sind. Generell lassen sich aufgrund der Art und Länge der verwendeten DNA-Moleküle Oligonukleotid- und cDNA-Arrays unterscheiden.

Am weitesten verbreitet ist der Einsatz von cDNA-Microarrays zur Studie differentieller Genexpression. Das Spektrum reicht von der Untersuchung von Modellorganismen wie Hefe oder *Drosophila* bis hin zur Typisierung verschiedenster Tumorgewebe. Solche Experimente basieren zumeist auf cDNA-Bibliotheken, in denen die klonierten cDNA-Fragmente mittels PCR amplifiziert und anschließend auf den Glasobjektträger aufgebracht werden. Die isolierte mRNA des betreffenden Organismus oder Gewebe wird markiert und gegen die immobilisierten Sonden auf dem Microarray hybridisiert. Mindestens zwei mRNA-Fractionen, die die exprimierten Gene unter bestimmten Konditionen repräsentieren, müssen für ein solches Experiment miteinander verglichen werden. Je höher die Expression eines Gens ist, desto größer die Transkriptmenge der zugehörigen mRNA-Spezies in der isolierten Gesamt-mRNA und dementsprechend stärker ist das zugehörige Hybridisierungssignal auf dem DNA-Chip. Klassischerweise werden zur Markierung der zu vergleichenden mRNA-Fractionen die Farbstoffe Cyanine5 und Cyanine3 verwendet, man spricht dabei auch von so genannten *Rot-Grün-Experimenten*. Die Auswertung solcher Versuche bedarf aufgrund ihrer Komplexität einer umfangreichen, computergestützten Datenanalyse (z.B. Fellenberg *et al.*, 2001, Noordewier & Warren, 2001).

3.3.3 SNP-Nachweis mittels Microarrays

Im Gegensatz zu cDNA-Chips finden Oligonukleotid-Microarrays ihre Anwendung zumeist in der Mutationsanalyse (Hacia, 1999), insbesondere dem Nachweis von Einzelnukleotidmutationen, so genannter *SNPs* (*single nucleotide polymorphisms*). Generell bieten dabei Array-basierte Methoden die Möglichkeit zur parallelen Analyse vieler Ansätze. Insbesondere hochauflösende Chips, bei denen durch *in situ*-Synthese der Oligonukleotide eine sehr hohe Sondendichte erreicht wird, erfüllen dieses Kriterium (Affymetrix). Je nach Fragestellung können entweder eine Vielzahl an Mutationen in einer einzelnen Ziel-DNA, oder aber wenige Mutationen in einem Gemisch verschiedener Proben untersucht werden. Die Diskriminierung wird entweder durch Hybridisierung oder enzymatische Reaktion vermittelt, wobei diese enzymatischen Reaktionen auch separiert in einer Flüssigphase ablaufen und im Anschluss auf dem Microarray nachgewiesen werden können (Sortierung). Die Sichtbarmachung der Diskriminierungsprodukte basiert zumeist auf Fluoreszenzdetektion. Es lassen sich in diesem Zusammenhang eine Vielzahl an Methoden unterscheiden. Einen guten Überblick geben zum Beispiel Syvänen (1999 und 2001) oder Gut (2001). Im Folgenden sollen nur einige Ansätze kurz beschrieben werden.

3.3.4 Direkte Hybridisierung

Die direkte Hybridisierung macht sich den Effekt zunutze, dass vollständig komplementär gepaarte DNA-Hybride (*perfect match*) eine größere Stabilität besitzen als solche mit einer Einzelnukleotid-Fehlpaarung (*mismatch*). Solche Fehlpaarungen haben den größten destabilisierenden Einfluss, wenn sie sich in der Sequenzmitte des Hybrid-Moleküls aus Oligonukleotidsonde und Ziel-DNA befinden. Durch stringente Reaktionsbedingungen wird die Dissoziation unvollständig komplementärer Ziel- und Sonden-DNA erreicht.

Ein weiterer Ansatz bedient sich dem Vergleich mit einer Referenz-DNA, die völlig komplementär zu den immobilisierten Sonden ist. Durch unterschiedliche Fluoreszenzmarkierung kann sie zusammen mit der zu untersuchenden Ziel-DNA kompetitiv hybridisiert und direkt verglichen werden (Hacia & Collins, 1999).

Zur Erhöhung der Bindungsaffinität und -spezifität können anstelle der Oligonukleotidsonden auch Nukleinsäure-Analoga eingesetzt werden. LNA (Locked Nucleic Acid) beispielsweise enthalten eine 2'-O,4'-C-Methylenbindung im Zuckerring. Wie physiologische DNA

unterliegen auch sie den Regeln der Watson-Crick-Basenpaarung und können zur Hybridisierung eingesetzt werden. Weit verbreitet sind auch PNA (Peptide Nucleic Acid)-Sonden, sie verfügen unter Beibehaltung der Basenspezifität über ein ladungsfreies, peptidähnliches, synthetisches Rückgrat (Nielsen *et al.*, 1991). Die fehlende Abstoßung der Phosphatgruppen ermöglicht eine höhere Hybridisierungsstabilität zwischen PNA- und Nukleinsäuresonden (Weiler *et al.*, 1997).

3.3.5 Sortierungsreaktion

Bei diesem Verfahren erfolgt eine Trennung von Diskriminierungs- und Nachweisreaktion. Eine spezifische, enzymatische Reaktion findet räumlich getrennt in Flüssigphase statt. Durch Hybridisierung auf dem Microarray werden die gebildeten Diskriminierungsprodukte identifiziert oder „sortiert“ (Chen *et al.*, 2000).

Bei den dafür verwendeten Oligonukleotidsonden handelt es sich um artifizielle, von der Ziel-DNA unabhängige Sequenzen, die als *zip* oder *tag* bezeichnet werden. Um Kreuzhybridisierungen auszuschließen, werden sie so gewählt, dass sie bei möglichst geringer Sequenzhomologie möglichst ähnliche thermodynamische Eigenschaften aufweisen. Jede dieser Sonden dient dem Nachweis einer Sequenzvariation.

Das System basiert auf dem Einsatz spezifischer Primer in der Diskriminierungsreaktion, die entweder komplementär zur Ziel-DNA verlängert oder ligiert werden. Die Sequenz am 5'-Ende dieser Primer entspricht in komplementärer Form einer Sonde auf dem Array. Erfolgt eine enzymatische Markierung in der Flüssigphase, können nach Hybridisierung gegen die immobilisierten Oligonukleotide Fluoreszenzsignale an der Festphase detektiert werden. Die Signalmuster ermöglichen einen Rückschluss auf die Sequenz der untersuchten Probe. Der Vorteil solcher Arrays liegt darin, dass sie universell für verschiedenste Anwendungen eingesetzt werden können.

3.3.6 Enzymatische Assays

Bei Verfahren, die auf **Ligation** beruhen, werden zwei Oligonukleotide nur dann miteinander ligiert, wenn die Nukleotide an der Ligationstelle komplementär zur Ziel-DNA sind. Die Diskriminierung erfolgt über das Ligationsprodukt (Landegren *et al.*, 1988).

Die Diskriminierung durch **allelische Primer Extension** basiert auf der Eigenschaft der DNA-Polymerase, nur dann einen Nukleinsäurestrang zu verlängern, wenn die letzte Base am 3'-Ende in beiden Strängen komplementär vorliegt. Verlängert werden immobilisierte Oligonukleotide an der Chipoberfläche, deren 3'-Ende die zu untersuchende Base repräsentiert. Weicht dieses Nukleotid von der Sequenz der Ziel-DNA ab, unterbleibt die enzymatische Reaktion. Für jede nachzuweisende Mutation müssen abhängig von der Fragestellung bis zu vier verschiedene Sonden eingesetzt werden, die die möglichen Sequenzvariationen widerspiegeln. Der Nachweis erfolgt durch den Einbau eines oder mehrerer, anteilig fluoreszenz-markierter dNTPs (Erdogan *et al.*, 2001).

Ligations- wie Verlängerungsreaktionen können sowohl in Lösung als auch an der Festphase stattfinden.

3.3.7 Arrayed Primer Extension

Die Arrayed Primer Extension (im Folgenden kurz **APEX** oder **Primer Extension** genannt) nach Kurg *et al.* (2000) dient dem Nachweis von Einzelnukleotid-Mutationen, so genannter *single nucleotide polymorphisms* (SNPs). Die Methode basiert auf einem Oligonukleotid-Microarray, der relevante Sequenzabschnitte einer zu untersuchenden Region abbildet. Die Oligonukleotide (weiterhin auch als **Primer** oder **Sonden** bezeichnet) sind über ihr 5'-Ende an eine beschichtete Glasoberfläche gebunden. Am freien 3'-Ende werden sie von einer DNA-Polymerase komplementär zum gebundenen PCR-Produkt verlängert. Für die Verlängerung der Primer stehen ausschließlich fluoreszenz-markierte Didesoxynukleotide zur Verfügung, so dass die Reaktion nach Einbau eines Nukleotids abbricht. Der Microarray wird anschließend mit Hilfe eines konfokalen Laser-Scanners bei vier verschiedenen Wellenlängen ausgelesen. Durch verschiedene Fluoreszenz-Markierung der vier möglichen ddNTPs, kann aufgrund der detektierten Signale unmittelbar auf die Sequenz der zu untersuchenden Ziel-DNA (zumeist ein PCR-Produkt) rückgeschlossen werden.

APEX unterscheidet sich von der allelischen Primer Extension in der Lage der verwendeten Oligonukleotide. Diese enden statt direkt an der nachzuweisenden Position schon ein Nucleotid zuvor. Eingebaut wird diejenige Base, die es nachzuweisen gilt in Form eines Didesoxynucleotids. Die Methode ermöglicht den spezifischen Nachweis unbekannter Punktmutationen innerhalb einer zumindest teilweise bekannten Nucleinsäuresequenz unter Aufbringung einer möglichst geringen Sondenanzahl. Eine Anwendung findet das Verfahren beispielsweise in der Detektion von SNPs, die im Zusammenhang mit Krebsentwicklung stehen (Shumaker *et al.*, 2001, Tonisson *et al.*, 2000). Vom so genannten Minisequencing (Pastinen *et al.*, 1997) spricht man, wenn die verwendeten Sonden innerhalb einer Ziel-DNA um jeweils nur eine Nucleotidposition verschoben sind. Es wird dabei wie beim herkömmlichen Sequenzieren eine zusammenhängende Sequenz generiert.

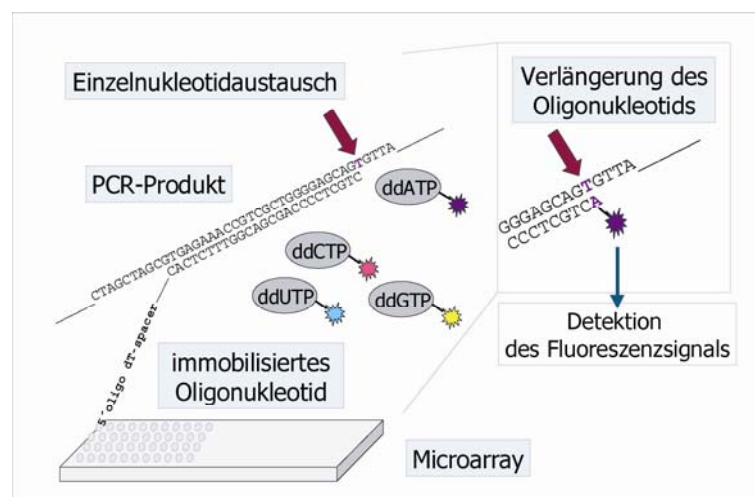


Abbildung 1: Funktionsprinzip der APEX-Reaktion

Komplementäre, enzymatische Verlängerung einer Oligonukleotidsonde durch Fluoreszenz-markierte ddNTPs zum Nachweis eines Einzelnucleotidaustausches innerhalb der Ziel-DNA

4 Problemstellung und Zielsetzung

Die erworbene Immunschwäche AIDS wurde erstmals Anfang der 80er Jahre beschrieben. Übertragen und ausgelöst durch das humane Immundefizienzvirus (HIV) ist sie heute die am weitesten verbreitete tödliche Infektionskrankheit weltweit. Derzeit sind schätzungsweise ca. 39 Mio. Menschen mit dem Virus infiziert (UNAIDS, 12/2004).

Die Einführung von Inhibitoren der HIV-Enzyme reverse Transkriptase (RT) und Protease (PR) und deren Anwendung in Kombinationstherapie haben die Behandlung von AIDS-Patienten entscheidend verbessert. Allerdings erschwert das Vorkommen resistenter Viren zunehmend den erfolgreichen Einsatz dieser Substanzen. Die Entstehung von resistenten Virusvarianten ist die Folge von Mutationen, die aufgrund der hohen Replikations- und Mutationsrate des HIV auftreten und zu einer verringerten Empfindlichkeit der Enzyme gegenüber antiviralen Medikamenten führen.

Genotypische Resistenzbestimmungen in PR und RT werden gegenwärtig durch herkömmliche Sequenzierung nach Sanger nachgewiesen. Sie ist als klinische Routineuntersuchung relativ zeit- und kostenintensiv. Eine Alternative dazu könnte die **APEX-Technologie (Arrayed Primer Extension)** sein. Sie ermöglicht durch die Kombination von Hybridisierung und enzymatischer Reaktion den spezifischen Nachweis von unbekanntem Punktmutationen innerhalb einer teilweise bekannten Nukleinsäuresequenz. Das erfordert die Etablierung der Methode für den Nachweis von Resistenz-assoziierten Mutationen innerhalb der Sequenz von PR und RT, die den derzeitigen Wissensstand reflektieren.

Auf der Grundlage von Datenbankinformationen soll eine möglichst hohe Sequenzabdeckung durch den Oligonukleotid-Microarray erreicht werden, wobei die immense Variabilität des HIV-Genoms eine Herausforderung an das Sondendesign darstellt.

Zur Validierung steht definiertes Patientenmaterial zur Verfügung. Es werden sowohl Subtyp B- als auch nicht-Subtyp B-Proben untersucht, um eine Subtyp-übergreifende Tauglichkeit des Microarrays zu bestätigen. Die APEX-Daten werden zum Vergleich mit den jeweiligen Sequenzierungsdaten (nach Sanger) herangezogen und bezüglich Konkordanz und Sensitivität bewertet.

Ziel ist es, basierend auf der APEX-Technologie ein schnelles und kostengünstiges Verfahren zur genotypischen Resistenzbestimmung zu entwickeln, das in der klinischen Diagnostik eingesetzt werden kann und Hochdurchsatz-Untersuchungen vereinfacht.

5 Material und Methoden

5.1 Material

5.1.1 Geräte und Laborausstattung

DNA-Konzentrationsbestimmung

Photometer Ultrospec 2000 Amersham, Freiburg

Elektrophorese

Elektrophorese-Netzgerät Renner, Darmstadt

Elektrophoresekammer, midi Renner, Darmstadt

Elektrophoresekammer, mini Renner, Darmstadt

Videosystem Geldoc 1000 Biorad, München

Herstellung der Microarrays

MicroGrid II Biorobotics, Cambridge, UK

Hybridisierungsöfen UniEquip, Martinsried

SDDC-2 Chip Writer Engineering Services, Kanada

SMP3- / SMP3B-Pins TeleChem, Sunnyvale, CA, USA

Stereolupe Renner, Darmstadt

Ultraschall-Wasserbad, Sonorex Bandelin, Berlin

PCR

ThermoCycler PTC-200 MJ Research, Waltham MA, USA

Primer Extension

Heizblock Grant, Cambridge, UK

Schüttler NeoLab, Heidelberg

ThermoCycler PTC-200, TwinTower MJ Research, Waltham MA, USA

Scanner

ScanArray 4000XL	Perkin Elmer, Boston, USA
ScanArray 5000	Perkin Elmer, Boston, USA

Software

GenePix 4.0	Axon Instruments, Union City, Kanada
Matlab	The Math Works Inc., Natick, USA
MCHiPS	Kurt Fellenberg, DKFZ Heidelberg
Primer Premier 5	Premier Biosoft Int., Palo Alto, CA, USA
Spotconverter	Kurt Fellenberg, DKFZ Heidelberg

Zentrifugen

Eppendorf Zentrifuge 5810R	Eppendorf, Hamburg
Biofuge pico	Haraeus, Hanau
Vakuum-Konzentrator	H.Saur Laborbedarf, Reutlingen

5.1.2 Chemikalien und Chemikaliensätze

Die allgemeinen Laborchemikalien wurden in der höchstmöglichen Reinheit von den Firmen Sigma-Aldrich (Deisenhofen), Merck (Darmstadt) und Roth (Karlsruhe) bezogen.

Agarose	Biozym, Hess. Oldendorf
Alconox®	Alconox, USA
Aqua ad iniectabilia	Braun, Melsungen
Cyanine3-ddCTP	Perkin Elmer, Boston, USA
Cyanine5-ddUTP	Perkin Elmer, Boston, USA
Desoxyadenosin-Triphosphat	MBI Fermentas, St. Leon-Roth
Desoxycytosin-Triphosphat	MBI Fermentas, St. Leon-Roth
Desoxyguanosin-Triphosphat	MBI Fermentas, St. Leon-Roth
Desoxythymidin-Triphosphat	MBI Fermentas, St. Leon-Roth
Desoxyuridin-Triphosphat	MBI Fermentas, St. Leon-Roth
Fluorescein-12-ddGTP	Perkin Elmer, Boston, USA
High Pure PCR Template Preparation Kit	Roche, Mannheim

Micro Spotting Solution, ArrayIt™(2x)	TeleChem, Sunnyvale, CA, USA
QIAmp Viral RNA Mini Kit	Qiagen, Hilden
QIAquick PCR Purification Kit	Qiagen, Hilden
RnaseOUT™ (RNase-Inhibitor, 40 U/μl)	Gibco, Eggenstein
SAL Spotting Solution (2x)	Asper Biotech, Tartu, Estland
Spotting Solution I (2x)	Quantifoil, Jena
TexasRed®-5-ddATP	Perkin Elmer, Boston, USA

5.1.3 Verbrauchsmaterialien

Eppendorfgefäße, 0,5 ml, 1,5 ml, 2,0 ml	Eppendorf, Hamburg
Falcon, 15 ml, 50 ml	Becton Dickinson, Heidelberg
Frame-Seal™ Kleberahmen	ABgene, Hamburg
Glasobjektträger	Menzel, Braunschweig
Klebefolien für Mikrotiterplatten	Biozym, Hess.Oldendorf
Mikrotiterplatten, 96er, 384er, V-Boden	Genetix, München
Nexterion Slide E (Eypoxysilan)	Schott, Jena
Nexterion Slide A (Aminosilan)	Schott, Jena
PCR-Reaktionsgefäße, 8er-Streifen, 0,2 ml	Biozym, Hess.Oldendorf
Skalpelle	NeoLab, Heidelberg
UV-Küvette Uvette, 100μl, 220-1600 nm	Eppendorf, Hamburg

5.1.4 Enzyme

Shrimp Alkaline Phosphatase (1 U/μl)	Amersham, Freiburg
SuperScriptII Reverse Transkriptase (200 U/μl)	Gibco, Eggenstein
Thermoprime Plus DNA Polymerase (5 U/μl)	ABgene, Hamburg
Thermolabile Uracil N-Glykosylase (1U/μl)	Biozym, Hess. Oldendorf
Thermo Sequenase DNA Polymerase (32 U/μl)	Amersham, Freiburg

5.1.5 Verwendete Puffer und Lösungen

DNA-Auftragspuffer	0,25 % Bromphenolblau 0,25 % Xylencyanol 30 % Glycerol in H ₂ O _{reinst}
Färbebad für Agarosegele	0,5 % (v/v) Ethidiumbromid 10 mg/ml
2 x Natriumcarbonat-Spottinglösung	0,05 M Natriumcarbonat 0,1 M Natriumcitrat 3 M Betain (steril filtriert) pH 9,5
2 x Natriumphosphat-Spottinglösung	6 x SSC 300 mM NaPO ₄ -Puffer 3 M Betain (steril filtriert) pH 9
NaPO₄-Puffer	600 mM Na ₂ HPO ₄ 600 mM NaH ₂ PO ₄ pH 8,5
50 x TAE-Puffer	20 mM Tris/Acetat 50 mM Na ₂ EDTA pH 8,3
1 x TE-Puffer	10 mM Tris/HCl, pH 8,3 1mM EDTA pH 8,3
20 x SSC-Lösung	3 M NaCl 0,3 M Na ₃ Citrat, pH 7

10 x UNG-Reaktionspuffer	500 mM Tris/HCl pH 9
	200 mM (NH ₂) ₄ SO ₄

5.1.6 Oligodesoxynukleotide

Sämtliche Oligonukleotide wurden von den Firmen biomers (Ulm) oder MWG (Ebersberg) synthetisiert und mittels HPLC gereinigt. Die Stocklösungen wurden in einer Konzentration von 100 µM in H₂O_{reinst} bei -20°C gelagert.

Immobilisierte Oligonukleotide

Soweit nicht anders angegeben, tragen alle Oligonukleotide, die zur Immobilisierung auf dem Array verwendet wurden, am 5'-Ende einen Oligo-d(T)-Spacer (bestehend aus 25-30 Thymidinen), sowie eine C6-Aminomodifikation. Diese dient zur kovalenten Bindung an der Silanoberfläche und soll gewährleisten, dass das freie 3'-Ende zur Prozessierung durch die Polymerase zur Verfügung steht. Die Länge der Sequenzen variiert zwischen 23 und 30 Basen, zuzüglich des Oligo-d(T)-Spacer beträgt die Gesamtlänge 50 Basen. In den Namen spiegeln sich die Region (z.B. PR für Protease) und die darin abgefragte Aminosäure bzw. die korrelierende Nukleotidposition wieder. Die Buchstaben f (forward) und r (reverse) kennzeichnen die Orientierung der Oligonukleotide. Die jeweils letzte Position gibt diejenige Base an, die bei Verlängerung komplementär zur Referenz eingebaut wird. Ausnahmen sind Positions- und Extensionskontrollsonden, auf ihren Aufbau wird an entsprechender Stelle im Ergebnisteil näher eingegangen. Alle Sequenzen sind im Anhang aufgeführt, den Abkürzungen der Basen liegt der IUB-Code zugrunde.

PCR-Primer

Zur Herstellung der HIV-spezifischen PCR-Produkte wurden Primer verwendet, deren Sequenzen von der Abteilung Virologie des Hygiene-Instituts an der Universität Heidelberg bereitgestellt wurden. Amplifiziert wurden Sequenzen, die entweder den Bereich der HIV-Protease (#120/121), der Reversen Transkriptase (#123/124) oder beider Enzyme abdecken (#118/119). In Tabelle 1 (oben) sind die Sequenzen, die Lage im Plasmid pNL43 (5.1.7), sowie die erwartete Amplifikatlänge angegeben.

Die Primer für die Amplifikation der MDR1- und CCR5-Regionen wurden mit Hilfe des Programms *Primer Premier 5* abgeleitet (Tabelle 1, Mitte und unten).

Tabelle 1: Verwendete PCR-Primer

Name	Sequenz 5'-3'	Lage in pNL43	Amplifikatlänge
# 118 / Prot-1	ATGACAGCATGTCAGGGAGT	1828-1847	1493 bp
# 119 / Pol-8	GTATGTCATTGACAGTCCAG	3321-3302	
# 120 / Prot-3	GGGAAGATCTGGCCTTCCTACAA	2092-2116	559 bp
# 121 / Prot-4	GGGCCATCCATTCGTGGC	2605-2588	780 bp
# 123 / Pol-2	CACCTGTCAACATAATTGGA	2491-2510	
# 124 / Pol-4	ACTGTCCATTTATCAGGATG	3271-3252	

Name	Sequenz 5'-3'	Lage im <i>MDR1</i> -Gen
mdr1_f1	GTTTCCGGTTTGGAGCCT	3288-3305
mdr1_f2	GCCGTGGGGCAAGTCAGT	3380-3397
mdr1_f3	CGTGGGGCAAGTCAGTTC	3382-3399
mdr1_r1	CCGGTCCGGTGGGATAGT	3549-3566
mdr1_r2	GCCTTCCGTGCTGTAGCT	3479-3496

Name	Sequenz 5'-3'	Lage im <i>CCR5</i> -Gen
CCR5_f1	GTGTTTGCTTTAAAAGCCAGG	443-463
CCR5_f2	TTAAAAGCCAGGACGGTCAC	451-471
CCR5_r1	CCAGCCCCAAGATGACTATCT	635-615
CCR5_r2	GGATTCCCGAGTAGCAGATGA	695-675

5.1.7 Probenmaterial

Das Probenmaterial zur Etablierung der Microarrays wurde von der Abteilung Virologie des Hygiene-Instituts der Universität Heidelberg zur Verfügung gestellt.

Als Referenz für die Optimierungsphase des Microarrays liegt die Sequenz des Plasmids pNL43 zugrunde. Die Plasmid-DNA wurde ausschließlich zu Amplifikations-Zwecken in PCR-Reaktionen (10 pg/Ansatz) eingesetzt.

Außerdem standen PCR-Produkte aus dem routinediagnostischen Labor, die Sequenzen der HIV-Protease und Reversen Transkriptase abdeckten, zur Verfügung. Die Proben waren anonymisiert, die zugehörigen Sequenzdaten lagen für die Validierung vor. Alle PCR-Produkte wurden zur Analyse unter angegebenen Bedingungen (5.2.4) reamplifiziert.

5.2 Methoden

5.2.1 Isolierung genomischer DNA aus Blut

Die Isolierung genomischer DNA aus eigenem Blut wurde mit Hilfe eines Kits der Firma Roche (High Pure PCR Template Preparation Kit) nach Angaben des Herstellers durchgeführt. Hierfür wurden jeweils 200µl Blut mit dem gleichen Volumen an mitgeliefertem Bindungspuffer versetzt und für 10 min bei 70°C ein Proteinase K-Verdau durchgeführt. Die Ansätze wurden dann zusammen mit 100µl Isopropanol jeweils auf eine der Säulen gegeben, bei 8000 x g abzentrifugiert und der Durchfluss verworfen. Nach Zugabe von 500µl *Inhibitor Removal Buffer* wurde wiederum unter gleichen Bedingungen zentrifugiert und der Durchfluss verworfen. Es folgten zwei Waschschriffe mit je 500µl Waschpuffer und ein Trockenzentrifugieren der Säule bei 13 000 x g für 10 Sekunden. Die DNA wurde in je 200µl vorgewärmten Elutionspuffer (70°C) von der Säule eluiert, in einem sterilen Eppendorfgefäß aufgefangen und anschließend bei -20°C gelagert. Die Konzentration der isolierten DNA wurde im UV-Photometer bestimmt. Die Template-Konzentration für die PCR lag bei 1 µg / Ansatz.

5.2.2 Isolierung viraler RNA

Um eine genotypische Resistenzbestimmung durchzuführen, muss zunächst die virale RNA aus Plasmaproben HIV-infizierter Patienten isoliert werden. Hierfür wurde ein Kit der Firma Qiagen verwendet (5.1.2). Nach Angaben des Herstellers wurden je 500 µl Plasma zusammen mit 2 ml AVL Lyse-Puffer für 10 min bei Raumtemperatur inkubiert. Danach wurden je 2 ml 100% EtOH zugegeben und die Mischung auf eine der Säulen pipettiert. Das Anbinden der RNA an das Säulenmaterial folgte durch Zentrifugation bei 8000 rpm bis das vollständige Volumen geladen war. Danach wurde zweimal mit 500 µl der mitgelieferten Puffer AW1 und AW2 gewaschen und im Anschluss für 3 min bei 13 000 rpm trocken zentrifugiert. Die RNA wurde mit 50 µl AVE-Puffer eluiert. Dieser Vorgang, sowie die folgende cDNA-Synthese wurden in Laboren durchgeführt, die der Sicherheitsstufe S 2 unterliegen.

5.2.3 Reverse Transkription

Durch reverse Transkription erfolgt eine Umschreibung der isolierten viralen RNA (5.2.2) in cDNA. Die cDNA-Synthese wird mit dem Primer #171 JD5574(-) durchgeführt (5'-CTAGTGGGATGTGTACTTCTGAACTTATTTT-3').

Um eine Hybridisierung des Primers an die komplementäre RNA zu erreichen, wird der Ansatz zunächst für 10 min bei 65°C inkubiert. Die Zugabe der restlichen Komponenten erfolgt auf Eis, in Tabelle 1 ist der Reaktionsansatz angegeben.

Nach 1,5 h reverser Transkription bei 42°C wird die Reaktion durch Inaktivierung des Enzyms (10 min, 72°C) abgebrochen. Die cDNA kann nun als Matrize zur Amplifikation mittels PCR (5.2.4) verwendet werden.

Tabelle 2: Reaktionsansatz für cDNA-Synthese

Komponenten	eingesetzte Menge	Endkonzentration
Primer #171 (10 µM)	1 µl	0,5 µM
H ₂ O _{dest}	1 µl	
RNA (Konz. je nach Versuch)	10 µl	je nach Versuch
12 µl Ansatz für 10 min bei 65°C inkubieren		
Reaktionspuffer (5 x)	4 µl	1 x
DTT (0,1 M)	2 µl	10 mM
dNTPs (je 2,5 mM)	1 µl	125 nM
RnaseOUT™ (40U/µl)	0,5 µl	20 U/µl
SuperScript II (200 U/µl)	0,5 µl	100 U/Ansatz
Gesamtvolumen	20 µl	

5.2.4 Polymerase-Kettenreaktion

Die Polymerase-Kettenreaktion (PCR) ist ein sehr sensibles Verfahren zur *in vitro*-Amplifikation kurzer DNA-Abschnitte. Sie eignet sich zum spezifischen Nachweis kleinster DNA-Mengen, in der Regel unter der Voraussetzung, dass die Nukleotid-Sequenzen an den Rändern der Zielregion bekannt sind. Das Prinzip der PCR ist die enzymatische Vermehrung eines DNA-Abschnittes zwischen zwei Oligonukleotid-Primern, die gegenläufig an die komplementären DNA-Stränge gebunden sind.

Der Ablauf einer PCR kann in drei Schritte gegliedert werden. Zuerst wird die doppelsträngige DNA bei 94-95°C denaturiert. Unter geeigneten Hybridisierungsbedingungen lagern sich die Primer an ihre komplementären Zielregionen an (Annealingphase). Die hitzestabile Taq-Polymerase kann nun, ausgehend vom 3'-OH-Ende der Primer, komplementäre Sequenzen synthetisieren. Man bezeichnet diesen Abschnitt der PCR als Extensionsphase. Der gesamte Zyklus wird 30-50 mal wiederholt.

Aufgrund der hohen Sensitivität dieser Methode muss großer Wert auf die Vermeidung von Kontaminationen gelegt werden. Für die PCR-Ansätze und die Verdünnungen der Oligonukleotide wurde ausschließlich steriles H₂O_{reinst} verwendet. Um falsch-positive Ergebnisse auszuschließen, wurde zu jedem Ansatz eine Negativkontrolle durchgeführt, die anstelle der Matrizen-DNA H₂O_{reinst} enthielt.

Tabelle 3 beschreibt den Ansatz für eine PCR-Reaktion in 50µl Volumen. 15 % der dTTPs im Nukleotid-Gemisch wurden durch dUTP ersetzt. Der Einbau von dUTP während der Amplifikation ist Voraussetzung für den enzymatischen Verdau der DNA mit Uracil N-Glykosylase im Anschluß an die PCR (5.2.8).

Tabelle 3: PCR-Ansatz für 50 µl Volumen

Komponenten	eingesetzte Menge	Endkonzentration
H ₂ O _{reinst}	34,25 µl	
10 x Puffer	5,0 µl	1 x
dNTPs (5 mM, 15 % dUTP)	2,5 µl	250 µM
forward-Primer (10 µM)	1,5 µl	300 nM
reverse-Primer (10 µM)	1,5 µl	300 nM
Taq-Polymerase (5 U/ µl)	0,25 µl	1,25 U / Ansatz
Template-DNA oder H ₂ O _{reinst} (Negativkontrolle)	5 µl	je nach Versuch -
Gesamtvolumen	50 µl	

Falls nicht anders angegeben, wurde für diese Reaktionen folgendes Programm verwendet (Tabelle 4).

Tabelle 4: Verwendetes PCR-Profil

Schritt	Dauer	Temperatur
1. Aktivierung der Polymerase	2 min	94 °C
2. Denaturierung	30 s	94 °C
3. Annealing	30 s	56 °C
4. Extension	2:30 min	72 °C
5. abschließende Extension	10 min	72 °C

Die Schritte 2.-4. werden 30-mal wiederholt.

Die entstandenen PCR-Produkte konnten über Gel-Elektrophorese ihrer Größe nach aufgetrennt und durch Anfärbung mit Ethidiumbromid und Betrachtung unter UV-Licht detektiert werden (5.1.1).

5.2.5 Agarosegel-Elektrophorese

Diese Technik dient der Auftrennung von DNA-Stücken in einem Agarose-Gel nach ihrer Größe mit Hilfe einer angelegten elektrischen Spannung.

Die Agarose wurde als 2% (w/v)-Suspension in 1 x TAE-Puffer (5.1.5) angesetzt und durch Aufkochen gelöst. Die noch flüssige Agarose wurde in die Gelschlitten gegossen, die Bildung der Probenaschen erfolgte durch Einstecken eines Gelkamms. Nach Erstarren der Agarose konnte der Gelkamm entfernt und das Gel zur Auftrennung von DNA-Fragmenten genutzt werden. Die Elektrophoresekammern (5.1.1) waren mit 1 x TAE-Puffer befüllt.

Jeweils 2 µl der PCR-Ansätze wurden mit je 0,2 Volumen DNA-Auftragspuffer (5.1.5) in die Probenaschen pipettiert. Als Kontrolle wurde ein Längenstandard aufgetragen, anhand dessen Banden im Vergleich die Größe und die ungefähre Menge des PCR-Produkts abgelesen werden konnte. Von diesem Marker wurden je 6 µl in die dafür vorgesehenen Geltaschen gegeben. Die angelegte Spannung betrug 10 V / cm Gel.

Anschließend wurde das Gel in einem Ethidiumbromidbad (5.1.5) für ca. 20 min angefärbt.

Ethidiumbromid dient dem Sichtbarmachen der DNA-Banden im Gel, es hat die Fähigkeit, in DNA zu interkalieren und kann unter UV-Licht nachgewiesen werden. Die Dokumentation der Gele erfolgte bei einer Wellenlänge von 254 nm (5.1.1). Alle Gelbilder in dieser Arbeit sind invers dargestellt.

Elektrophoresemarker:

Als Längenstandard wurde der GeneRuler™ 100bp DNA Ladder Plus (MBI Fermentas, St. Leon-Roth) verwendet. Die 500 bp-Bande enthält in der aufgetragenen Menge von 6 µl 66,5 ng DNA. Dieser Wert wurde zur Konzentrationsabschätzung der DNA-Banden herangezogen. Bei Darstellungen von Gelbildern wird nur die jeweils erwartete Bandengröße angegeben.

5.2.6 Isolierung von DNA-Amplifikaten aus PCR-Ansätzen

Zur weiteren Verwendung der PCR-Produkte muss diese von den anderen Komponenten des PCR-Ansatzes getrennt und gereinigt werden. Hierfür wurde ein Kit der Firma Qiagen verwendet (5.1.2). Alle Zentrifugationsschritte (5.1.1) wurden bei 13 000 rpm (17 900 x g) durchgeführt. Nach Angaben des Herstellers wurden 1 Volumen an zu reinigendem PCR-Ansatz mit 5 Volumen des Puffers *PB* gemischt und auf eine der mitgelieferten Säulen pipettiert, maximal allerdings 800 µl. Die DNA wurde an das Säulenmaterial gebunden, der Durchfluss durch Zentrifugation abgetrennt und verworfen. Die Säule wurde danach mit 0,75 µl Waschpuffer *PE* beladen und zentrifugiert. Nach Verwerfen des Durchflusses wurden Reste des Waschpuffers durch erneutes Zentrifugieren entfernt. Zum Eluieren der DNA wurden 30 µl H₂O_{reinst} direkt auf die Oberfläche des Säulenmaterials gegeben, für 5 Minuten bei Raumtemperatur inkubiert und ein letztes Mal zentrifugiert. Im Anschluss wurde ein Aliquot der eluierten DNA durch Auftragen auf ein Agarosegel kontrolliert.

5.2.7 Photometrische Konzentrationsbestimmung von Nukleinsäurelösungen**Konzentrationsbestimmung von Oligonukleotidlösungen**

Durch photometrische Messung wurden die Konzentrationen der Oligonukleotide bestimmt und Stammlösungen von 100 µM eingestellt. Die Messungen wurden bei einer Wellenlänge von 260 nm, was dem Absorptionsmaximum der Basen der Nukleinsäure entspricht, durchgeführt. Es besteht ein linearer Zusammenhang zwischen der Absorption bei 260 nm und der Konzentration der Nukleinsäurelösung. Durch Kenntnis des Molekulargewichts und

ausgehend von durchschnittlich 33 µg einzelsträngiger DNA / ml bei 1 OD₂₆₀ konnten die Lösungen der Oligonukleotide auf einheitliche Konzentrationen eingestellt werden.

Konzentrationsbestimmung von PCR-Produkten

Die Messungen im Photometer wurden für PCR-Produkte auf die gleiche Weise wie für Oligonukleotidlösungen durchgeführt. Für die Berechnung der Konzentration von doppelsträngiger DNA lag bei 1 OD₂₆₀ ein Wert von 50 µg / ml zugrunde.

5.2.8 Enzymatischer Verdau der PCR-Produkte

Um sterische Hinderung bei der Hybridisierung der PCR-Produkte an die immobilisierten Oligonukleotide auf dem Microarray zu verringern, wurde die amplifizierte DNA zunächst enzymatisch fragmentiert. Das dazu verwendete Enzym Uracil N-Glykosylase (UNG, 5.1.4) hydrolysiert die N-glykosidische Bindung zwischen der Desoxyribose und Uracil. Außerdem wurden durch zusätzliche Inkubation mit einer alkalischen Phosphatase (sAP, 5.1.4) freie dNTPs dehydrolysiert. Dadurch soll verhindert werden, dass möglicherweise noch vorhandene dNTPs während der Verlängerungsreaktion auf dem Microarray mit den dabei eingesetzten ddNTPs konkurrieren.

Die Konzentration der eingesetzten DNA wurde entweder photometrisch bestimmt (5.2.7) oder durch Vergleich mit dem verwendeten Elektrophoresemarker (5.2.5) abgeschätzt. Das Volumen für einen Reaktionsansatz betrug 25 µl (Tabelle 5). Die Inkubation erfolgte für 1h bei 37°C, die anschließende Inaktivierung der Enzyme für 10min bei 95°C.

Tabelle 5: Reaktionsansatz für einen UNG- /sAP-Verdau

Komponenten	eingesetzte Menge	Endkonzentration
H ₂ O _{dest}	0,5 µl	
10 x UNG-Reaktionspuffer	2,5 µl	1 x
UNG (0,1 U/µl)	1,0 µl	0,1 U/Ansatz
sAP (1 U/µl)	1,0 µl	1 U/Ansatz
PCR-Produkt (ca.40-50 ng/µl)	20 µl	je nach Versuch
Gesamtvolumen	25 µl	

Anschließend wurde durch Kontrolle mittels Agarosegel-Elektrophorese sichergestellt, dass die DNA in fragmentiertem Zustand vorliegt.

5.2.9 Herstellung der Microarrays

Die verwendeten Primer wurden zunächst in Natriumcarbonat-Spottinglösung (5.1.5) auf eine Konzentration von 50 μM verdünnt und entsprechend ihrer späteren Anordnung auf dem Chip in eine Mikrotiterplatte pipettiert. Mit Hilfe von Pipettierrobotern (SDDC-2 Chip Writer oder MicroGrid II; 5.1.1), die nach dem Kontaktverfahren arbeiten, wurden die Oligonukleotide auf Epoxysilan-beschichtete Glaträger aufgebracht. Durch Verwendung von SMP 3-Nadeln (5.1.1) wurde ein Volumen von ca. 1,0 nl pro Spot und Spotdurchmesser von ca. 100 μm erreicht. Der Abstand zwischen den Spots betrug je 200 μm .

Zusätzlich zu den HIV-spezifischen Oligonukleotiden sind in jedem Array weitere Sonden zur Kontrolle von Anordnung und Kopplung vorhanden. Diese so genannten Positionskontrollen sind fluoreszenz-markierte Oligonukleotide, sie markieren die Eckpunkte des Arrays. Die Kontrollsonden wurden in einer Konzentration von 12,5 μM ebenfalls in Natriumcarbonatpuffer aufgebracht.

Zur Immobilisierung der Oligonukleotide wurden die Microarrays im Anschluß an den Spottingvorgang zunächst für 1h bei 60°C im Hybridisierungsöfen und dann über Nacht bei Raumtemperatur inkubiert. Die Fläche, die der Array auf dem Glaträger einnimmt, muss vor der Weiterbehandlung mit einem Diamantstift gekennzeichnet werden. Unmittelbar nach dem Aufbringen der Oligonukleotide sind die Flüssigkeitstropfen auf der Oberfläche noch deutlich zu erkennen.

5.2.10 Blocken der Oberfläche

Um eine unspezifische Bindung der zu untersuchenden DNA an den Glaträger zu verhindern, muss zunächst die Epoxysilanoberfläche deaktiviert werden. Nicht gebundene Oligonukleotide und Spottingpuffer werden durch Waschen mit heißem $\text{H}_2\text{O}_{\text{reinst}}$ entfernt. Tabelle 6 zeigt die Abfolge der Waschschritte, die jeweils direkt vor der Primer Extension-Reaktion durchgeführt wurden.

Für alle Waschvorgänge, sowohl vor als auch nach der APEX-Reaktion (Tabelle 8), wurden die Microarrays zunächst in einem Plastikhalter für maximal 20 Glaträger befestigt und unter Schütteln gewaschen. Das Volumen der jeweiligen Waschlösungen betrug ca. 250 ml.

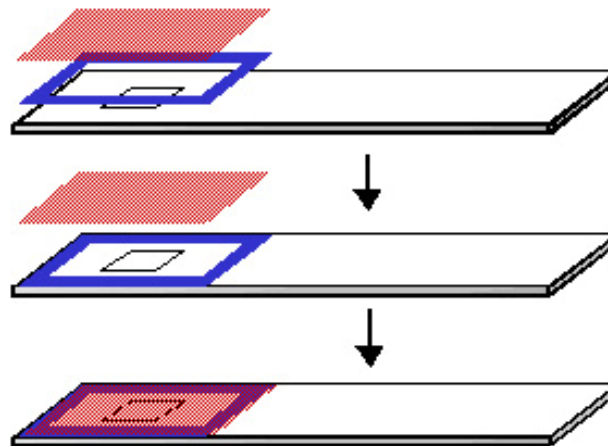
Tabelle 6: Vorbehandlung der Microarrays

Waschlösung	Temperatur	Dauer
H ₂ O _{reinst}	ca. 95°C	1 x 1 min
100 mM NaOH	Raumtemperatur	1 x 10 min
H ₂ O _{reinst}	ca. 95°C	2 x 1 min

Nach dem abschließenden Waschschrift wurden die Microarrays unmittelbar mit Druckluft getrocknet.

5.2.11 Vorbereitung der Microarrays

Vor der APEX-Reaktion muss eine Reaktionskammer (5.1.3) auf den Chip aufgeklebt werden, die ein Volumen von 20-25 µl Reaktionsansatz fasst. Der PCR-Ansatz wird in diese Reaktionskammer pipettiert, die danach durch eine Abdeckfolie verschlossen wird. Abbildung 2 stellt diesen Vorgang dar. Es muss gewährleistet sein, dass der Kleberahmen fest auf dem Glsträger anhaftet, um ein Austrocknen der Flüssigkeit während der Reaktion zu verhindern.

**Abbildung 2: Aufbringen der Reaktionskammer auf den Microarray**

Rot: Abdeckfolie zum Verschließen der Kammer
Blau: Kleberahmen, der die Kammer bildet und den Array umschließt.

5.2.12 Arrayed Primer Extension (APEX)

Tabelle 7 zeigt die eingesetzten Mengen und Konzentrationen für eine Primer Extension-Reaktion. Als Template wurde DNA aus dem UNG- / sAP-Verdau, ohne vorherige Reinigung aus dem Reaktionsansatz, eingesetzt. Als DNA-Polymerase wurde eine Thermo Sequenase (5.1.4) gewählt, die sich besonders gut für den Einbau der farbmarkierten ddNTPs eignet.

Folgende Farbstoff-Nukleotid-Kombinationen wurden verwendet (nach Kurg *et al.*, 2000):

- ddATP-TexasRed
- ddCTP-Cyanine3
- ddGTP-Fluorescein
- ddUTP- Cyanine5

Vor Zugabe in den Reaktionsansatz wurde die Ziel-DNA zunächst für 10 min bei 95°C im Heizblock denaturiert. Alle folgenden Pipettierschritte erfolgten auf Eis. Jeweils 20 µl Ansatz wurden in eine Reaktionskammer auf dem Microarray gegeben und mit einer Abdeckfolie luftblasenfrei verschlossen (5.2.11).

Tabelle 7: APEX-Reaktionsansatz

Komponenten	eingesetzte Menge	Endkonzentration
10 x Thermo Sequenase-Reaktionspuffer	2,0 µl	1 x
ddATP-TexasRed (50 mM)	0,5 µl	1,25 µM
ddCTP-Cy3 (50 mM)	0,5 µl	1,25 µM
ddGTP-Fluorescein (50 mM)	0,5 µl	1,25 µM
ddUTP-Cy5 (50 mM)	0,5 µl	1,25 µM
Thermo Sequenase (4 U/µl)	1,0 µl	4 U / Ansatz
verdautes PCR-Produkt	10 µl	je nach Versuch
H ₂ O _{reinst}	5,0 µl	-

Gesamtvolumen 20 µl

Die APEX-Reaktion wurde ähnlich einer PCR in einem Thermocycler mit entsprechendem Aufsatz für Objektträger durchgeführt (5.1.1). Standardmäßig wurde folgender Ablauf verwendet: Zunächst wurde die DNA auf den Glasträgern bei 95°C für 30 sec denaturiert. Die Hybridisierung an die komplementären Oligonukleotidsequenzen am Chip erfolgte für 10 min

bei 40°C. Für die spezifische Verlängerung durch die Polymerase wurde die Temperatur für 15 min auf 65°C eingestellt. Die Heiz- bzw. Kühlrate des Gerätes lag bei 1°C/Sekunde. Zeitspannen, die dadurch bedingt zum Heizen oder Kühlen notwendig sind, kommen zu den angegebenen Zeiten hinzu. Für einige Experimente gilt ein abweichendes Temperatur-Profil, dies ist an den entsprechenden Stellen im Ergebnisteil vermerkt.

Anschließend wurden die Microarrays mit H₂O_{dest} und Alconox® (enthält u.a. SDS) gewaschen, noch freie ddNTPs und Reste des Reaktionsansatzes, sowie die Ziel-DNA wurden dadurch entfernt. Das Waschprotokoll ist in Tabelle 8 angegeben, danach wurden die Chips wiederum mit Druckluft getrocknet.

Tabelle 8: Nachbehandlung der Microarrays

Waschlösung	Temperatur	Dauer
H ₂ O _{reinst}	ca. 95°C	1 x 1 min
0,3 % Alconox®	Raumtemperatur	1 x 10 min
H ₂ O _{reinst}	ca. 95°C	2 x 1 min
0,2 x SSC, pH 8	Raumtemperatur	1 x 1 min

5.2.13 Detektion der elongierten Oligonukleotide

Zum Nachweis der eingebauten ddNTPs können die Microarrays unmittelbar mit Hilfe eines konfokalen Vierfarb-Laserscanners (5.1.1) ausgelesen werden. Vorhandene Fluoreszenzsignale werden so detektiert und können innerhalb des Arrays lokalisiert werden. Die Wellenlängen für Anregung und Emission sind auf die vier verwendeten Fluoreszenzfarbstoffe abgestimmt, wie in Tabelle 9 gezeigt.

Tabelle 9: Detektion der Fluoreszenzfarbstoffe

Fluoreszenzfarbstoff	Anregungswellenlänge	Emissionswellenlänge
Cyanine3	532 nm	570 nm
Cyanine5	635 nm	670 nm
Fluorescein	488 nm	522 nm
TexasRed	594 nm	614 nm

5.2.14 Quantifizierung der Signalintensitäten

Für die quantitative Auswertung der Signalintensitäten stand eine entsprechende Software (GenePix 4.0, siehe 5.1.1) zur Verfügung.

Die detektierten Bilder wurden in Falschfarben wiedergegeben, wobei jedem Farbstoff eine Farbe zugeordnet wurde (Cy3/grün, Cy5/rot, Fluorescein/blau, TexasRed/gelb). Über diese Falschfarbenbilder wurde ein Raster gelegt und die einzelnen Punkte angepasst. Die Software errechnet die Signalintensitäten und ordnet jeder Position einen spezifischen Namen zu. Das Ergebnis wird in einer umfangreichen Datei (*genepix result* oder *gpr file*) ausgegeben. Sie beinhaltet unter anderem Datensätze wie Median, Mittelwert und Standardabweichung für die Intensitäten und Werte für den jeweiligen lokalen Hintergrund. Diese Rohdaten wurden für die anschließende Auswertung herangezogen (5.2.15).

5.2.15 Auswertung der Daten

Weil für die statistische Auswertung der durchgeführten Vierfarb-Experimente keine geeignete Software zur Verfügung stand, wurde zusammen mit Aslihan Ay im Rahmen einer Diplomarbeit ein Programm etabliert, das die statistisch signifikante Verarbeitung der generierten Daten ermöglicht. Dieses Analyseprogramm umfasst die Hintergrundkorrektur der Signalintensitäten, das Ausfiltern nicht signifikanter Signale und mehrere Normalisierungsschritte. Die Ergebnisse werden in *result file* und *sequence file* ausgegeben. Die inhaltlichen Anforderungen, die diese Software erfüllen sollte, wurden in Zusammenarbeit erarbeitet und werden deshalb im Ergebnisteil dieser Arbeit aufgeführt.

Die Programmierschritte, die dieser Auswertung im Detail zugrunde liegen, sind in der Diplomarbeit von Aslihan Ay (Fachhochschule Bingen, Januar 2006) ausführlich angegeben.

6 Ergebnisse

Für die Entwicklung eines Genotypisierungs-Assays für HIV-Sequenzen, der auf einer Hybridisierung gegen Oligonukleotide beruht, stellt die immense Heterogenität des viralen Genoms eine große Herausforderung dar. Zwar sind die Zahl der relevanten Positionen und die Größe der kodierenden Regionen für Protease und Reverse Transkriptase überschaubar, doch macht die fast uneingeschränkte Variabilität eine unfassbare Anzahl an Sequenzkombinationen möglich, die als Sonden abgebildet werden müssen. Mutationen innerhalb der Primerbindungsstellen sind zwar für den eigentlichen Nachweis nicht von Interesse, müssen aber, um eine Hybridisierung zu gewährleisten, im Design der Primer berücksichtigt werden.

Die Etablierung setzt also zunächst Datenbankrecherchen und die Überlegungen voraus, welche Positionen nachgewiesen werden sollen (6.1) und welche Sequenzen die Grundlage für das Primerdesign bilden sollen (6.2). Im Folgenden wurden die Bedingungen für die praktische Durchführung der APEX-Reaktion zum Nachweis der HIV-Sequenzen optimiert (6.3 + 6.4). Um die Abbildung der ausgewählten Sequenzen auf dem Microarray zu realisieren, mussten die verwendeten Oligonukleotide modifiziert werden (6.5). Für Protease und Reverse Transkriptase wurden jeweils ein Microarray zur Detektion der resistenz-assoziierten Positionen entwickelt, die Arrays wurden durch Untersuchung definierter Patientenproben evaluiert (6.6- 6.9). Zur Auswertung der generierten Microarray-Daten mussten geeignete Strategien entwickelt werden (6.10).

6.1 Datenbankrecherche und Auswahl der Sequenzen

6.1.1 Auswahl der nachzuweisenden Positionen

Im Gegensatz zu herkömmlichen Sequenzierungsmethoden, beschränkt sich ein Punktmutations-Nachweis auf die Detektion einzelner Positionen. Voraussetzung für die Etablierung eines Genotypisierungs-Assays ist die Auswahl der entsprechenden Positionen im HIV-Genom. Der Microarray soll alle diejenigen Mutationen umfassen, die von der Stanford HIV-Datenbank (<http://hivdb.stanford.edu/cgi-bin/PIResiNote.cgi>) als Resistenz-

assoziiert eingestuft wurden, unabhängig vom Grad der beschriebenen Resistenz. Ausnahmen sind Aminosäure 60 und 93 der Protease, die nicht aufgrund der Datenbankinformation, sondern aus der gegenwärtigen Literatur ausgewählt wurden. Für die Protease ergab sich somit eine Zahl von 26, für die Reverse Transkriptase von 33 Aminosäuren, deren Austausch infolge von Mutationen mit einer veränderten Empfindlichkeit gegenüber Inhibitoren korreliert sein kann. In diesem Zusammenhang spricht man auch von so genannten *Schlüsselmutationen* oder *-positionen*.

Tabelle 10 zeigt die ausgewählten Aminosäurepositionen, die im Zusammenhang mit Inhibitorresistenzen relevant sind.

Tabelle 10: Resistenz-assoziierte Aminosäurepositionen innerhalb der HIV-PR und -RT

Protease

10	20	23	24	30	32	33	36	46	47	48	50	53	54	60	62	63	71	73	76	77	82	84	8	90	93
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	---	----	----

Reverse Transkriptase

41	44	62	65	67	69	70	74	75	77	98	100	101	103	106	108	115	116	118	151
179	181	184	188	190	210	215	219	225	227	230	236	238							

6.1.2 Sequenzverarbeitung

Laden der Sequenzen

Die Grundlage für den HIV-Chip bilden Sequenzen, die über die Los Alamos-Sequenzdatenbank zur Verfügung standen (<http://hiv-web.lanl.gov/>). Für die Protease lag bereits ein Alignment aller zum damaligen Zeitpunkt vorhandener Sequenzen vor (Anfang 2003). Für die Reverse Transkriptase wurde über das *search interface* dieser Datenbank gezielt nach allen verfügbaren Sequenzen gesucht, die die entsprechende Region abdeckten (p51, RT) und eine Mindestlänge von 100 Basen aufwiesen. Es wurden keine Einschränkungen den Subtyp, die geographische Region oder Patientendaten betreffend vorgenommen.

(http://hiv-web.lanl.gov/components/hiv-db/combined_search_s_tree/search.html)

Die Sequenzen wurden in Textdateien geladen, es handelte sich dabei um insgesamt 15 380 Datensätze, für die im Folgenden ein Alignment vorgenommen werden musste.

Sequenzalignment

Mit Unterstützung der Abteilung Bioinformatik-HUSAR (DKFZ-Heidelberg) wurden die Rohdaten weiter bearbeitet. Das Alignment wurde mit Hilfe des Programms Clustalw durchgeführt, das allerdings limitiert ist auf das Vergleichen von höchstens 1000 Datensätzen, von denen jeder wiederum nicht mehr als 10 000 Zeichen umfassen darf. Aus diesem Grund wurden die knapp 16 000 Sequenzen zunächst nach der Reihenfolge, in der sie geladen wurden, in 16 Dateien aufgeteilt.

Anschließend wurde die Häufigkeit der einzelnen Nukleotide an jeder Position des Alignments mit Hilfe des Programms *prophecy* bestimmt (<http://www.emboss.org>). Die Ausgabedateien des verwendeten Programms (*prophecy files*) können mittels herkömmlicher Tabellenkalkulations-Programme (z.B. microsoft excel) geöffnet und bearbeitet werden. Jede dieser Dateien besteht aus insgesamt 26 Spalten und höchstens 10 000 Zeilen, wobei jede der Spalten einen Buchstaben des Alphabets angibt, also Spalte 1=A, 2=B, usw. Die Zeilen spiegeln die fortlaufende Sequenz, ähnlich einem Sequenziergel, wieder. Abbildung 3.1 soll dies verdeutlichen, jeder Base in der Sequenz wird an entsprechender Position der Wert 1 zugeordnet, allen anderen eine Null. Von oben nach unten in 5'-3'-Richtung gelesen, ergibt sich die Sequenz *ACTGG*. Analoges gilt bei Vergleichen mehrerer Sequenzen, die Zeilensumme ergibt die Gesamtanzahl der betrachteten Sequenzen (Abbildung 3.2). Der Konsensus bezeichnet diejenige Sequenz, die am häufigsten vertreten ist, bei einer repräsentativen Anzahl entspricht er dem Wildtyp.

1.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Summe	Konsensus
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	A
0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	C
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	T
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	G
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	G

2.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Summe	Konsensus
934	0	0	0	0	0	62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	996	A
8	0	14	0	0	0	947	0	0	0	1	0	0	8	0	0	0	1	0	16	0	0	0	0	1	0	996	G
938	0	1	0	0	0	55	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	996	A
0	0	34	0	0	0	58	0	0	0	0	0	3	0	0	0	0	0	0	900	0	0	0	0	1	0	996	T
660	0	14	0	0	0	310	0	0	0	0	0	5	0	0	0	7	0	0	0	0	0	0	0	0	0	996	A

Abbildung 3: Auszug aus einer Ausgabedatei des Alignments für RT-Sequenzen

1.: beim Betrachten nur einer Sequenz (zur Demonstration)

2.: Alignment von 996 Sequenzen

Von den 26 Spalten, die jeden Buchstaben des Alphabets abdecken, wurden nur 11 Spalten, nämlich A, C, G, K=G+T, M=A+C, N=A+C+G+T, R=A+G, S=G+C, T, W=A+T und Y=C+T, zur weiteren Bearbeitung herangezogen. Da nicht alle Sequenzen die gleiche Länge und den gleichen Anfangspunkt haben, treten Überlappungen auf. Zunächst musste also in allen 16 Dateien der Anfang der RT-Sequenz definiert werden. Danach wurden die Datensätze spaltenweise addiert und der jeweilige Basenanteil prozentual am Gesamtvorkommen in der Datenbank angegeben. Abbildung 4 zeigt in Leserichtung von oben nach unten die ersten drei Triplets der RT, die recht hoch konserviert sind (Konsensus CCC ATT AGT). In dieser Weise konnte für jede Position der RT die prozentuale Verteilung der möglichen Basen ersehen werden. Für das Design der Oligonukleotidsequenzen wurden nur diejenigen Sequenzen berücksichtigt, die einen in Abhängigkeit von der Variabilität festgelegten Schwellenwert überschreiten (PR: 5%, RT 10%, siehe 6.2.1).

A	C	G	K=G+T	M=A+C	N	R=A+G	S=G+C	T	W=A+T	Y=C+T
0,03%	95,13%	0,01%	0,00%	0,00%	4,48%	0,00%	0,01%	0,31%	0,00%	0,02%
0,03%	95,32%	0,03%	0,00%	0,00%	4,48%	0,00%	0,01%	0,12%	0,00%	0,00%
0,75%	89,41%	0,03%	0,01%	0,04%	4,53%	0,00%	0,02%	4,37%	0,00%	0,83%
95,01%	0,30%	0,07%	0,00%	0,02%	4,50%	0,02%	0,00%	0,07%	0,01%	0,00%
0,24%	0,06%	0,06%	0,00%	0,00%	4,46%	0,00%	0,00%	95,16%	0,00%	0,02%
1,99%	3,49%	0,02%	0,00%	0,00%	4,48%	0,00%	0,00%	89,48%	0,22%	0,32%
95,16%	0,06%	0,12%	0,00%	0,01%	4,48%	0,02%	0,01%	0,14%	0,02%	0,00%
0,06%	0,07%	95,28%	0,02%	0,00%	4,49%	0,00%	0,02%	0,06%	0,00%	0,00%
0,05%	1,88%	0,06%	0,03%	0,00%	4,45%	0,01%	0,00%	93,21%	0,00%	0,31%

Abbildung 4: Prozentuale Basenverteilung für Triplett 1-3 der Reversen Transkriptase

Ausgehend von dieser Anfangssequenz und unter Berücksichtigung möglicher Insertionen, wurden die Positionen / Triplets nummeriert und die Resistenz-assoziierten Codons markiert (6.1.1). Insertionen lassen sich dadurch erkennen, dass die Gesamtanzahl der Sequenzen im Vergleich deutlich geringer ist, da nur wenige Sequenzen diese zusätzlichen Nukleotide enthalten. Basierend auf den Sequenzalignments wurde für die PR gleichermaßen vorgegangen, wie für die RT beschrieben.

6.2 Chiplayout

Theoretisch benötigt die Detektion mittels APEX nur eine oder, im Falle der Abdeckung beider DNA-Stränge, zwei Sonden (= Primer) pro nachzuweisende Nukleotidposition, allerdings nur unter der Voraussetzung, dass die Sequenzregion strangauf- bzw. abwärts konserviert vorliegt. Sequenzvariationen, die in dem Bereich liegen, den die Sonde abdeckt, sind nicht zwangsläufig für den Nachweis von Interesse. Für das Design der Sonden müssen sie allerdings berücksichtigt werden, um die Hybridisierung an variable Ziel-DNA zu gewährleisten. Aufgrund der ausgeprägten Variabilität des HI-Virus ist die Verwendung von mehreren Oligonukleotiden für ein zu detektierendes Nukleotid unumgänglich. Im Einzelfall einer Analyse werden jedoch nur jeweils eine bis wenige Sonden an das jeweilige PCR-Produkt hybridisieren. Handelt es sich um Probenmaterial unbekannter Sequenz, kann nicht vorausgesagt werden, welche Sonden enzymatisch verlängert werden.

Die äußerst hohe Sequenzvariabilität bedingt eine Einschränkung der Abdeckung der Variationen durch die Primer auf dem Microarray. Durch jede variable Nukleotidposition innerhalb der Primer potenziert sich die Anzahl der Sequenzmöglichkeiten.

6.2.1 Festlegen der Sequenzgrundlage

Ausgehend von den in 6.1.2 beschriebenen Verteilungen konnten die theoretisch notwendigen Sondenzahlen zum Nachweis der Resistenz-assoziierten Positionen für verschieden hohe Sequenzabdeckung und beide DNA-Stränge errechnet werden. Der Mutationsnachweis mittels APEX erfordert das Enden der Sonden an der -1 Position des nachzuweisenden Nukleotids. Abhängig von reverse- oder forward-Orientierung unterscheiden sich die Variationsmöglichkeiten zum Teil erheblich. Es wurde für das Ableiten der Oligonukleotide an jeder nachzuweisenden Position zunächst immer auf die Orientierung zurückgegriffen, die die wenigsten Variationen aufwies. Die Möglichkeit der Detektion über beide Stränge wurde wegen der hohen Variabilität aus Kostengründen nicht verfolgt.

HIV-Protease

Für die PR wurden in der Theorie die Sondenzahl für unterschiedlich hohe Sequenzabdeckungen (99-90%) errechnet (Tabelle 11). Selbst bei Beschränkung auf Abdeckung des weniger variablen Stranges, wären zur Berücksichtigung aller Varianten, die

häufiger als 1% vertreten sind, noch mehr als 200 000 Oligonukleotide notwendig. Vernachlässigt man alle Sequenzen <2%, sind es 9196 Sonden, bei einer Abdeckung von 97% 2635 und bei 96% immerhin noch 961 Oligonukleotide. Aus den Zahlen in Tabelle 11 geht hervor, dass eine Abdeckung durch Oligonukleotide erst ab 95% sinnvoll erscheint, will man das eigentliche Ziel eines kostengünstigen Resistenz-Assays nicht aus den Augen verlieren.

Tabelle 11: Theoretische Anzahl der Sonden bei verschieden hoher Sequenzabdeckung (PR)

Aus der Tabelle gehen die theoretischen Sondenzahlen hervor.
Den Sequenzen liegt jeweils nur der DNA-Strang mit der geringeren Variabilität zu Grunde.

Sequenzabdeckung	theoretisch notwendige Sondenanzahl	
	20mere (für einen Strang)	30mere (für einen Strang)
99%	ca. 206 000	
98%	9196	
97%	2635	
96%	961	
95%	384	831
90%	159	
	1515 (beide Stränge)	

HIV-Reverse Transkriptase

Gleichermaßen wurde für die codierende Region der Reversen Transkriptase vorgegangen, allerdings beschränkt auf die Variationen >95% bzw. >90%. Aus den errechneten Summen geht hervor, dass die RT-Region im Vergleich zur PR eine noch deutlich höhere Sequenzvielfalt aufweist. Eine Sequenzabdeckung von 95% würde eine Sondenanzahl von 128 728 bei einer Länge von 20mer-Oligonukleotiden verlangen, für 30mere steigt die Zahl der Variationsmöglichkeiten auf über 3 Millionen. Aufgrund dieser immensen Vielzahl an Kombinationen musste die Sequenzabdeckung auf 90% weiter eingeschränkt werden. Berücksichtigt werden nun nur noch Sequenzen, die häufiger als 10% in der Datenbank vorkommen. Wie aus Tabelle 12 hervorgeht, ergeben sich dann trotzdem immer noch 2142 (20mere) bzw. 171 677 Sonden (30mere), die erforderlich sind, um alle theoretisch möglichen Kombinationen zu berücksichtigen.

Tabelle 12: Theoretische Anzahl der Sonden bei verschiedener hoher Sequenzabdeckung (RT)

Aus der Tabelle gehen die theoretischen Anzahlen bei unterschiedlicher Sondenlänge hervor. Den Sequenzen liegt jeweils nur der DNA-Strang mit der geringeren Variabilität zu Grunde.

Sequenzabdeckung	theoretisch notwendige Sondenanzahl	
	20mere (für einen Strang)	30mere (für einen Strang)
95%	128 728	>3 Millionen
90%	2142	171 677

Diese immensen Zahlen vermitteln einen Eindruck von der enormen Variabilität des viralen Genoms. Aufgrund dessen mussten neben der Einschränkung auf Sequenzen mit einer bestimmten Häufigkeit in der Datenbank weitere Strategien entwickelt werden, die eine Verringerung der benötigten Primeranzahlen ermöglichen sollten (6.5).

6.2.2 Erweiterung um beschriebene Mutationen

Nach der Limitierung der Sequenzvariationen auf ein Vorkommen von >10% (RT) bzw. >5% (PR), mussten im Folgenden allerdings wieder Mutationen ergänzt werden, denen in der Literatur die Assoziation mit einer Resistenzentwicklung nachgesagt wird, unabhängig von ihrer Präsenz in der Datenbank. Diese Vervollständigungen betreffen natürlich nur die nachzuweisenden Triplets, nicht die Regionen strangauf- oder abwärts.

Hierfür wurde wiederum auf die Stanford HIV-Datenbank zurückgegriffen. Unter <http://hivdb.stanford.edu/cgi-bin/PIResiNote.cgi> (Stand 13.08.2004) findet man zu jeder Resistenz-assoziierten Triplettsposition innerhalb der PR und RT mögliche Aminosäure-Veränderungen und deren Auswirkungen auf das Resistenzprofil. Diese Informationen liegen allerdings nur auf Amino-, nicht auf Nukleinsäureebene vor. Resistenzmutationen im HIV-Genom werden im Allgemeinen mit einer Zahl und zwei Buchstaben beschrieben. Die Zahl bezeichnet die Position des veränderten Codons. Der erste Buchstabe steht vor der Zahl und gibt die durch das entsprechende Tripletts kodierte Aminosäure innerhalb der Proteinsequenz im Wildtyp-Virus an. Der zweite Buchstabe nach der Zahl kennzeichnet die durch die Mutation gebildete Aminosäure.

Mit Hilfe von Abbildung 5 soll am folgenden Beispiel das weitere Vorgehen erläutert werden. Für Tripletts / AS 10 der Protease beschreibt die Datenbank die Mutationen L10I, L10V, L10F

und L10R, also den Austausch von Leucin durch Isoleucin, Valin, Phenylalanin oder Arginin (Abbildung 5, Spalte 2 und 4). Innerhalb der zugrunde liegenden Wildtyp-Sequenz (NL 4-3) wird diese Aminosäure durch das Triplet CTC codiert (Spalte 3). Isoleucin kann durch die Triplets ATT, ATC und ATA dargestellt werden (Spalte 5). Eine Berücksichtigung aller möglichen Dreiergruppen für das Sondendesign hätte wiederum zu einem außerordentlichen und nicht realisierbaren Anstieg der benötigten Anzahl geführt. Deshalb wurde von der Annahme ausgegangen, dass der Austausch der Aminosäure durch eine möglichst einfache Mutation zustande kommt, also im einfachsten Fall durch die Veränderung nur eines der drei Nukleotide. Diese Möglichkeit wurde als *wahrscheinlichstes Codon* (Spalte 6) angenommen und, soweit nicht über dem festgelegten Schwellenwert repräsentiert, in die Liste der vorkommenden Sequenzvariationen integriert, auf der das Sondendesign basiert. Im Beispiel L10I wurde also von einer Mutation von CTC zu ATC ausgegangen, da diese nur den Austausch des ersten Nukleotids voraussetzt. Gleiches gilt für L10V (CTC-GTC), L10F (CTC-TTC) und L10R (CTC-CGC), wobei für Arginin das zweite Nukleotid verändert werden muss. Für alle Resistenz-assoziierten Triplets / Aminosäurepositionen der PR und RT (insgesamt 59) wurde analog vorgegangen.

Triplet / AS	WT	Codon NL 4-3	Mutation	mögliche Codons	wahrscheinlichstes Codon
10	Leucin	CTC	Isoleucin	ATT ATC ATA	ATC
			Valin	GTT GTC GTA GTG	GTC
			Phenylalanin	TTT TTC	TTC
			Arginin	CGT CGC CGA CGG AGA AGG	CGC

Abbildung 5: Erweitern der Sequenzvariationen anhand beschriebener AS-Veränderungen (am Beispiel von Triplet / AS 10 innerhalb der HIV-Protease)

Die in der Datenbank vorkommenden Sequenzvariationen werden durch beschriebene Resistenz-assoziierte Mutationen ergänzt. Erklärung siehe Text.

Durch diese Ergänzungen berechnen sich indessen die theoretisch erforderlichen Sondenzahlen für den Microarray neu. Tabelle 13 zeigt die nach oben korrigierten Werte. Die Zeilen 1+2 geben noch einmal die ursprünglichen Zahlen im Vergleich zu denen an, die unter Einbeziehung der beschriebenen Mutationen errechnet wurden (Zeile 3+4). Für die PR-Region erhöhen sie sich in Abhängigkeit von der Primerlänge auf 4410 (20mere) bzw. 66264 (30mere). Im Falle der RT reduziert sich die Summe zunächst bei Verwendung von 20meren auf 2378, also in etwa vergleichbar mit 2142 Sonden, die bei gleicher Länge zur Abdeckung 90% der Sequenzen notwendig sind. Verlängert man in der Theorie auf eine Länge auf 30 Basen, steigt die Anzahl immens auf einen Wert von ca. 4 Millionen.

Tabelle 13: Erhöhung der Sondenzahl durch Berücksichtigung bekannter Mutationen

Gezeigt sind die theoretisch erforderlichen Sondenzahlen bei einer Primerlänge von 20- bzw. 30Basen
 Zeile 1+2: bei einer Sequenzabdeckung von 95% (PR) und 90% (RT)
 Zeile 3+4: unter Einbeziehung von Mutationen, die als Resistenz-assoziiert gelten (Frequenz-unabhängig)

Sonden	PR-Sequenzen 95%	RT-Sequenzen 90%
20mere	384	2142
30mere	861	171 677
20mere + Mutationen	4410	2378
30mere + Mutationen	66264	~4 Millionen

Die unmittelbare Darstellung aller Sequenzen auf dem Microarray, wäre, je nach Sondenlänge und Sequenzauswahl, finanziell und technisch nur schwer oder nicht möglich.

6.3 Optimierung der APEX-Reaktion für HIV-Sequenzen

Die Etablierung der APEX-Reaktion zum Nachweis Resistenz-assoziiierter Mutationen im HIV-Genom wurde anhand einer Modellsequenz durchgeführt. Dafür wurde keine synthetische Sequenz ausgewählt, sondern Oligonukleotidsonden, die innerhalb der HIV-Protease Nukleotid 60 in reverser Orientierung abfragen und der Wildtyp-Sequenz entsprechen (NL4-3, s. Anhang). Ausgehend von den von Kurg *et al.* (2000) beschriebenen Bedingungen für die Primer Extension wurde zunächst kein Ergebnis erzielt. Aus diesem Grund musste die APEX-Reaktion zunächst angepasst werden. Anhand der Modellreaktion wurden folgende Reaktionsparameter und Herstellungsbedingungen der Microarrays optimiert:

- Oberflächenbeschichtung der Glasträger
- Spottingpuffer
- Sondenlänge und 5'-C6-Aminomodifizierung
- Nachbehandlung der Microarrays nach dem Spotten
- Temperaturverlauf der Reaktion

Hierfür wurden parallel zahlreiche Versuche durchgeführt. Davon sind im Folgenden diejenigen gezeigt, die zu den letztlich am besten geeigneten Versuchsbedingungen geführt haben.

Generell gelten für die Optimierungsversuche folgende Parameter. Um den Einfluss der Sondenlängen bzw. der T_m zu testen, variiert die Längen der Sonden je nach Versuch zwischen 20-30 Basen spezifischem 3'-Sequenzanteil. Ein Oligo-d(T)-Spacer zwischen 20 und 30 Nukleotiden soll sterische Hinderung an der Festphase minimieren. Durch diesen Spacer haben alle Sonden eine konstante Gesamtlänge von 50 Basen. Zur Überprüfung der Reaktion wird als Ziel-DNA ein Wildtyp-PCR-Produkt (NL4-3) eingesetzt, kommt es bei komplementärer Verlängerung durch die Polymerase zum Einbau eines ddCTPs. Alle APEX-Optimierungsversuche basieren auf der Amplifikation der Protease-Sequenz aus Plamid pNL4-3. Nach der APEX-Reaktion wurden die Arrays bei allen vier Wellenlängen ausgelesen, erwartungsgemäß konnten nur bei 532 nm spezifische Signale detektiert werden. Für Diagramme gilt: die Rohdaten wurden durch Subtraktion des lokalen Hintergrunds korrigiert. Die gezeigte Signalintensität wurde immer aus dem Mittelwert der jeweiligen Spottingreplikate errechnet, Fehlerbalken ergeben sich aus der zugehörigen Standardabweichung. Bei Darstellung von Microarrays als Falschfarbenbilder, korreliert die Signalintensität mit den gezeigten Farbspektren von blau über grün und rot bis hin zur Sättigung im weißen Bereich. Soweit nicht anders angegeben, entsprechen die Versuchsbedingungen den im Methodenteil beschriebenen.

6.3.1 Temperaturprofil

Zunächst musste ein geeigneter Temperaturverlauf für die APEX-Reaktion etabliert werden. Kurg *et al.*, 2000 beschreiben für die Durchführung der APEX-Reaktion ein einstufiges Temperaturprofil von 15 min bei 58°C. Unter diesen Bedingungen konnte allerdings keine Verlängerung der verwendeten Testsonden (20-25 Basen spezifischer 3'-Sequenzanteil) erreicht werden. Da das Temperaturoptimum der verwendeten ThermoSequense bei 72°C liegt, wurden anschließend zunächst verschiedene Temperaturen (58°C, 60°C, 70°C und 72°C) für die Primer Extension getestet. Die Reaktionszeit lag jeweils bei 15 min. Durch Inkubation bei nur einer Temperatur konnte für kein Experiment eine Verlängerung der Oligonukleotide durch fluoreszenzmarkierte ddNTPs nachgewiesen werden. Auch bei Wiederholung der Versuchsanordnung mit einer vorgeschalteten Denaturierung für 30 sec bei 95°C konnten keine Signale detektiert werden. Im Folgenden wurden deshalb verschiedene zweistufige Temperaturprofile durchgeführt, die sich aus einem Hybridisierungs- und einem Verlängerungsschritt bei unterschiedlicher Temperatur zusammensetzten. Tabelle 14 zeigt

eine Übersicht über die verwendeten Profile, dabei gibt der erste, niedrigere Wert die Hybridisierungstemperatur an, der zweite und höhere die Temperatur für die Verlängerungsreaktion. Die Experimente wurden in einem ThermoCycler mit einem Aufsatz für Glasobjektträger durchgeführt, der sich analog einem herkömmlichen PCR-Cycler programmieren lässt. Der Hybridisierungsschritt hatte eine Dauer von jeweils 10 min, die enzymatische Verlängerung von je 15 min. Insgesamt wurden 12 zweistufige Experimente im Doppelansatz durchgeführt, die Hybridisierungstemperaturen von 20°C, 30°C, 40°C und 50°C wurden jeweils kombiniert mit den Temperaturen für die Extension (60°C, 65°C und 70°C). Allen Temperaturabläufen ging ein Denaturierungsschritt von 30 sec bei 95°C voraus (in Tabelle 14 nicht gezeigt).

Tabelle 14: Übersicht über die getesteten Temperaturprofile für die Primer Extension-Reaktion

Zu Optimierungszwecken wurden verschiedene Temperaturprofile für die APEX-Reaktion getestet.
 erste Temperaturangabe: Hybridisierung, jeweils für 10 min.
 zweite Temperaturangabe: enzymatische Verlängerung jeweils für 15 min.

20°C / 60°C	20°C / 65°C	20°C / 70°C
30°C / 60°C	30°C / 65°C	30°C / 70°C
40°C / 60°C	40°C / 65°C	40°C / 70°C
50°C / 60°C	50°C / 65°C	50°C / 70°C

Abbildung 6 zeigt zwei Microarrays nach dem Auslesen bei 532 nm. Die Oligonukleotide sind in jeweils fünf Replikaten immobilisiert und tragen keine 5'-C6-Aminomodifikation. Zusätzlich sind Kontrollen in das Raster integriert, zum einen die so genannten Positionskontrollen (6.4.1, weiß markiert, jeweils erste und letzte Reihe), diese dienen der Orientierung innerhalb des Arrays. Zum anderen handelt es sich um Sonden (20-25mere), die aufgrund ihrer Lage nicht mit ddCTP verlängert werden. Sie werden als Negativkontrollen bezeichnet (unterer Teil der Arrays, grau markiert).

Gezeigt wird die Durchführung der APEX-Reaktion unter den Ausgangsbedingungen, wie in der Literatur (Kurg *et al.*, 2000) angegeben (einstufig: 58°C, links in der Abbildung) und denen, die im Vergleich die stärksten Signale lieferten (zweistufig: 40°C / 65°C, rechts in der Abbildung). Der linke Array zeigt ausschließlich Signale, die nicht stärker sind als die Negativkontrollen im jeweils unteren Teil der Arrays, während der rechte im oberen Teil deutliche Signale erkennen lässt. Gleichzeitig wurde der Einfluss verschieden langer Oligonukleotide zur Verlängerung getestet (20-25mere). Die Signalintensitäten nehmen unter den verwendeten Bedingungen mit wachsender Sondenlänge zu. Die Negativkontrollen

werden auch unter optimierten Temperaturbedingungen nicht durch die Polymerase verlängert, die Reaktion verläuft also spezifisch.

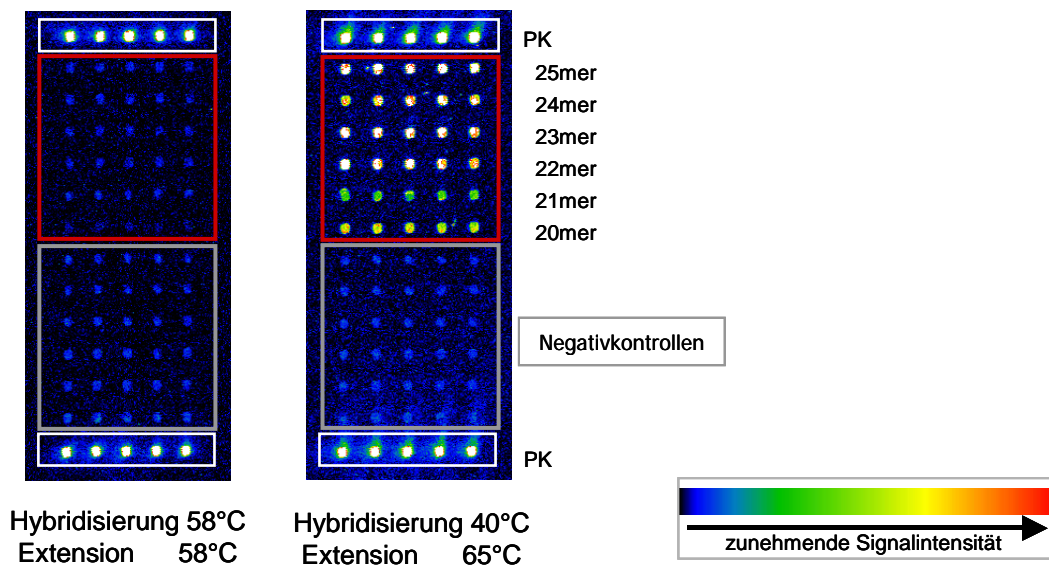


Abbildung 6: Vergleich verschiedener Temperaturprofile zur Durchführung der APEX-Reaktion

Enzymatische Verlängerung von Sonden unterschiedlicher Länge (rot) unter verschiedenen Bedingungen

links: Hybridisierung und Verlängerung bei 58°C (Kurg *et al.*, 2000)

rechts: Hybridisierung bei 40°C, Verlängerung bei 65°C (optimiert)

PK (weiß): Positionskontrollen (6.4.1)

Negativkontrollen (grau): Sonden, die aufgrund ihrer Sequenz nicht mit Cy3-ddCTP verlängert werden.

Für alle anderen zweistufigen Temperaturabläufe (nicht gezeigt), konnten Ergebnisse unterschiedlicher Qualität erreicht werden, in jedem Fall konnten aber spezifische Signale detektiert werden. Während aus niedrigeren Hybridisierungstemperaturen eher unspezifische Signale resultierten, bewirkte die Erhöhung auf 50°C eine Abschwächung der Signalintensitäten. Letzteres gilt auch für die Erhöhung der Extensions-Temperatur auf 70°C. Dies ist zurückzuführen auf eine im Vergleich niedrigere T_m der Oligonukleotidsonden und dem Dissoziieren von der Ziel-DNA. Die Reaktionseffizienz bei 60°C und 65°C war vergleichbar; um unspezifischer Verlängerung vorzubeugen, wurde die höhere Temperatur bevorzugt. Für alle folgenden Experimente wurde, soweit nicht anders angegeben, das ausgewählte Temperaturprofil (40°C / 65°C) verwendet.

6.3.2 Spottingpuffer

Die Herstellungsbedingungen der Microarrays können das Ergebnis der Primer Extension-Reaktion entscheidend beeinflussen. Eine wesentliche Rolle spielt dabei der Puffer, in dem die Sonden auf der Chipoberfläche immobilisiert werden.

Getestet wurden hierfür sechs verschiedene Spottinglösungen (Abbildung 7), von denen drei kommerziell erhältlich sind (1-3) und drei gemäß den Angaben im Methodenteil hergestellt wurden (4-6). Auf Epoxysilan-beschichteten Glasoberflächen wurden Sonden mit und ohne 5'-C6-Aminomodifikation und mit verschiedenen langen spezifischen 3'-Sequenzanteilen immobilisiert (20, 25 und 30mere). Abbildung 7 stellt diesen Teil der Ergebnisse dar.

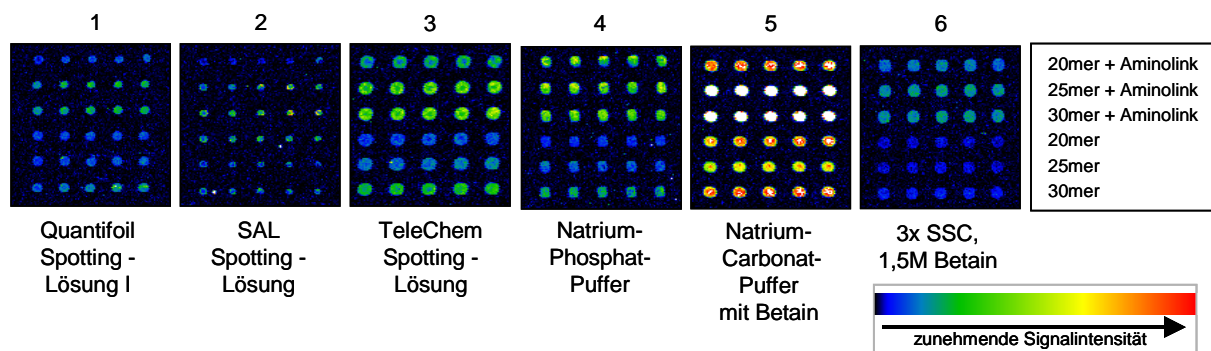


Abbildung 7: Vergleich verschiedener Spottinglösungen (I)

Enzymatische Verlängerung von Oligonukleotidsonden verschiedener Länge mit und ohne 5'-C6-Aminolink in unterschiedlichen Spottinglösungen auf Epoxysilan-Oberflächen. Darstellung des bei 532 nm ausgelesenen Microarrays in Falschfarben.

Im direkten Vergleich zeigen sich die schwächsten Signalintensitäten beim Immobilisieren der Sonden in der Spotting-Lösung der Firma Quantifoil (1), der Firma Asper (2) und in der SSC / Betain-Lösung (6). Zudem kristallisiert die verwendete SAL-Spottinglösung (2) aus und bildet sehr kleine, unregelmäßig geformte Spots. Durch Zusatz von Betain (5+6) erreicht man hingegen gleichmäßige Spotformen und -größen. Vergleichsweise stärkere Intensitäten können durch das Immobilisieren in der Spottinglösung der Firma TeleChem (3) und in Natrium-Phosphatpuffer (4) erzielt werden, wobei die Spotdurchmesser mit der kommerziell erhältlichen Lösung etwas größer ist. Die deutlich besten Signalintensitäten wurden bei Verwendung von Natrium-Carbonat-Spottingpuffer (pH 9, mit Betain) erreicht (5). Alle Felder (1-6) befanden sich auf einem Microarray und wurden unter gleichen technischen Bedingungen (Laserstärke und Verstärkung) ausgelesen.

Ein Teil der in Abbildung 7 gezeigten Daten wird in Diagramm 1 dargestellt. Es zeigt einen Vergleich der gemessenen mittleren Signalintensitäten bei einer Sondenlänge von 30 Basen für die verschiedenen Spottinglösungen und macht zudem den Unterschied zwischen Sonden mit und ohne 5'-C6-Aminomodifikation deutlich. Außer für die TeleChem-Spottinglösung sind die Signalintensitäten für Sonden mit 5'-C6-Aminomodifikation höher im Vergleich zu den unmodifizierten Sonden. Das Resultat konnte in vier Wiederholungen reproduziert werden. Es ist deutlich zu erkennen, dass mit Sonden, die in Natrium-Carbonatpuffer immobilisiert wurden, die höchsten Signalstärken erreicht wurden (violett), während die anderen signifikant schwächere Intensitäten aufweisen. Für alle weiteren Experimente wurden die Sonden auf den Miroarrays in Natrium-Carbonatpuffer immobilisiert.

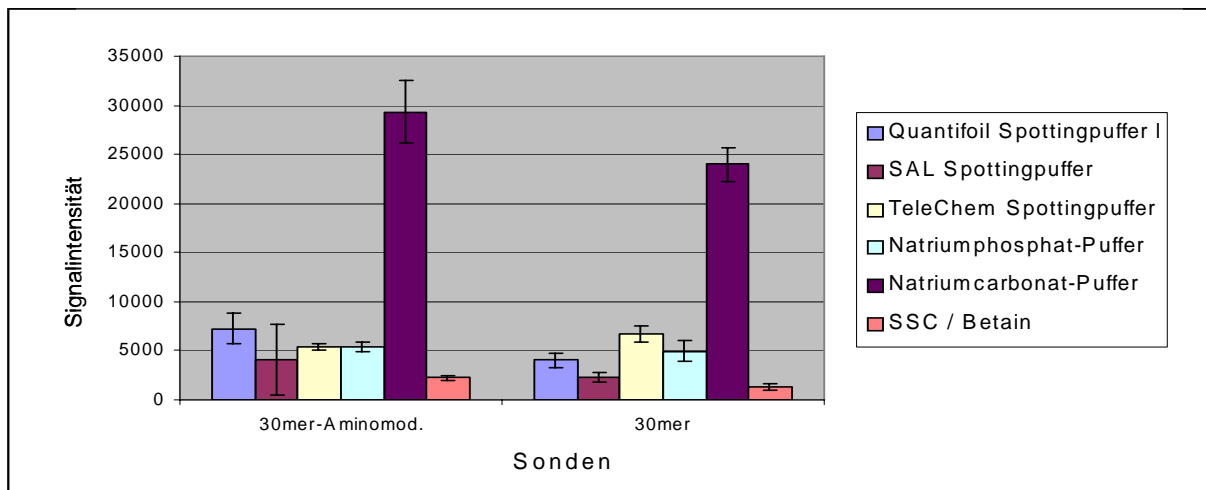


Diagramm 1: Vergleich verschiedener Spottinglösungen (II)

Gemessene Signalintensitäten nach Verlängerung von 30meren mit und ohne 5'-C6-Aminomodifikation in unterschiedlichen Spottinglösungen auf Epoxysilan-Oberflächen. Vergleich der gemessenen Intensitätswerte.

6.3.3 Oberflächen

Um optimale Ergebnisse für die Reaktion zu erzielen, wurden parallel zu den Spottingpuffern auch verschiedenen Oberflächenmodifizierungen getestet. Da diese beiden Parameter abhängig voneinander sind, wurden verschiedene Oberflächen- / Spottingpuffer-Kombinationen getestet. Abbildung 8 zeigt nur den Vergleich zwischen Amino- und Epoxysilan beschichteten Glaträgern (Schott Nexterion Slide A bzw. E). Sonden verschiedener Länge wurden in Natriumcarbonat-Puffer (6.3.2) in je fünf Replikaten aufgebracht. Die beiden Microarrays wurden mit gleicher Laserintensität ausgelesen. Auf der

Aminosilan-Oberfläche (rechts) können abhängig von den Sonden nur schlechte bis mittlere Signal- / Hintergrundverhältnisse erreicht werden. Dagegen zeigen alle Sonden auf der Epoxysilan-Oberfläche (links) Signale im Sättigungsbereich unter den gleichen Bedingungen. Von allen getesteten Kombinationen lieferte die Immobilisierung der Sonden in Natriumcarbonat-Spottinpuffer auf Epoxysilan-Oberflächen die besten Ergebnisse.

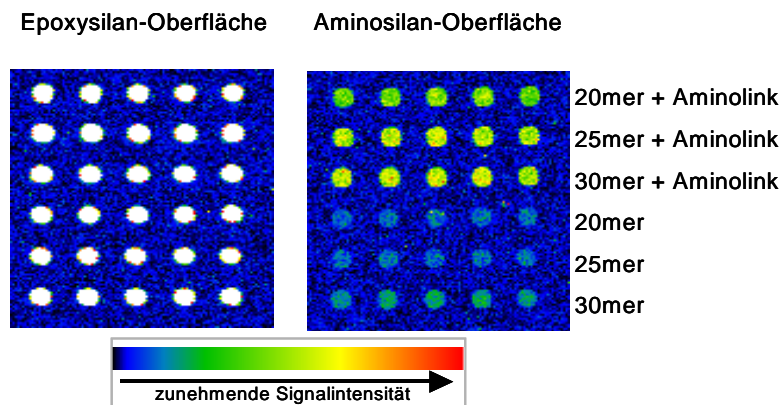


Abbildung 8: Vergleich von Epoxy- und Aminosilan-beschichteten Glasoberflächen

Einfluss unterschiedlicher Oberflächenchemie auf die Signalstärke. Verwendet wurden Sonden unterschiedlicher Länge mit und ohne 5'-C6-Aminomodifikation.

6.3.4 Aminomodifikation und Sondenlänge

Eine wichtige Rolle für die Optimierung der APEX-Reaktion spielt die Länge der verwendeten Sonden bzw. ihre Schmelztemperatur und die Art der Anbindung an die Oberfläche. Um sterische Hinderung bei der Verlängerungsreaktion zu vermindern, wurden abweichend von Kurg *et al.* (2000) Sonden mit 5'-Oligo-d(T)-Spacer verwendet. Dieser Spacer ist nicht komplementär zur HIV-spezifischen Ziel-DNA und soll nur den Mindestabstand von der Festphase gewährleisten. Weiterhin wurde der Einfluss von 5'-C6-Aminolinkern zur Anbindung an den Glasträger auf die Signalintensität nach Primer Extension-Reaktion getestet.

Verglichen wurden Oligonukleotide mit einer Länge von 20, 25 und 30 Basen mit und ohne 5'-C6-Aminomodifikation. Die angegebene Länge bezieht sich nur auf den spezifischen Sequenzanteil am 3'-Ende, durch verschiedene Längen des 5'-Oligo-d(T)-Spacers beträgt die Gesamtlänge jeweils 50 Basen. Die Schmelztemperaturen der Sonden betragen 44°C (20mer), 57°C (25mer) und 64°C (30mer), der G/C-Gehalt lag zwischen 40 und 44%. Die

Hybridisierung fand bei 40°C, die Verlängerungsreaktion bei 65°C statt (6.3.1). Diagramm 2 zeigt den Vergleich der Mittelwerte der Signalintensitäten aus fünf Replikaten.

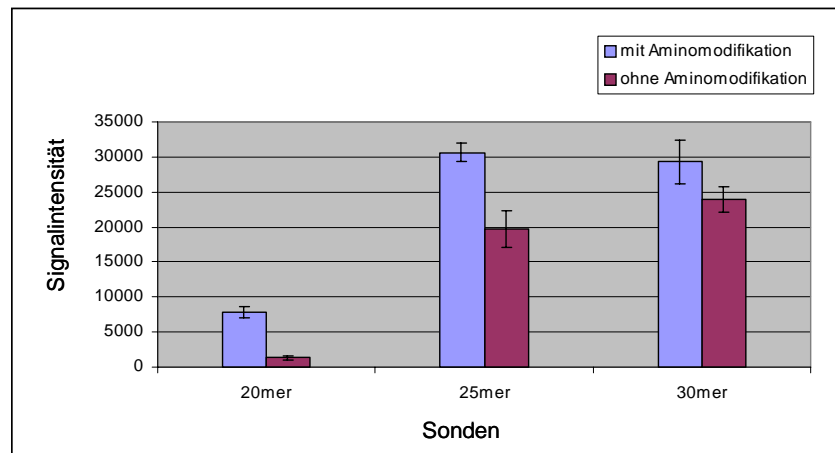


Diagramm 2: Vergleich der Signalintensitäten von 5'-aminomodifizierten und nicht modifizierten Sonden verschiedener Länge

Immobilisiert wurden 20-, 25- und 30mer-Sonden mit und ohne 5'-C6-Aminomodifikation auf Epoxysilan-Oberflächen in Natriumcarbonat-Puffer. Vergleich der Signalstärken nach der APEX-Reaktion bei 532 nm.

Aufgrund der niedrigen T_m und den gewählten Reaktionsbedingungen zeigen die verlängerten 20mer-Oligonukleotide erwartungsgemäß eine wesentlich schwächere Signalintensität nach der APEX-Reaktion als die Sonden mit höherer Schmelztemperatur. Zurückzuführen ist dies auf ein vorzeitiges Dissoziieren von der Ziel-DNA bei Erhöhung der Temperatur von 40°C (Hybridisierung) auf 65°C (Extension). Durch die höhere T_m des Primers kann die Intensität der gemessenen Signale deutlich gesteigert werden (25 und 30mer). Zwischen Extensionstemperatur und Schmelztemperatur der Sonden bestehen Wechselwirkungen. In dieser Arbeit werden Oligonukleotide mit leicht variierender T_m verwendet, wichtig für die Stabilität der Reaktion ist, dass die Hybridisierung und Verlängerung auch dann gewährleistet bleiben.

Bei Vergleich von 5'-aminomodifizierten (blau) und unmodifizierten (rot) Oligonukleotiden kann man leicht erkennen, dass eine Aminomodifikation die Signalintensität deutlich erhöht. Durch die gezielte Anbindung des 5'-Endes über die reaktive NH_2 -Gruppe an die endständige Epoxygruppe der Oberflächenchemie verbessert sich die Kopplungseffizienz und -spezifität.

6.3.5 Nachbehandlung der Microarrays

Ebenfalls wichtig für das Erzielen optimaler Resultate für die APEX-Reaktion ist die Nachbehandlung der Microarrays nach dem Spotten. Verwendet wurden kommerziell erhältliche Epoxysilan-beschichtete Glasobjektträger (Slide E) der Firma Schott Nexterion (Mainz). Das vom Hersteller vorgeschriebene Protokoll empfiehlt die Inkubation der Chips nach dem Spotten über Nacht in einer feuchten Kammer, also in einer Umgebung mit gesättigter Luftfeuchtigkeit bei Raumtemperatur.

Um eine möglichst hohe Spotdichte zu erreichen, werden in dieser Arbeit Natriumcarbonat-Spottingpuffer, sowie Spotgrößen von ca. 100 µm und -abstände von 200 µm verwendet. Dies führt bei dieser Nachbehandlung allerdings zur Wasseraufnahme und daraus resultierend zu einem Ineinanderlaufen der einzelnen Spots. Um dies zu verhindern, wurden die Microarrays im Anschluss an den Spottingvorgang trocken bei Raumtemperatur oder trockener Hitze unterschiedlich lange inkubiert. Verglichen wurden folgende Nachbehandlungen:

- Verwendung der Arrays für APEX nach einstündiger Inkubation bei Raumtemperatur
- Verwendung nach einstündiger Inkubation bei 60°C im Hybridisierungssofen
- Inkubation über Nacht bei Raumtemperatur, dann direkte Verwendung
- Inkubation für 1h bei 60°C, dann über Nacht bei Raumtemperatur

Diagramm 3 zeigt einen Vergleich der daraus resultierenden, mittleren Intensitätswerte. Es wurden 30mer-Oligonukleotide mit und ohne 5'-C6-Aminomodifikation in jeweils fünf Replikaten auf die Oberfläche immobilisiert. Alle Chips wurden mit der gleichen Laserintensität ausgelesen.

Die schwächsten Signale lieferten die Microarrays, die direkt nach Herstellung für die APEX-Reaktion verwendet wurden (blau). Durch Inkubation für 1h bei 60°C erreicht man eine Erhöhung der Signalintensität (rot). Beste Signal- / Rauschverhältnisse liefern Arrays, die über Nacht aufbewahrt wurden, wobei es wiederum keinen erkennbaren Unterschied macht, ob ein Inkubationsschritt von 1h bei 60°C vorangeht oder nicht (gelb und türkis). Wichtig für die kovalente Anbindung an die Oberfläche ist offensichtlich ein Mindestzeitrahmen, in dem die Reaktion ablaufen kann. Das Experiment wurde zweimal wiederholt, die Ergebnisse konnten reproduziert werden. Wie schon in vorangegangenen Versuchen, liefern Oligonukleotide mit Aminomodifikation am 5'-Ende höhere Signalintensitäten als unmodifizierte Sonden gleicher Sequenz.

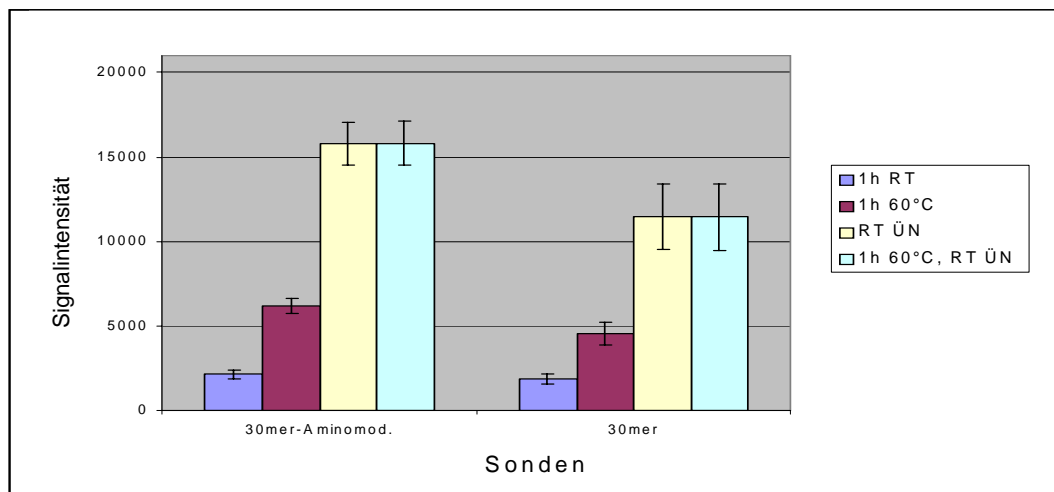


Diagramm 3: Unterschiedliche Nachbehandlung der Microarrays nach dem Spotting

Unterschiedliche Behandlung der Microarrays unmittelbar nach dem Spotting und vor der APEX-Reaktion

6.4 Kontrollsonden

Die Verwendung von Microarrays setzt die eindeutige Identifizierung der Sonden innerhalb des Arrays und eine Qualitätskontrolle einzelner Arbeitsschritte voraus. Zu diesem Zweck wurden zusätzliche Kontrollsonden verwendet. Hierfür wurden HIV-fremde Sequenzen ausgewählt, um eine kompetitive Hybridisierung auszuschließen (*Neurospora crassa*, Gen *al-1*).

6.4.1 Positionskontrollen

Zur Orientierung innerhalb des Arrays wurden so genannte *Positionskontrollen (PK)* an drei von vier Eckpunkten immobilisiert. Es handelte sich dabei um Oligonukleotide, die analog zu den verwendeten ddNTPs mit den Fluoreszenzfarbstoffen Cy3, Cy5, TexasRed und Fluorescein markiert waren. Die Positionskontrollen wurden gemischt und in einer Konzentration von je 12,5 μM in Spottingpuffer an die Oberfläche angebunden. Sie hatten eine Gesamtlänge von 50 Basen, bestehend aus 25 Thymidinen am 3'-Ende und folgenden 25 Basen spezifischer Sequenz (5'-TGGGTATAAACGATTCGTAGATAA-25(T)-3'). Im Gegensatz zu den HIV-spezifischen Sonden, tragen die Kontrollen eine C6-Aminomodifikation zur Anbindung an die Chipoberfläche am 3'-Ende, um einer unspezifischen Verlängerung durch

die Polymerase entgegen zu wirken. Die Farbmarkierung erfolgte am freien 5'-Ende der Sonden und erlaubt die Kontrolle von Anordnung und Kopplung unabhängig von der Verlängerungsreaktion.

6.4.2 Extensionkontrollen

Zusätzlich zu den Positionskontrollen kamen weitere Sonden zum Einsatz, die eine Überprüfung der Primer Extension-Reaktion gewährleisten sollten. Dabei handelt es sich um Oligonukleotide, die aufgrund ihrer Sequenzanordnung in der Lage sind, Haarnadelstrukturen auszubilden. Durch diese Sekundärstruktur fungiert die Sequenz gleichzeitig als Primer und als Template und wird durch die Polymerase verlängert. Diese Reaktion verläuft unabhängig von der eingesetzten Ziel-DNA. Abbildung 9 zeigt den schematischen Aufbau dieser im Folgenden als Extensionskontrollen oder *self extension oligonucleotides (SE)* bezeichneten Sonden.

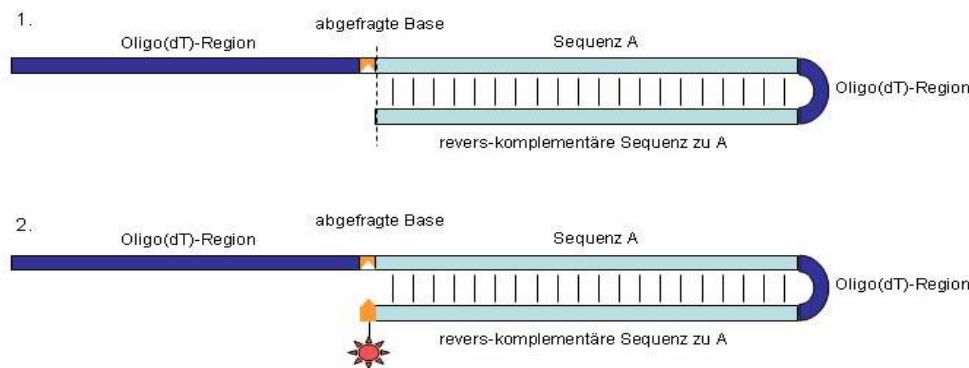


Abbildung 9: Schematische Darstellung der Extensionskontrollen

Zur Kontrolle der Verlängerungsreaktion wurden Oligonukleotide eingesetzt, die aufgrund ihrer Sequenz die Fähigkeit zur Ausbildung von Haarnadelstrukturen besitzen (1). Sie sind in der Lage, sich ohne Ziel-DNA selbst zu verlängern (2).

Wie die übrigen Sonden, tragen die Extensionskontrollen am 5'-Ende eine C6-Aminomodifikation und einen Oligo-d(T)-Spacer, der den Abstand zur Glasoberfläche gewährleistet. Als spezifische Nukleotidfolge (Sequenz A) wurde wie im Fall der Positionskontrollen eine Sequenz gewählt, die keine Komplementarität zu HIV zeigt. Nach einer weiteren, kurzen Oligo-d(T)-Folge, wiederholt sich diese Sequenz A in revers-komplementärer Orientierung, so dass es unter bestimmten Temperaturbedingungen

zur Formation eines partiellen Doppelstranges kommt (Abbildung 9.1). Tabelle 15 zeigt die Sequenzen der Extensionskontrollen. Die beschriebenen Sequenzabschnitte sind durch Leerzeichen verdeutlicht.

Tabelle 15: Sequenzen der Extensionskontrollen

Name	Sequenz 5'-3'
SE_A	TT T GTCCAGATGCTTCGCATTGACG TTT CGTCAATGCGAAGCATCTGGAC
SE_C	TT G GTCCAGATGCTTCGCATTGACG TTT CGTCAATGCGAAGCATCTGGAC
SE_G	TT C GTCCAGATGCTTCGCATTGACG TTT CGTCAATGCGAAGCATCTGGAC
SE_T	TT A GTCCAGATGCTTCGCATTGACG TTT CGTCAATGCGAAGCATCTGGAC

Bei enzymatischer Verlängerung wird ein farbmarkiertes ddNTP komplementär zu einer Base eingebaut, die sich zwischen dem Oligo-d(T)-Anhang und Sequenz A befindet (Abbildung 9.2). Es wurden vier verschiedene Kontrollsonden verwendet, die sich ausschließlich an dieser Position unterscheiden und jeweils eine der vier möglichen Basen beinhalten. Bei der Auswahl der Sequenz war zu beachten, dass nur eine Hybridisierungsmöglichkeit bei der Ausbildung des doppelsträngigen Bereichs besteht. Die Kontroll-Oligonukleotide wurden in gleicher Konzentration und gleichem Spottingpuffer wie die HIV-spezifischen Sonden aufgebracht und konnten aufgrund einer entsprechend gewählten Schmelztemperatur unter identischen Reaktionsbedingungen verlängert werden.

Wegen möglicher sterischer Hinderung durch Bildung der Haarnadelstruktur wurden Oligo-d(T)-Spacer in zwei verschiedenen Längen, 30 und 50 d(T), getestet und miteinander verglichen. Eine Sequenz von 30 Thymidinen am 5'-Ende erwies sich als ausreichend, um die Selbstverlängerung zu ermöglichen (Daten nicht gezeigt). Die Namen der Sonden setzen sich zusammen aus der Bezeichnung *SE* für *self extension* und dem erwarteten Signal, für die entsprechende Base, die eingebaut wird. Die Sequenz *SE_C* beispielsweise trägt also an der Position der abzufragenden Base (Diagramm 4) ein Guanin. Bei komplementärer Verlängerung kommt es zum Einbau eines Cy3-ddCTPs, das bei der entsprechenden Farbstoff-spezifischen Wellenlänge nachgewiesen werden kann.

Diagramm 4 zeigt die Auswertung einer Versuchsanordnung, bei der die vier Kontrollsonden in zwölf Replikaten auf der Chipoberfläche immobilisiert und in Abwesenheit von Ziel-DNA, verlängert wurden. Die Signale wurden bei den vier entsprechenden Wellenlängen detektiert.

Die SE-Kontrollen unterscheiden sich jeweils lediglich durch eine Base (Tabelle 15), die nicht innerhalb des doppelsträngigen Bereichs liegt und somit das Hybridisierungsverhalten nicht beeinflusst. Es konnte also gezeigt werden, dass die Unterschiede der Signalintensitäten nur durch unterschiedliche Quantenausbeuten der verschiedenen Fluoreszenzfarbstoffe

zustande kommen. Anhand dieses Diagramms wird deutlich, dass innerhalb der Datenauswertung eine Normalisierung notwendig ist.

Aus dem Diagramm geht hervor, dass die vier Sonden Template-unabhängig verlängert werden können und die gemessenen Signale mit dem erwarteten Nukleotid-Einbau korrelieren. Die Fehlerbalken errechnen sich aus der Standardabweichung der mittleren Signalintensität von zwölf Replikaten. Diagramm 4 zeigt eine von vier analog durchgeführten Primer Extensions-Reaktionen, für die reproduzierbare Ergebnisse erzielt wurden.

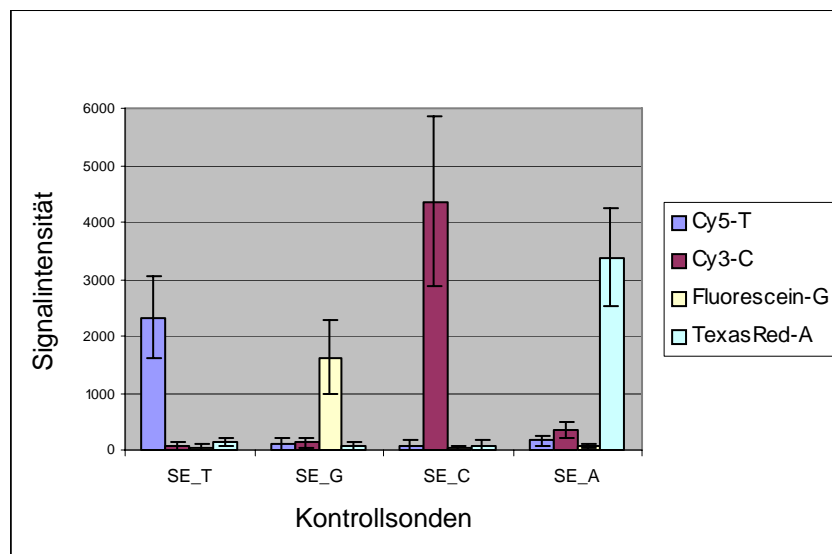


Diagramm 4: Spezifische Signale bei Verlängerung der SE-Kontrollsonden

Vergleich der Signalintensitäten, die durch enzymatische Verlängerung der Extensionskontrollen erreicht werden. Je nach abzufragender Base wird bei entsprechender Wellenlänge ein spezifisches Signal detektiert

6.5 Degenerierte Sonden

Die immense Variabilität der HIV-Sequenzen (6.2) erforderte Strategien, die Anzahl der notwendigen Oligonukleotide auf dem Array zu minimieren, um mit möglichst geringem Sondaufwand eine möglichst breite Sequenzabdeckung zu erreichen. Dazu wurden zwei Lösungsansätze verfolgt, die auf der Modifikation der Oligonukleotidsonden beruhen mit dem Ziel mehrere Sequenzen in einem Spot darzustellen.

Dies kann zum einen erreicht werden durch das Einführen von Sequenzmischungen an einzelnen Positionen („Wobble“). Während der schrittweisen Synthese von DNA-Oligomeren besteht die Möglichkeit, zu bestimmten Zeitpunkten ein Gemisch von zwei, drei oder aller vier Monomeren in die Reaktion einfließen zu lassen. Es kommt an diesen Synthesestufen zu

einem anteiligen Einbau der zugegebenen Nukleotide in die wachsende Nukleinsäurekette. In der Folge sind in der Lösung Sequenzvarianten enthalten, die sich ausschließlich an den betreffenden Positionen unterscheiden, ansonsten aber identisch sind. Solche Modifikationen in einer Sequenz werden in diesem Zusammenhang auch als so genannte *Wobble* bezeichnet. In dieser Art hergestellte Oligonukleotide können als Sonden eingesetzt werden. Nach dem Immobilisieren stellt ein Spot auf dem Array verschiedene Sequenzen dar.

Zum anderen besteht die Möglichkeit, so genannte universelle Basen (z. B. Inosin) als „Platzhalter“ an denjenigen Positionen der Primer einzusetzen, an denen die Sequenz variabel ist. Die Base Inosin (Abbildung 10) zeichnet sich durch die Fähigkeit aus, in einem Nukleinsäure-Doppelstrang mit jeder ihr gegenüberliegenden Base Wasserstoffbrückenbindungen ausbilden zu können.

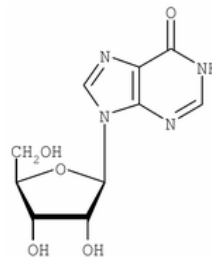


Abbildung 10: Strukturformel von Inosin

Es sollte untersucht werden, inwieweit die Anzahl der Oligonukleotide durch Inosin und „Wobble“-Modifikationen verringert werden kann. Zu diesem Zweck wurde anhand einer „Modellsequenz“ die systematische Sondenmodifizierung und der Einfluss auf die Signalstärke getestet. Gewählt wurde keine synthetische Sequenz, sondern diejenige zum Nachweis von Position 211, Aminosäure 71 (innerhalb der HIV-Protease) in reverser Orientierung. Betrachtet man alle PR-Resistenz-relevanten Positionen (6.1.1) und geht von allen Sequenzen mit einer Präsenz >95% in der HIV-Stanford Datenbank aus, unterliegt diese Region der höchsten Variabilität und eignet sich daher als Modellsystem.

6.5.1 „Wobble“-Modifikationen

Um den Einfluss mehrerer „Wobble“ auf die Signalintensität zu testen, wurden 20mer-Sonden abgeleitet, die die theoretisch vorkommenden Sequenzvarianten an dieser Position abbilden (Tabelle 16). Dabei wurden systematisch die nicht konservierten Positionen durch die

möglichen „Wobble“-Mischungen ersetzt. Unter den oben beschriebenen Bedingungen sind 13 der 20 Nukleotidpositionen der Sequenz variabel. Diese wurde durch die entsprechenden „Wobble“ ersetzt, um die Sequenzvarianten abzubilden, die tatsächlich auftreten können. Die Modifizierungen wurden sowohl vom 3′- als auch vom 5′-Ende des Oligomers kommend eingeführt. Der Sondenname bezeichnet die zunehmende Anzahl der Positionen, die durch „Wobble“ ersetzt sind und die Information, von welchem Ende aus nummeriert wurde. Mit „3“ oder „5“ sind das 3′- bzw. 5′-Ende gemeint (Nr. 2-13). Eine Ausnahme ist W13, hier sind alle 13 nicht-konservierten Nukleotide ersetzt, die Angabe des Sequenzendes entfällt also (Nr. 14). Als Kontrolle dient eine unmodifizierte Sonde, die Position 211 abfragt und vollständig komplementär zur Wildtyp-Sequenz ist (Nr.1).

Tabelle 16: Sequenzen der Modell-Sonden mit zunehmender „Wobble“-Anzahl

Positionen, die keine eindeutigen Sequenzen repräsentieren, sind rot markiert. Die Abkürzungen entsprechen dem IUB-Code.

Nr.	Name	Sequenz 5′-3′ ohne Oligo d(T)-Anteil	Anzahl der „Wobble“
1	WT_reverse_C	TACTAATACTGTACCTATAG	0
2	_{5W1}	YACTAATACTGTACCTATAG	1
3	5W2	YAYTAATACTGTACCTATAG	2
4	5W3	YAYYAATACTGTACCTATAG	3
5	5W4	YAYYARTACTGTACCTATAG	4
6	5W5	YAYYARYACTGTACCTATAG	5
7	5W6	YAYYARYACTGWACCTATAG	6
8	5W7	YAYYARYACTGWRCCCTATAG	7
9	5W8	YAYYARYACTGWRCYTATAG	8
10	5W9	YAYYARYACTGWRCCYYATAG	9
11	W10	YAYYARYACTGWRCCYYDTAG	10
12	5W11	YAYYARYACTGWRCCYYDYAG	11
13	5W12	YAYYARYACTGWRCCYYDYDG	12
14	W13	YAYYARYACTGWRCCYYDYDR	13

1	WT_reverse_C	TACTAATACTGTACCTATAG	0
2	_{3W1}	TACTAATACTGTACCTATAR	1
3	3W2	TACTAATACTGTACCTATDR	2
4	3W3	TACTAATACTGTACCTAYDR	3
5	3W4	TACTAATACTGTACCTDYDR	4
6	3W5	TACTAATACTGTACCYDYDR	5
7	3W6	TACTAATACTGTACYYDYDR	6
8	3W7	TACTAATACTGTRCYYDYDR	7
9	3W8	TACTAATACTGWRCCYYDYDR	8
10	3W9	TACTAAYACTGWRCCYYDYDR	9
11	3W10	TACTARYACTGWRCCYYDYDR	10
12	3W11	TACYARYACTGWRCCYYDYDR	11
13	3W12	TAYYARYACTGWRCCYYDYDR	12
14	W13	YAYYARYACTGWRCCYYDYDR	13

Die aufgeführten Sonden wurden in jeweils zwölf Replikaten auf der Chipoberfläche immobilisiert und unter den optimierten Bedingungen, die im Methodenteil beschrieben sind, enzymatisch verlängert. Als Ziel-DNA wurde die amplifizierte Protease-Sequenz aus Plasmid pNL4-3 eingesetzt. Ausgehend von der Wildtyp-Sequenz wurde für alle Sonden eine Verlängerung um ein ddCTP erwartet.

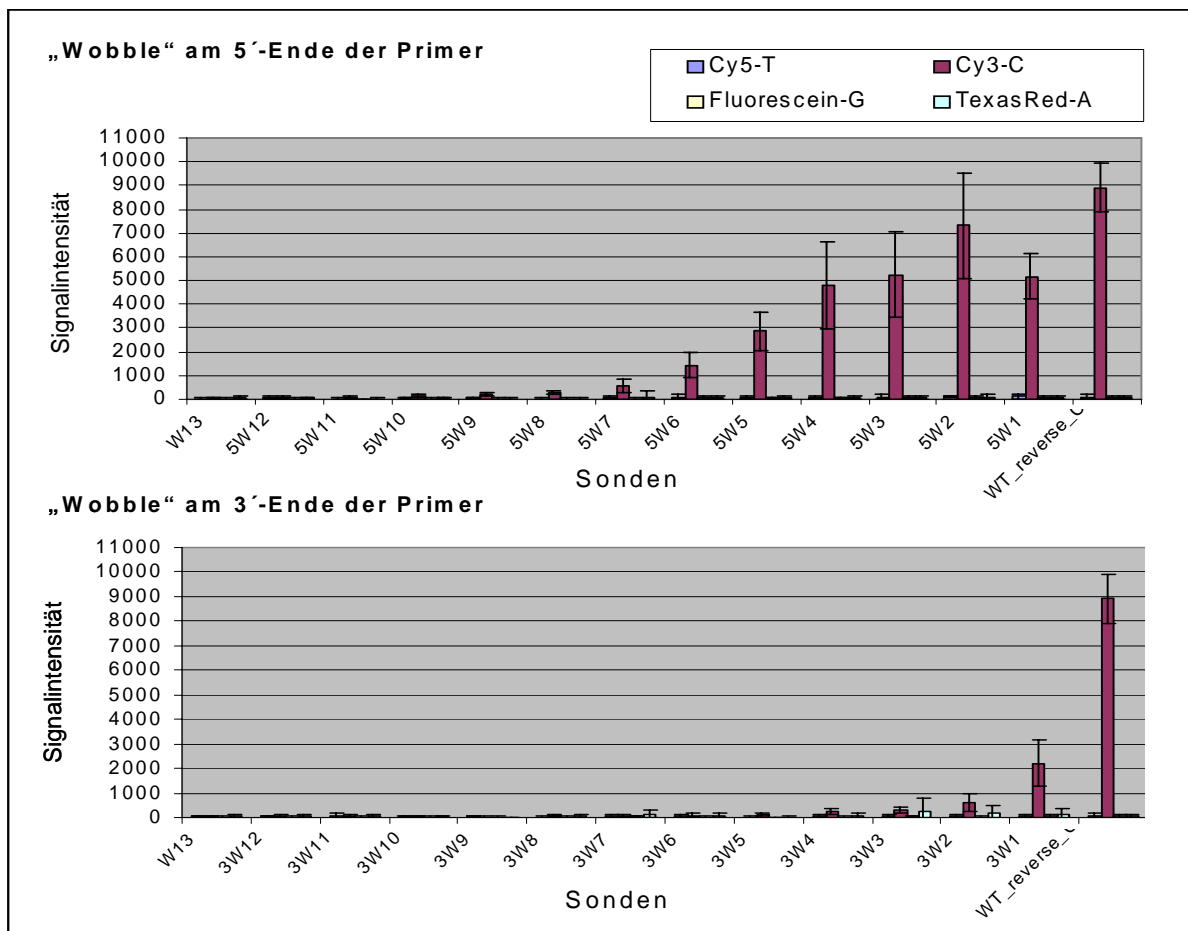


Diagramm 5: Einfluss von „Wobble“-Positionen auf die Signalstärke

Enzymatische Verlängerung von Sonden mit ansteigender Anzahl an „Wobble“-Modifikationen vom 5'-Ende (oben) bzw. vom 3'-Ende (unten). Vergleich der gemessenen Signalintensitäten bei 532 nm.

Diagramm 5 zeigt die gemessenen und gemittelten Signalstärken nach Subtraktion des lokalen Hintergrundes für ein Experiment. Die Fehlerbalken ergeben sich aus der errechneten Standardabweichung für die zwölf Spotreplikate. Das Ergebnis konnte in vierfacher Wiederholung des Versuchs reproduziert werden. Die Wildtyp-Kontrolle zeigt spezifisch das erwartete Signal. Die Sonden 5W1-5W6 (im oberen Teil) zeigen im Vergleich zur Kontrolle eine abnehmende Signalintensität mit ansteigender Zahl der eingeführten Modifikationen. Für 5W7-W13 können keine signifikanten Signale mehr detektiert werden.

Im Vergleich zu „Wobbles“ am 5'-Ende ist der störende Einfluss auf die Verlängerungsreaktion am 3'-Ende erwartungsgemäß deutlich stärker (Diagramm 5, unterer Teil). Schon das Einführen einer Modifikation lässt die Signalstärke im Vergleich zur Kontrollsonde um etwa zwei Drittel abnehmen (3W1). Dieser Effekt setzt sich für die Sequenz 3W2 fort. Die Signale für die Oligonukleotide 3W3-W13 zeigen gegenüber dem Hintergrund keine Signifikanz. Der Versuch wurde viermal wiederholt und konnte reproduziert werden.

6.5.2 Inosin

Gestestet wurde der Einfluss dieser universellen Base auf die enzymatische Verlängerung in Abhängigkeit der Position innerhalb des Oligonukleotids. Analog der Verwendung von „Wobbles“ hat der Austausch eines oder mehrerer der ersten vier Nukleotide am 3'-Ende durch Inosin jeweils einen gravierenden Effekt auf die Signalstärke. Bei diesen Oligonukleotiden konnte im Vergleich mit dem entsprechenden unmodifizierten Primer mit ansonsten gleicher Sequenz kaum oder keine Verlängerung mehr gemessen werden. Ersetzungen in der Mitte der Sonden oder am 5'-Ende haben hingegen keinen negativen Einfluss auf die Primer Extension. Im Sondendesign für den späteren Genotypisierungs-Chip wurden vermehrt solche Positionen durch Inosin ersetzt, die drei oder alle vier variable Nukleotide repräsentierten ($B=C+G+T$, $D=A+G+T$, $H=A+C+T$, $V=A+C+G$ und $N=A+C+G+T$).

6.6 Sondendesign

Die Einführung von *Wobble*-Modifikationen oder Inosin (6.5) erlaubt die Verringerung der Sondenanzahl. Es wurde gezeigt, dass drei bis vier möglichen Positionen innerhalb jedes Oligonukleotids so ersetzt werden können. Dadurch reduziert sich die unter 6.2 errechneten Zahlen um den Faktor 2^3 bis 2^4 . Sondenlängen von 30 Basen und entsprechender T_m zeigten bei der Etablierung optimale Ergebnisse. Im Einzelfall, wenn die Sequenz eine sehr hohe Variabilität aufweist, wurde diese Länge auf bis zu 20 Basen verkürzt. Dadurch können für die Verwirklichung eines HIV-Microarrays realistische Sondenanzahlen erreicht werden. Das

daraus erarbeitete Sondendesign richtete sich im Einzelfall nach Anzahl und Lage der variablen Positionen innerhalb der Primersequenz.

Basierend auf der gewonnenen Sequenzgrundlage wurden die Sonden zum Nachweis der Resistenz-assoziierten Triplets innerhalb von PR und RT abgeleitet. Die Detektion mittels Primer Extension erfordert Oligonukleotide, deren freies 3'-Ende an Position -1 des nachzuweisenden Nukleotids enden und der Verlängerung durch die Polymerase zur Verfügung stehen. Je nach Sequenzvariabilität wurden bestimmte Positionen innerhalb der Sonden, möglichst weit am 5'-Ende gelegen, durch bis zu vier *Wobble* und / oder höchstens zwei Inosin pro Primer ersetzt. Generell wurden entsprechend dem IUB-Code die Basenkombinationen K, S, W, M, Y und R durch „Wobble“ ersetzt, während V, B, H, D und N durch Inosin abgebildet wurden. Dabei betrug der Mindestabstand der Modifikationen zum zu verlängernden 3'-Ende mindestens 5 Basen. Die Länge der Oligonukleotide variierte zwischen 23- und 30meren, alle Sequenzen waren 5'-C6-Amino-modifiziert, um eine optimale Kopplung an die Oberfläche des Microarrays zu gewährleisten. Die Sequenzen sind im Anhang aufgeführt, die Nomenklatur ist im Methodenteil unter 5.1.6 erklärt.

6.7 Protease-Microarray

6.7.1 Auswahl der Patientenproben

Für die Validierung des PR-Chips standen insgesamt 96 PCR-Produkte zur Verfügung, von denen 93 aus Patientenmaterial generiert wurden. Dafür wurden aus Plasma-Proben die virale, genomische RNA isoliert und mittels reverser Transkription in DNA umgeschrieben. Im Folgenden wurden die codierenden Regionen der HIV-Protease und -Reversen Transkriptase spezifisch amplifiziert. Für diese PCR-Produkte lagen bereits Sequenzdaten vor, die durch Sequenzierung nach Sanger ermittelt worden waren. Im Folgenden werden diese als *Referenzsequenzen* bezeichnet, mit denen die Microarraydaten verglichen werden. Die übrigen drei PCR-Produkte wurden durch Amplifikation der entsprechenden Sequenzabschnitte aus dem Plasmid pNL4-3 erstellt.

Die Patientenproben wurden willkürlich ausgewählt, sie erfüllen keine besonderen Bedingungen wie beispielsweise die Vorbehandlung mit bestimmten Medikamenten oder die Zugehörigkeit zu einer ethnischen Gruppe oder einem Virus-Subtyp.

6.7.2 Vorbereitung der Proben

Alle PCR-Produkte wurden zunächst wie im Methodenteil beschrieben unter Einbau von dUTP und der Verwendung Protease-spezifischer Primer (#120/121) reamplifiziert. Die Kontrolle der Fragmentlängen, sowie der Negativkontrolle, die keine Matrizen-DNA enthielt, erfolgte über die Auftrennung im Agarosegel. Anschließend wurden die Amplifikate aus dem Reaktionsansatz isoliert und enzymatisch fragmentiert, um eine sterische Hinderung bei der Hybridisierung zu vermindern.

Das dazu verwendete Enzym Uracil N-Glykosylase hydrolysiert die N-glykosidische Bindung zwischen der Desoxyribose und Uracil. Die Länge und Anzahl der dabei entstehenden Fragmente ist zufällig und hängt vom Einbau des dUTPs während der PCR ab.

Außerdem wurden durch zusätzliche Inkubation mit einer alkalischen Phosphatase freie dNTPs dehydrolysiert. Möglicherweise noch vorhandene dNTPs würden in Konkurrenz mit den farbmarkierten ddNTPs der APEX-Reaktion treten und der Einbau sich zugunsten der nicht modifizierten dNTPs verschieben.

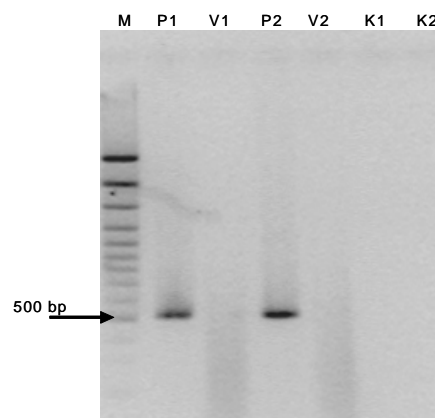


Abbildung 11: Elektrophoretische Auftrennung von Protease-spezifischen Amplifikaten in 2%-Agarose (enzymatisch verdaut und unverdaut)

Spur P1, P2: unverdaute, PR-spezifische PCR-Produkte; erwartete Fragmentlänge: 559 bp
Spur V1, V2: PCR-Produkte nach Inkubation mit Uracil N-Glykosylase; Fragmentlängen nicht definiert
Spur K1, K2: Negativkontrollen der PCR-Reaktion
M = 100 bp DNA-Elektrophoresemarker

Um einen Verlust der DNA während der Reinigung aus dem PCR-Ansatz auszuschließen und zur Kontrolle der enzymatischen Fragmentierung, wurden alle Proben als Doppelansatz in einem 2%-Agarosegel aufgetrennt. Abbildung 11 zeigt exemplarisch den Vergleich von jeweils verdauter (V1+2) und unverdauter DNA (P1+2) zweier Proben. Vor dem Verdau zeigen sich im Vergleich mit dem Marker Banden der erwarteten Größe. Danach lässt sich aufgrund des Schneidens in zufällige Fragmente nur noch ein durchgängiger, so genannter DNA-Schmier erkennen, die PCR-Produkte wurden also vollständig verdaut. K1+2 zeigen die Negativkontrollen der PCR-Reaktion ohne den Einsatz von Matrizen-DNA.

6.7.3 Evaluierung des Microarrays

Anhand dieser so gewonnenen 94 Proben und Sequenzen sollte der HIV-PR-Chip vergleichend mit der Standard-Sequenzierung bewertet werden. Die Durchführung der APEX-Experimente entsprach der im Methodenteil beschriebenen Versuchsanordnung. Nach der Reaktion konnten die Microarrays mittels Laser-Scanner bei vier verschiedenen Wellenlängen ausgelesen werden. Die Signalintensität korreliert mit dem gezeigten Farbspektrum von blau über grün und rot bis hin zur Sättigung im weißen Bereich.

Für jeden Microarray werden entsprechend der vier Wellenlängen (635 nm, 532 nm, 488 nm, 594 nm) vier Bilder ermittelt. Abbildung 12 zeigt eine typische Messung.

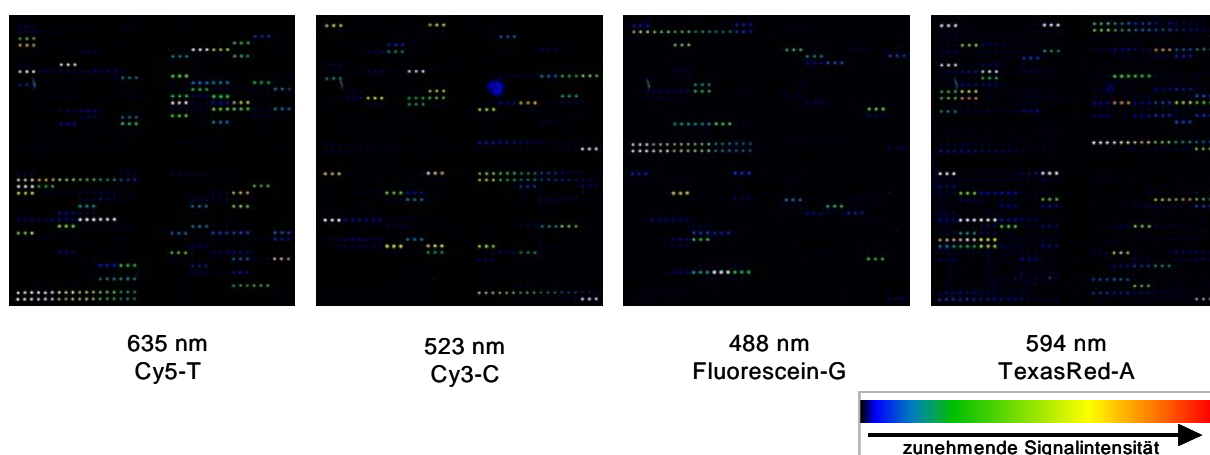


Abbildung 12.: Bei verschiedenen Wellenlängen ermittelte Bilder einer Vierfarbreaktion

Die Signale basieren auf der enzymatischen Verlängerung der immobilisierten Sonden durch fluoreszenzmarkierte ddNTPs bei den Wellenlängen 488, 532, 594 und 635 nm. Die Sonden sind in jeweils drei Replikaten aufgebracht.

Für die Signalquantifizierung wurde die Software *genepix 4.0* herangezogen. In der Darstellung wird jedem Fluoreszenzfarbstoff eine Farbe zugeordnet: Cy5-Signale werden rot, Cy3 in Grün abgebildet. Fluorescein entspricht blau und TexasRed-Signale werden in gelb dargestellt (Abbildung 13). Jeder Position auf dem Array wird ein spezifischer Name zugeordnet, neben anderen Werten die Intensität der Signale ermittelt und diese in einer Datei ausgegeben. Für jede Primer Extension-Reaktion lag also nach Versuchsdurchführung eine solche Ausgabedatei, das so genannte *genepix result file (gpr file)* vor, insgesamt also 94 Dateien. Die nachfolgende, ausführliche Datenanalyse ist unter 6.10 beschrieben. Für jedes PCR-Produkt wurde eine APEX-Reaktion durchgeführt.

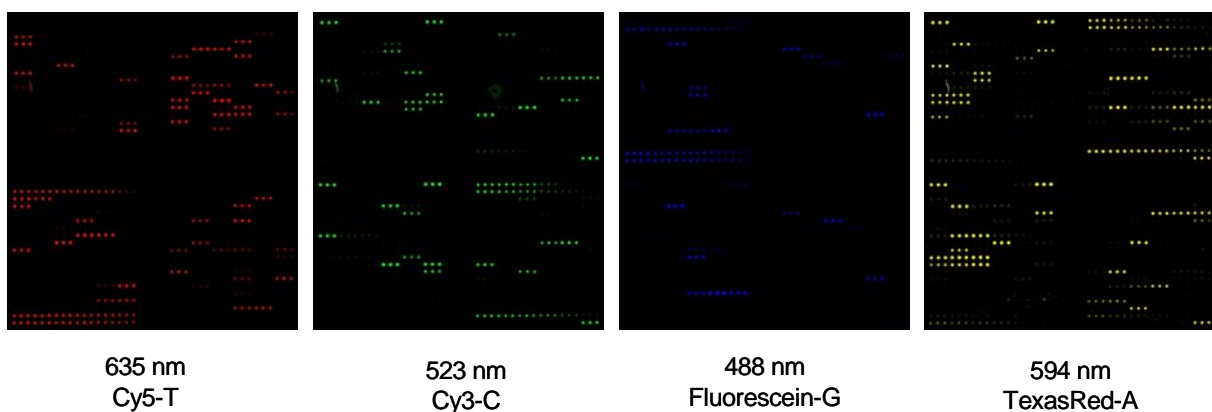


Abbildung 13: Darstellung des Ergebnis durch die Auswertungssoftware

Jedem Farbstoff wird durch die Software eine Farbe zugeordnet: Cy5-rot, Cy3-grün, Fluorescein-blau und exasRed-gelb. Die Signalintensität korreliert jetzt mit der Helligkeit der Spots. Die Sonden sind in jeweils drei Replikaten aufgebracht. Darstellung exemplarisch für ein Experiment

Auswertung der Ergebnisse

Die Sonden des Arrays fragen innerhalb der Protease-Region 26 Codons, das heißt insgesamt 78 Nukleotidpositionen ab. In Abhängigkeit mit der Übereinstimmung der jeweiligen Referenzsequenz wurden die Ergebnisse der APEX-Reaktionen an jeder dieser Positionen als falsch oder richtig definiert. Das Ergebnis wurde als Fehlerquote, also Abweichung von der Referenz, oder als Übereinstimmung in prozentualen Werten angegeben.

Ist die Basenzuordnung nicht eindeutig, wird im Folgenden auch von *Mischungen* oder partieller Übereinstimmung gesprochen. Das ist dann der Fall, wenn sich die Gesamtintensität auf zwei Signale verteilt, wobei das kleinere Signal mindestens einen Anteil von >30 % am Gesamtsignal haben muss. Bei Auftreten von Mischungen ist die Interpretation etwas schwieriger, da die Software nicht in der Lage ist, eine solche Situation automatisch zu erkennen. An diesen Stellen wurde manuell ausgewertet.

Generell können dann drei Sachverhalte vorliegen: erstens zeigen sowohl Microarray- wie auch Sequenzierungs-Daten nach Sanger die gleiche Mischung auf. Standardmäßig werden diese dann nach IUB-Code bezeichnet. Liegt auch dann eine Übereinstimmung vor, gilt das Ergebnis als richtig. Die Ursache kann im Vorhandensein verschiedener Virusvarianten in einem Patienten liegen, die sich an der entsprechenden Stelle in der Sequenz unterscheiden. Zweitens kann die Referenz ein eindeutiges Ergebnis liefern, während die Microarray-Daten eine Mischung zeigen (beispielsweise A vs. A/G). Oder aber drittens, die umgekehrte Situation, in der für die APEX-Reaktion eine eindeutige Basenzuordnung zu treffen ist, für die Sequenzierung hingegen nicht. Aus der Beurteilung der beiden letzten Situationen, je nachdem ob sie als falsch oder richtig definiert werden, ergeben sich Unterschiede für das Gesamtergebnis.

Verteilt sich die gemessene Gesamtintensität der Microarray-Messung eines Spots auf drei oder vier Teilsignale, kann keine Basenzuordnung vorgenommen werden. Die Sequenz an der entsprechenden Nukleotidposition wird dann als fehlerhaft gewertet. Gleiches gilt, wenn alle Sonden, die ein bestimmtes Nukleotid abfragen, aufgrund schwacher Signalintensitäten durch die Software ausgefiltert werden. Stimmen APEX-Daten und Referenz eindeutig nicht miteinander überein, wird ebenfalls das Ergebnis der Primer Extension-Reaktion als falsch definiert.

Nach dieser Vorgehensweise wurde jede abgefragte Position untersucht und eine Fehlerrate bzw. prozentuale Übereinstimmung für jede Probe bezüglich der Referenzsequenz ermittelt. Im Anschluss wurde wiederum die Fehlerrate aller analysierten Proben an der Gesamtanzahl, also 94, errechnet. Aus Diagramm 6 gehen die Ergebnisse der Auswertung von 94 analysierten PCR-Produkten hervor. Auf der x-Achse sind die Nukleotid- bzw. Codonpositionen dargestellt, die von den Sonden auf dem Array abgefragt werden. Die y-Achse zeigt die prozentuale Übereinstimmung mit der Referenzsequenz. Blaue Balken stellen die Werte dar, die die oben vorher beschriebenen Mischungen tolerieren. Bei strengerer Beurteilung gelten solche Mischungen als Fehler, die Übereinstimmung ist folglich geringer (rote Balken).

In verschiedenen Bereichen der Protease lassen sich unterschiedlich gute Ergebnisse mit der APEX-Analyse erzielen. Nicht jede Position lässt sich gleich gut detektieren, so gibt es etwa für die Nukleotide 28, 30, 108, 184, 264 und 277 prozentuale Abweichungen (~20 %), die deutlich höher sind, als für die übrigen Bereiche. Gleiches gilt für die Positionen, die die Aminosäuren 63 und 71 codieren. In dieser Region weisen die Array-basierten Daten mehr Fehler und somit eine geringere Übereinstimmung mit den Sequenzierungsdaten auf.

Umgekehrt zeigen andere Teile sehr hohe Konkordanz, z.B. die Nukleotidstellen, die Sequenzgrundlage der Aminosäuren 24-32 und 46-53 sind.

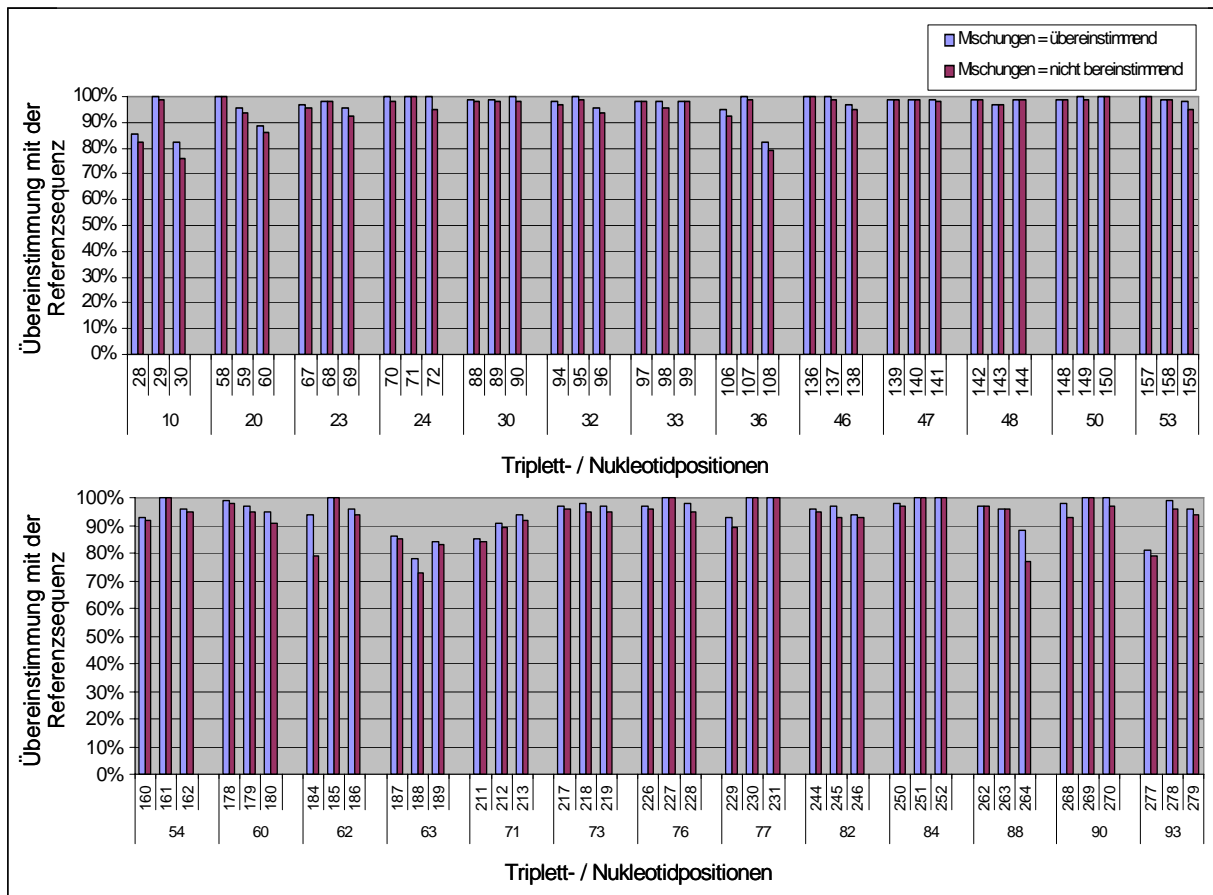


Diagramm 6: Vergleich APEX / Sequenzierung für den Protease-Chip

Dargestellt wird die prozentuale Übereinstimmung der APEX- mit den Sequenzierungs-Daten nach Sanger für 94 untersuchte PCR-Produkte aus Patientenmaterial.

Zur besseren Übersicht wurden dann die drei Werte, die die Nukleotidpositionen eines Codons bilden, für jedes Experiment zusammengefasst und in Diagramm 7 dargestellt.

Der Wert für einen Datenpunkt errechnet sich aus der Summe der falschen Messungen dividiert durch die Gesamtanzahl der Messungen (also $3 \times 94 = 282$). Die Fehlerbalken geben die Standardabweichung an, die sich aus den drei Messwerten für die drei Nukleotidpositionen ergeben. Sind sie groß, so bedeutet das, dass eine große Diskrepanz im Nachweis der drei Nukleotide vorliegt. Das gilt beispielsweise für Aminosäure 10, codiert durch die Positionen 28, 29 und 30, im Vergleich mit Diagramm 6 lässt sich erkennen, dass die Fehlerrate für die Positionen 28 und 30 sehr hoch liegt, 29 hingegen in fast allen Fällen korrekt detektiert werden kann. Umgekehrt hat Codon 30 keinen Fehlerbalken, da sich alle drei Nukleotide gleich effizient nachweisen lassen.

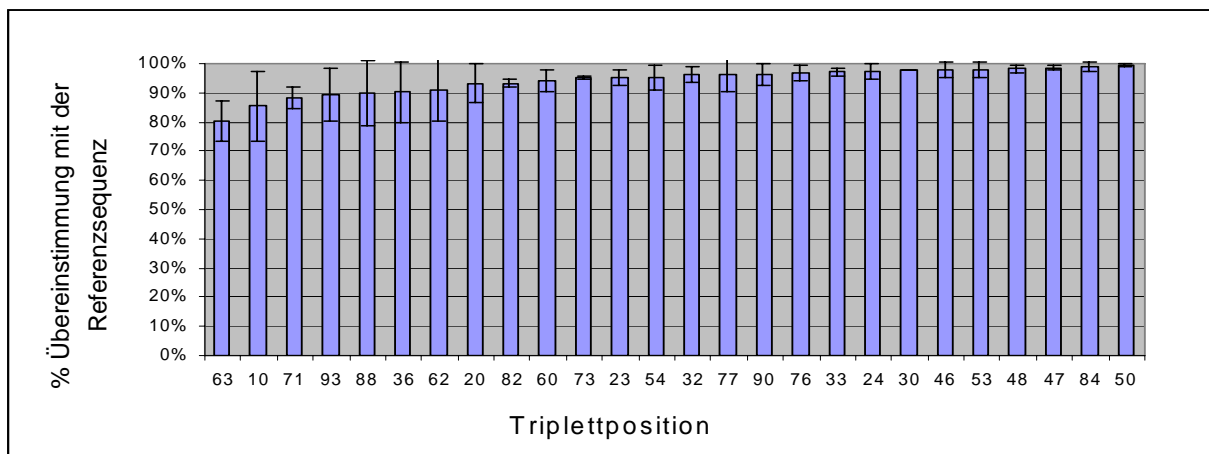


Diagramm 7: Übereinstimmung / Fehlerrate auf Triplet- / AS-Ebene

Gezeigt sind die prozentualen Übereinstimmungen der untersuchten Triplets (Erklärung siehe Text) unter Berücksichtigung aller analysierten PCR-Produkte. Auf der x-Achse sind die abgefragten Codons in aufsteigender Nachweiseffizienz angeordnet. Die prozentuale Übereinstimmung liegt zwischen 80% für das Triplet, das AS 10 codiert und 99% für AS 60.

Die Triplet-Positionen sind mit steigender Nachweiseffizienz auf der x-Achse angeordnet. Die höchste Fehlerrate zeigen die Triplets, die Aminosäure 63 codieren, gefolgt von 10, 71, 93 und 88. Die Triplet-Positionen 48, 47, 84 und 50 hingegen werden mit wenigen Ausnahmen immer richtig detektiert. Die Fehlerrate über alle analysierten PCR-Produkte und alle abgefragten Positionen (insgesamt $78 \times 94 = 7332$) liegt bei 8%. Es kann also eine Übereinstimmung der APEX-Resultate mit den Sequenzierungsdaten (nach Sanger) von 92% erreicht werden. Interpretiert man die partiell übereinstimmenden Mischungen ebenfalls als richtiges Ergebnis, erhöht sich die Übereinstimmung auf 94%. Von den 7332 untersuchten Positionen wurden nach Filtern der Daten 7167 zur weiteren Analyse herangezogen, unter den gewählten Bedingungen konnten also für 97,5% signifikante Signale gemessen werden. Der Anteil an falsch-negativen Resultaten liegt somit bei 2,5%. Betrachtet man nun ausschließlich die Positionen, für die eine Sequenzzuweisung stattgefunden hat, liegt die Konkordanz zwischen 94,2-96,2% (abhängig von der Toleranz gegenüber Mischungen). Falsch-positive Ergebnisse sind also zu 3,8-5,8% für die Gesamtfehlerrate verantwortlich.

6.7.4 Optimierung des Chiplayouts

Um zu untersuchen, ob sich durch Ergänzen weiterer Sonden die Rate der richtig bestimmten Positionen erhöhen lässt, wurden nach Auswertung der 94 Datensätze Veränderungen am

Sondenbestand des Arrays vorgenommen. Für sieben Nukleotidpositionen wurden Sonden ergänzt, die die Sequenz des Gegenstranges abdeckten. Aufgrund zu schwacher Signalintensitäten wurden die vorhandenen Primer immer oder häufig durch die Auswertungssoftware ausgefiltert. Außerdem wurden insgesamt acht Oligonukleotide aus dem Chip-Layout entfernt, weil sie offensichtlich unabhängig von der Ziel-DNA immer oder sehr häufig falsche Ergebnisse lieferten. Dies kann möglicherweise auf Kreuzhybridisierungen mit anderen Abschnitten innerhalb des Amplifikates oder durch Bildung von Sekundärstrukturen der Ziel-DNA zurückzuführen sein. Die ergänzten oder entfernten Sonden sind im Anhang entsprechend gekennzeichnet.

6.7.5 Reproduzierbarkeit der Ergebnisse

Um den Einfluss des optimierten SONDENSATZES im Vergleich zum ursprünglichen zu zeigen, wurden vier Proben mit einer unterdurchschnittlichen Nachweiseffizienz von weniger als 92% Übereinstimmung mit der Referensequenz ausgewählt. Da für die Validierung des PR-Microarrays jedes der 94 Amplifikate nur einmal analysiert wurde, soll zudem eine Aussage über die Reproduzierbarkeit der Ergebnisse getroffen werden. Mit den ausgewählten Proben wurde die Primer Extension-Reaktion mehrfach in verschiedenen Ansätzen, zu unterschiedlichen Zeitpunkten und sowohl mit der ursprünglichen (PE_5-8) als auch mit der veränderten Sondenzusammenstellung durchgeführt (PE_1-4).

Diagramm 8 zeigt exemplarisch die Ergebnisse für eine der vier mehrfach untersuchten Proben. Aufgetragen sind die Anzahl der Fehler gegen die Nukleotidpositionen, an denen sie auftreten. Bei achtfacher Wiederholung kann ein Fehler also höchstens achtmal auftreten, je höher der Balken, desto reproduzierbarer die falsch bestimmte Position. Umgekehrt ist für alle Stellen, an denen kein Balken erscheint, die Sequenz in jeder Analyse richtig nachgewiesen worden. An den Nukleotidpositionen 60, 89, 187 und 263 tritt in einem von acht Untersuchungen ein Fehler auf, Position 108 ist dreimal (von achtmal) falsch bestimmt. Die höchste Fehlerrate ist für Triplet 62 und 63 zu beobachten. Für Nukleotid 184 gibt die APEX-Reaktion in fünf, für 188 in sieben von acht Fällen ein falsches Ergebnis. Position 189 wurde sogar immer falsch detektiert. Dies bestätigt die Beobachtung, dass Triplet 63 der höchsten Fehlerrate unterliegt (Diagramm 6 und Diagramm 7).

Außerdem wurde in dieser Versuchsanordnung die ursprüngliche Sondenzusammenstellung (blaue Balken) mit dem neuen Chip-Layout (gelbe Balken) verglichen. Im Ganzen ist eine

Verringerung der Fehler zu beobachten; wurden mit dem ersten Sondensatz bei achtfacher Versuchswiederholung insgesamt 19 Positionen falsch bestimmt (blau), sind es mit dem zweiten nur noch acht (gelb). Diese wiederum treten nur in Bereich der Triplets 62 und 63 auf. Die rote Markierung zeigt die Triplett-Positionen, an denen die Oligonukleotide für den zweiten DNA-Strang ergänzt wurden (30, 36 und 62). Es lässt sich deutlich die Übereinstimmung mit der Verringerung der Fehleranzahl erkennen. Für Position 63 waren bereits im Vorfeld Sonden eingesetzt worden, die beide Orientierungen abdeckten, eine Ergänzung war hier nicht mehr möglich.

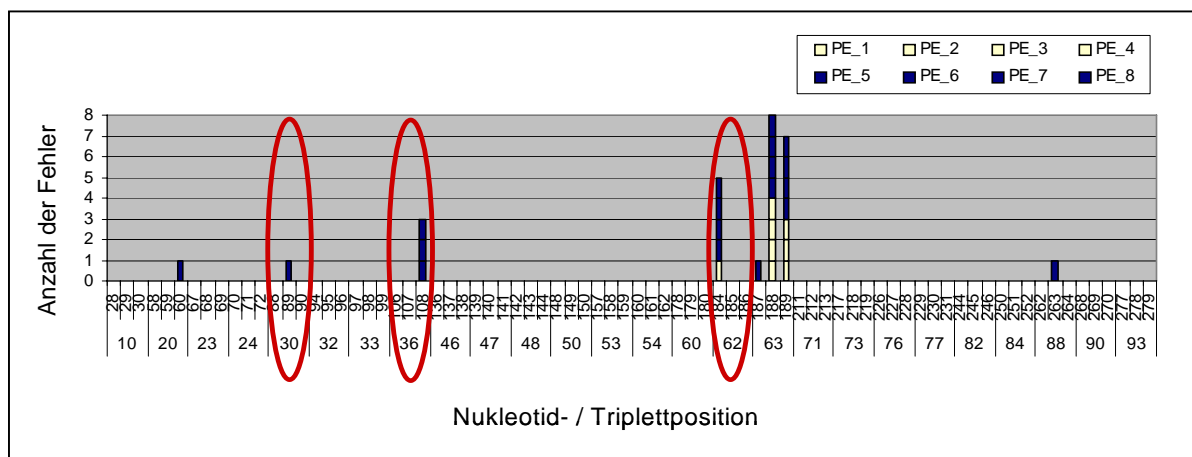


Diagramm 8: Reproduzierbarkeit der Ergebnisse und Vergleich der Chip-Layouts

Aus dem Diagramm gehen die Nukleotidpositionen hervor, an denen bei achtfacher Durchführung einer APEX-Reaktion (PE_1-9) wiederholt fehlerhafte Ergebnisse auftreten.

Gelb: verändertes Chip-Layout durch Entfernen / Ergänzen von Sonden (PE_1-4)

Blau: Verwendung des ursprünglichen Chip-Layouts (PE_5-8)

Rot markiert: Positionen, an denen Oligos im neuen Layout ergänzt wurden

Für die vier untersuchten PCR-Produkte, die aus den 96 Proben aufgrund ihrer unterdurchschnittlichen Nachweiseffizienz (<92%) ausgewählt wurden, konnte in jedem Fall eine Verbesserung des Ergebnisses durch die Verwendung der neuen Sondenzusammenstellung erreicht werden. Tabelle 17 zeigt die prozentualen Übereinstimmungen mit der jeweiligen Referenz im Vergleich, die Werte konnten zwischen 4% (Probe 1) und 8% (Probe 3) verbessert werden (Diagramme nicht gezeigt). Die durchschnittliche Erhöhung der Konkordanz liegt bei den vier untersuchten Proben also bei 5,5%.

Tabelle 17: Verbesserung der Nachweiseffizienz durch einen veränderten Sondensatz

Probe	1. Sondensatz	2. Sondensatz
1	86%	90%
2	88%	93%
3	86%	94%
4	92%	97%

6.7.6 Detektion von Minoritäten

Für die klinische Diagnostik ist es von Bedeutung, das Vorhandensein verschiedener Sequenzen in einer Mischung nachweisen zu können. Diese Situation tritt dann auf, wenn eine Probe unterschiedliche Viruspopulationen enthält. Um den Primer Extension-Chip dahingehend zu testen, wurden zu diesem Zweck zwei PCR-Produkte (WT / mut) in definierten Anteilen gemischt. Das zweite Amplifikat unterscheidet sich von der NL4-3-Sequenz nur durch Mutationen im Triplet 10 (Nukleotid 28 C => A) und Triplet 90 (Nukleotid 268 T => A). Die beiden PCR-Produkte wurden in den Verhältnissen 100/0, 75/25, 50/50, 25/75 und 0/100 in einer Gesamt-DNA-Menge von je 400µg / Ansatz eingesetzt. Die Primer Extension-Reaktion und das Auslesen der Microarrays wurden wie im Methodenteil beschrieben durchgeführt.

Diagramm 9 zeigt die Anteile der jeweils zwei Signale (WT / mut) am Gesamtsignal für die verschiedenen Mischungsverhältnisse für Position 28 (oberer Teil) und 268 (unterer Teil). Anteile <2 am Gesamtsignal konnten, bedingt durch die Software nicht angezeigt werden (siehe Verhältnis 100/0 und 0/100). Die Diagramme zeigen den Mittelwert der Spottingduplikate eines Chips, die Ergebnisse konnten in einem zweiten Experiment reproduziert werden.

Wird nur ein PCR-Produkt eingesetzt, ist eine eindeutige Detektion in jedem Fall möglich. Es wird deutlich, dass in Mischungen auch die weniger stark repräsentierte Sequenzvariante in einem Anteil von 25% sowohl für Position 28 als auch für 268 detektiert werden kann. Annähernd gleiche Signalanteile zeigt die 50/50-Mischung für Position 28, während für Position 268 der Einbau des TexasRed-ddA (mut) über den des Cy5-ddTs überwiegt. Bei Verhältnissen von ein zu zwei Drittel ist in allen Fällen deutlich, welche Variante stärker vertreten ist. Es wurde gezeigt, dass sich Mischungen als solche detektieren lassen. Eine Quantifizierung der eingesetzten Verhältnisse ist hingegen stark sondenabhängig und nicht generell möglich.

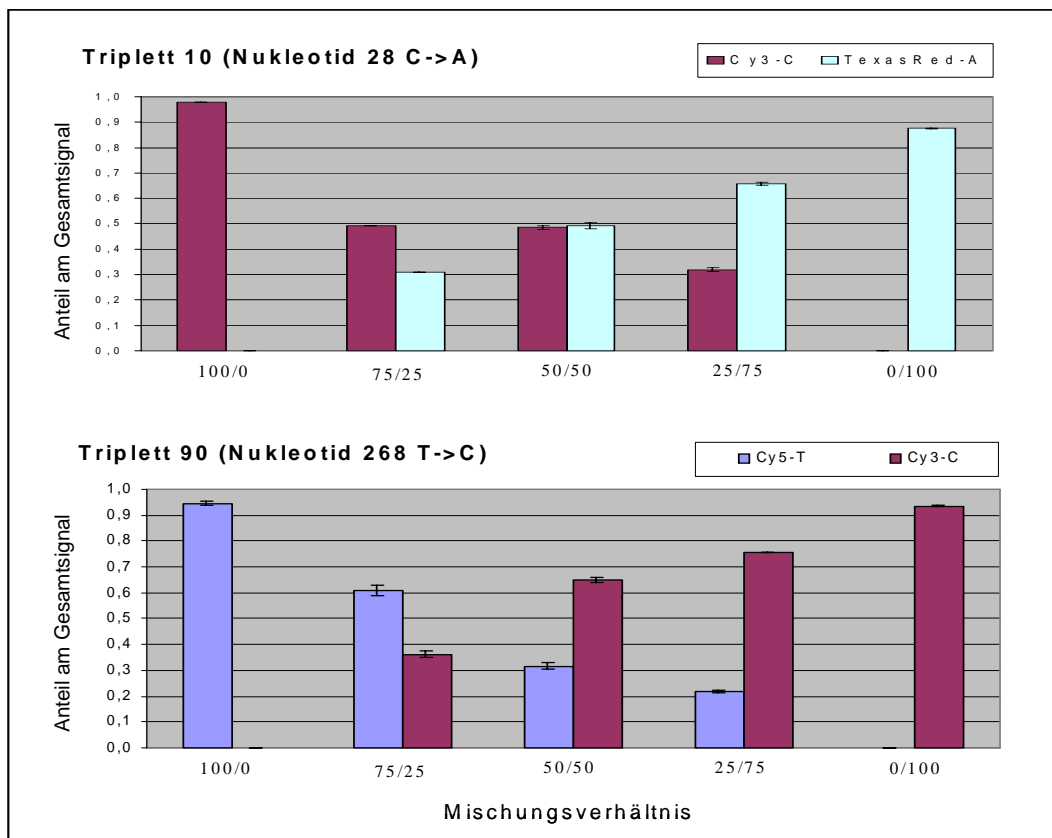


Diagramm 9: Nachweis unterschiedlicher Sequenzvarianten in Mischungen (WT vs Mutation)
 Aufgetragen sind die Anteile der Einzelsignale am Gesamtsignal gegen die Mischungsverhältnisse der PCR-Produkte.

6.7.7 Untersuchung von Proben mit komplexen Mutationsmustern

2003/2004 wurde u.a. vom Nationalen Referenzzentrum für Retroviren in Erlangen ein Ringversuch durchgeführt. Dabei wurden insgesamt sechs definierte Proben (zwei davon mit Mischungen) zur Resistenztestung an 23 verschiedene Labore in Deutschland und Luxemburg verschickt. Aus den dabei gewonnenen Daten geht unter anderem hervor, dass die Genotypisierung dann häufig ein falsches Ergebnis bestimmt, wenn die Proben komplexe Mutationsmuster aufweisen.

Auf diesen Gesichtspunkt hin wurden die bei der Validierung des APEX-Assays erhaltenen Ergebnisse im Folgenden untersucht. Es sollte ermittelt werden, ob auch hier Proben, die eine größere Abweichung zur Sequenz des Wildtyps zeigen, ein schlechteres Genotypisierungsergebnis (geringere Übereinstimmung APEX vs Sequenzierung) erzielen.

Dazu wurde zunächst die prozentuale Übereinstimmung zwischen der jeweiligen Referenzsequenz und dem Wildtyp mittels Alignment bestimmt (Huang & Miller, 1991). Diese Übereinstimmung ist bezogen auf die gesamte Protease-Sequenz, nicht nur auf Resistenz-relevante Positionen. Die Werte wurden für alle Proben in Diagramm 10 in aufsteigender Reihenfolge auf der x-Achse dargestellt. In Abhängigkeit davon auf der y-Achse ist die jeweils zugehörige Konkordanz zwischen APEX und Sequenzierung aufgetragen. Die aufsteigende schwarze Linie stellt den linearen Trend der korrelierten Daten dar. Für die analysierten Proben lässt sich sagen, dass die Nachweiseffizienz für Proben mit weniger Mutationen höher ist, als für Proben, deren Sequenz stark (bis zu 17%) vom WT abweicht. Das korreliert mit den Daten, die aus dem Ringversuch resultierten.

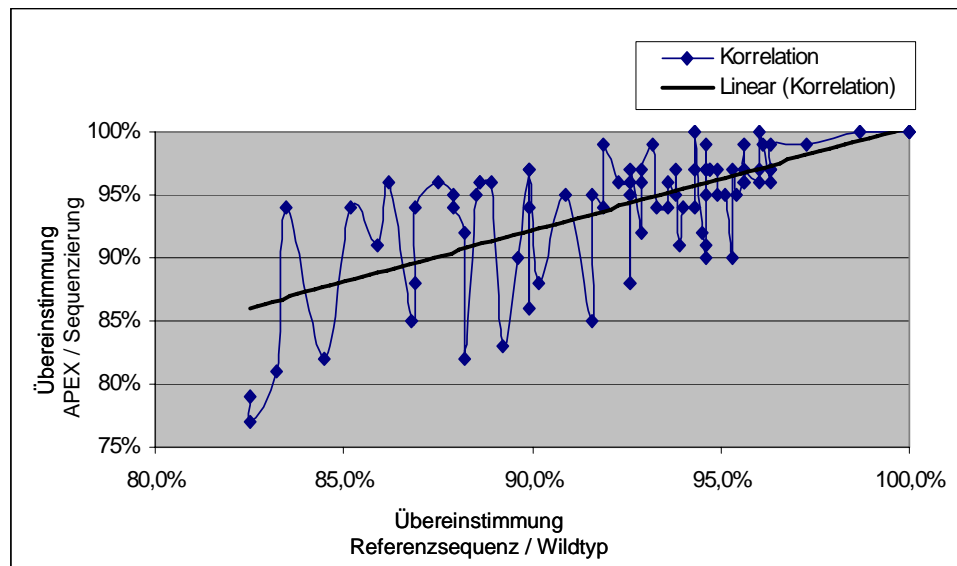


Diagramm 10: Korrelation von komplexen Mutationsmustern und Detektionseffizienz

Je weniger Mutationen eine Probe aufweist, desto besser ist das Ergebnis für die APEX-Untersuchung.
Erklärung siehe Text.

6.8 Reverse Transkriptase-Microarray

6.8.1 Auswahl der Patientenproben

Für die Validierung des RT-Microarrays wurde eine Teilmenge der Patientenproben verwendet, die schon für die Analyse des PR-Chips herangezogen wurden. Es handelte sich um insgesamt 48 PCR-Produkte, die für die Untersuchung mittels RT-Microarray eingesetzt

wurden. Die Bezeichnungen sind im Anhang angegeben, ebenso wie die Namen und Sequenzen der verwendeten Oligonukleotide. Die Vorbereitung der Proben, das Sondendesign und die Herstellung der Arrays erfolgte in gleicher Art wie für den PR-Microarray (6.7.2).

6.8.2 Reaktionsbedingungen

Bei der Durchführung der Primer Extension-Reaktion allerdings führte das bisher genutzte Temperaturprofil (30 sec, 95°C / 10 min, 40°C / 15 min 65°C) nur zu einer unzureichenden Signalbildung. Es wurden nur wenige Oligonukleotide auf dem Array verlängert und die erhaltenen Signalintensitäten waren derart schwach, dass es nicht möglich war, die Ergebnisse mit der etablierten Software auszuwerten. Aus diesem Grund wurde die Temperatur für den Verlängerungsschritt gesenkt, da die verwendeten Sonden aufgrund der höheren Variabilität im Durchschnitt kürzer waren und somit eine tendenziell eher niedrigere T_m vorlag.

Zunächst wurde die Temperatur auf 63°C, 61°C und 59°C gesenkt, Denaturierung und Hybridisierung blieben unverändert, sowie die übrigen Parameter, wie Konzentration der Reaktionskomponenten oder der Versuchsablauf. Doch auch unter diesen geänderten Bedingungen blieben ausreichende Signalintensitäten aus (Daten nicht gezeigt). Erst durch die Einführung eines mehrstufigen Temperaturverlaufs konnten Ergebnisse erzielt werden, die mit denen des PR-Microarrays vergleichbar waren.

Tabelle 18: Mehrstufiges Temperaturprofil zur Sonderverlängerung auf dem RT-Microarray

Schritt	Temperatur	Dauer
Denaturierung	95°C	30 sec
Hybridisierung	40°C	10 min
Verlängerung	68°C	15 min
Hybridisierung	40°C	10 min
Verlängerung	66°C	15 min
Hybridisierung	40°C	10 min
Verlängerung	64°C	15 min
Hybridisierung	40°C	10 min
Verlängerung	62°C	15 min
Hybridisierung	40°C	10 min
Verlängerung	60°C	15 min
Hybridisierung	40°C	10 min
Verlängerung	58°C	15 min

Die Heizrate des *in situ*-Thermocyclers betrug 1°C / Sekunde.

Tabelle 18 zeigt den Verlauf der verschiedenen Temperaturstufen. Zunächst erfolgt eine Denaturierung der Ziel-DNA für 30 Sekunden bei 95°C. Es folgt eine wiederkehrende Wiederholung von Hybridisierung und Verlängerung, wobei die Extensionstemperatur, ausgehend von 68°C pro Zyklus, schrittweise um 2°C bis 58°C verringert wird. Zwischengeschaltet ist jeweils eine Hybridisierung für 10 min bei 40°C. Die Reaktionszeit für die enzymatische Verlängerung beträgt nach wie vor jeweils 15 min. Mit diesem Temperaturprofil wurden die 48 PCR-Produkte zur Validierung des RT-Microarrays analysiert.

Abbildung 14 zeigt, dass die Ergebnisse mit denen des PR-Microarrays vergleichbar sind. Gezeigt ist die Darstellung eines bei vier Wellenlängen ausgelesenen Microarrays durch die Software *genepix 4.0*, die zur Quantifizierung der Signalintensitäten verwendet wurde.

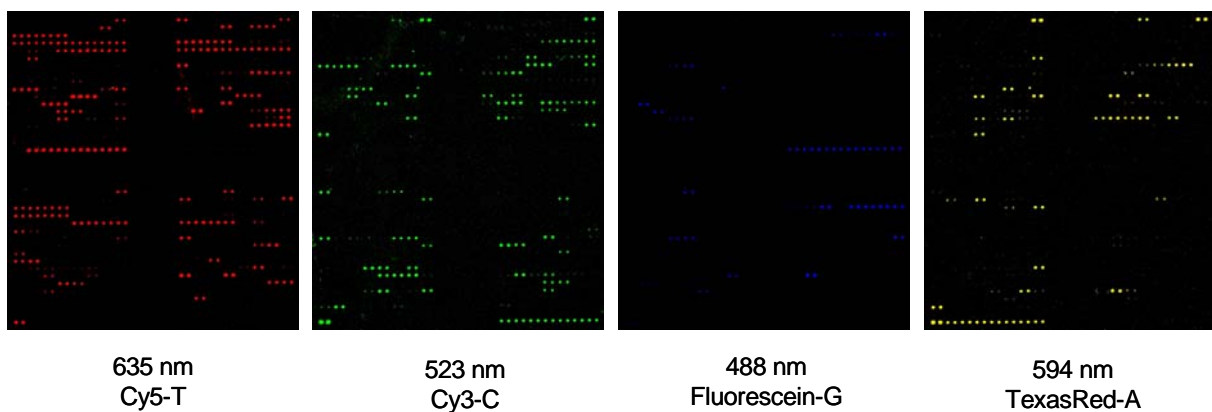


Abbildung 14: Darstellung des Ergebnis durch die Auswertungssoftware

Jedem Farbstoff wird durch die Software eine Farbe zugeordnet: Cy5-rot, Cy3-grün, Fluorescein-blau und TexasRed-gelb. Die Signalintensität korreliert mit der Helligkeit der Spots. Die Sonden sind in jeweils zwei Replikaten aufgebracht. Darstellung exemplarisch für ein Experiment

6.8.3 Evaluierung

Die ausführliche Datenanalyse wurde mit den unter 6.10 beschriebenen Programmen vorgenommen, sie entspricht in Darstellung und Interpretation den Ergebnissen des PR-Chips (6.7.3).

Diagramm 11 zeigt die Übereinstimmung der APEX-Ergebnisse mit den Sanger-Sequenzierungsdaten für 48 untersuchte Patientenproben auf der Ebene der Nukleotidpositionen. Wie beim PR-Microarray zeigt sich, dass die Übereinstimmung stark

Positions-abhängig ist. Auffällig ist vor allem die hohe Fehlerrate an Nukleotidposition 307 im Triplet 103. Besonders viele Mischungen werden detektiert an den Positionen 199 (Triplet 67) und 558 (Triplet 190). Generell zeigt sich im Vergleich zum Protease-Microarray für die Reverse Transkriptase eine geringere Konkordanz mit den Referenzsequenzen.

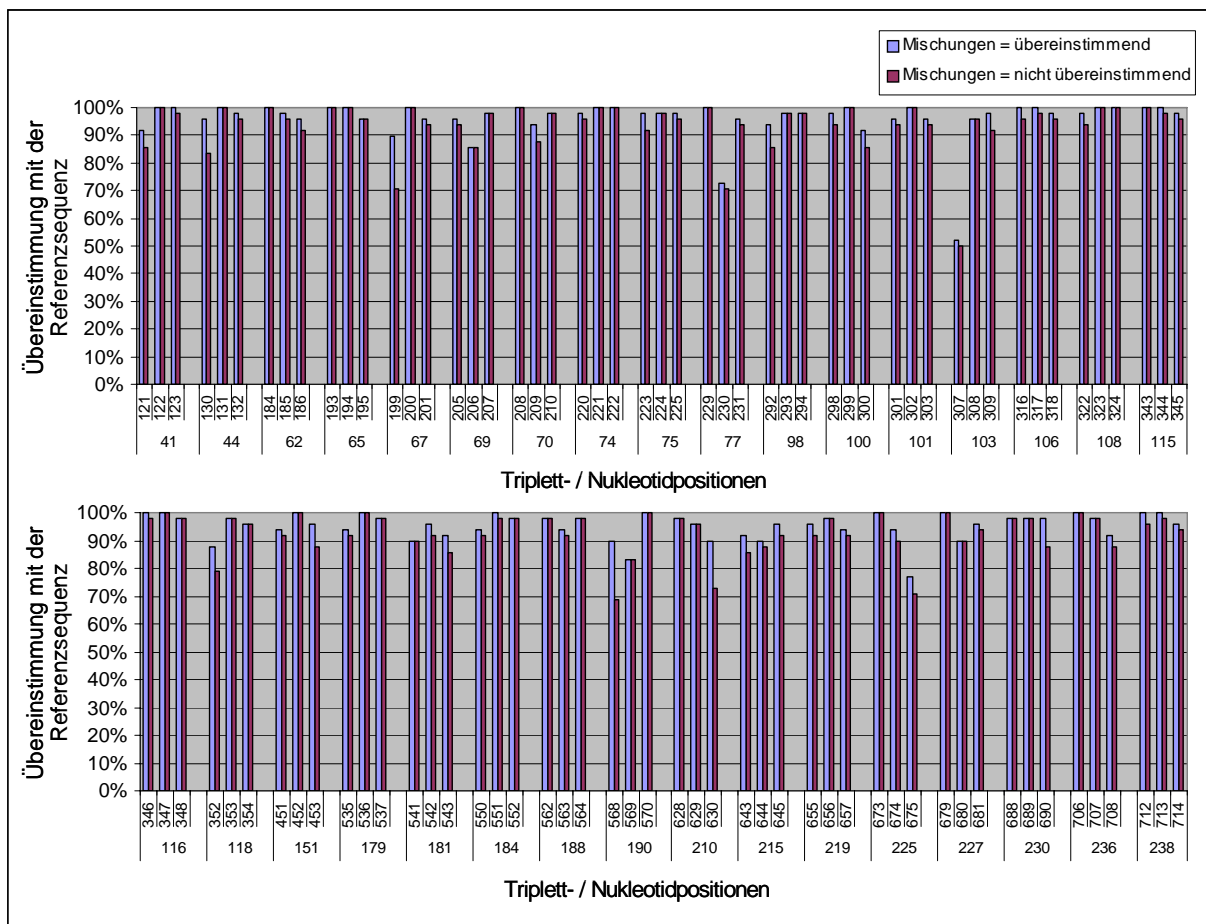


Diagramm 11: Vergleich APEX / Sequenzierung für den Reverse Transkriptase-Chip

Dargestellt wird die prozentuale Übereinstimmung der APEX- mit den Sequenzierungs-Daten nach Sanger für 48 untersuchte PCR-Produkte aus Patientenmaterial.

Für den gesamten Microarray lässt sich über alle untersuchten Positionen und alle analysierten PCR-Produkte (insgesamt $48 \times 99 = 4752$) eine Übereinstimmung mit den Referenzsequenzen (nach Sanger) von 86% erreichen, unter der Voraussetzung, dass *Mischungen* (wie unter 6.7.3 definiert) als falsches Ergebnis gelten. Toleriert man solche Resultate als richtig, erhöht sich die Übereinstimmung auf 90%. Die Fehlerrate der Microarray-Daten liegt also im Vergleich mit herkömmlichen Sequenzierungsdaten je nach Interpretation der Ergebnisse zwischen 10 und 14%.

Ausgewertet wurden 4518 von 4752 Positionen. Der Anteil an ausgefilterten (falsch-negativen) Positionen lag demnach bei 4,9%, also im Vergleich mit dem PR-Chip etwa doppelt so hoch. Damit ergibt sich in Abhängigkeit der Interpretation von partiell übereinstimmenden Sequenzzuweisungen eine Übereinstimmung von 91,2-94,6%. Die falsch-positiven Ergebnisse haben also einen Anteil von 5,4-8,8% an der Gesamtfehlerrate.

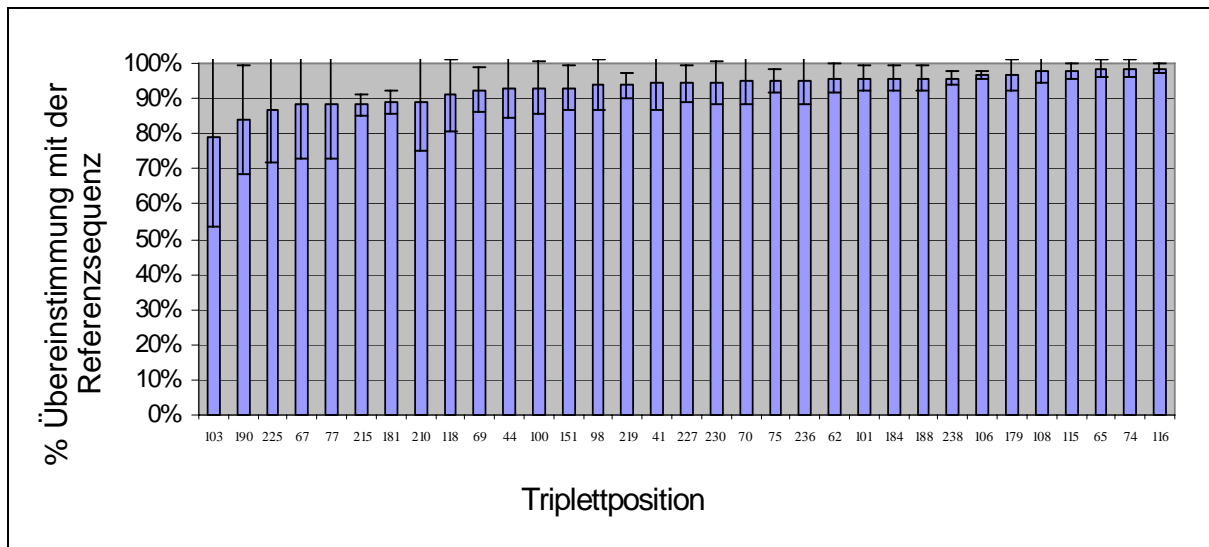


Diagramm 12: Übereinstimmung / Fehlerrate auf Triplet- / AS-Ebene

Gezeigt sind die prozentualen Übereinstimmungen der untersuchten Triplets unter Berücksichtigung aller analysierten PCR-Produkte. Auf der x-Achse sind die abgefragten Codons in aufsteigender Nachweiseffizienz angeordnet. Die prozentuale Übereinstimmung liegt zwischen 79% für das Triplet, das AS 103 codiert und 99% für AS 116.

Diagramm 12 zeigt die mittlere Übereinstimmung mit den Referenzsequenzen für alle Triplettpositionen des RT-Microarray. Sie liegt zwischen 79% für Triplet 103 und 99% für Position 116. Die Fehlerbalken geben die Standardabweichungen zwischen den drei Nukleotidpositionen an. Wie schon vorher erklärt, heißt das für große Fehlerbalken, dass sich die drei Nukleotide unterschiedlich gut detektieren lassen, während bei ähnlicher Nachweiseffizienz die Abweichungen nur gering sind.

6.8.4 Überprüfen der Referenzsequenzen

Um die Referenzsequenzen zu überprüfen, wurden stichprobenartig sechs der 48 PCR-Produkte, die zur Validierung des RT-Microarrays verwendet wurden, in einem

zweiten Labor nachsequenziert (Andreas Hunzicker, DKFZ). Von den Proben mit den Kürzeln HB, FW, ID, IS, JM und LN wurden die Amplifikate mit den Primern #123 Pol-2 und #124 Pol-4 nachträglich sequenziert. Die Ergebnisse wurden mit denen von der Abteilung Virologie (Hygieneinstitut, Universität Heidelberg) zur Verfügung gestellten Sequenzen mittels einer Software verglichen (<http://www.ch.embnet.org/cgi-bin/LALIGN>, Huang & Miller, 1991). Über eine Länge von 702 Basen lagen die Übereinstimmungen der jeweils zusammengehörigen Datensätzen bei 95,7% (FW), 95,9% (ID), 98,1% (LN), 98,3% (IS), 99,1% (JM) und 99,2% (HB). Trotz Anwendung der gleichen technischen Vorgehensweise wurde in keinem Fall eine Konkordanz von 100% erreicht.

6.9 Genotypisierung von Proben verschiedener Subtypen

Nimmt man mit Hilfe des *Rega HIV-1 Subtyping Tool* (Oliveira *et al.*, 2005) eine Bestimmung des Subtyps vor, so ergibt sich für die 94 zur Validierung des Microarrays herangezogenen Patientenproben folgende, in Tabelle 19 angegebene Verteilung.

(<http://dbpartners.stanford.edu/RegaSubtyping>)

Tabelle 19: Verteilung der HIV-Subtypen innerhalb der untersuchten Patientenproben

Subtyp	Anzahl (PR)	Anzahl (RT)
B	71	41
A	3	0
C	3	1
F1	1	1
CRF01_AE	2	2
CRF01_AG	1	0
CRF02_AG	5	1
CRF06_cpx	1	0
nicht bestimmt	7	2
insgesamt	94 Proben	48 Proben

Unter Berücksichtigung der Subtypzugehörigkeiten zeigen sich Unterschiede in den Ergebnissen hinsichtlich der Konkordanz von APEX- und Sequenzierungsdaten nach Sanger.

Protease-Chip

Insgesamt wurden für die Validierung des Protease-Microarrays 94 Patientenproben zufällig ausgewählt. Davon entsprechen 71, also knapp 76% dem HIV-Subtyp B, die restlichen Proben verteilen sich auf andere Subtypen und rekombinante Formen, sieben Proben davon konnten nicht bestimmt werden (Tabelle 19). In den Subtyp B-Proben wurden also 5538 Nukleotidpositionen ($71 \times 78 = 5538$) durch den Chip abgefragt. Davon lieferten 4,9-6,5% ein falsches Ergebnis oder wurden durch die Software ausgefiltert. Die beiden Werte beziehen sich wiederum auf das Zulassen oder Nichtzulassen von Mischungen. Im Vergleich dazu schneiden die 23 nicht-Subtyp B-Proben deutlich schlechter ab. Über alle analysierten Positionen ($23 \times 78 = 1794$) liegt die Fehlerrate je nach Interpretation der Daten hier bei 9,5-12,3%.

Reverse Transkriptase-Chip

Eine ähnliche Situation ergibt sich bei analogem Rechengang für die Resultate des RT-Microarrays. Der Anteil an Subtyp B-Proben liegt hier höher, nämlich bei 87,5% (41 von 48). Die davon betrachteten Positionen ($41 \times 99 = 4059$) weisen eine Übereinstimmung zwischen 87,3-90,5%, d.h. eine Fehlerrate von 9,5-12,7% auf. Bei nur sieben Proben handelt es sich um nicht-Subtyp B-Sequenzen. Für diesen geringen Anteil sind die Ergebnisse der APEX-Experimente aber wiederum deutlich schlechter. 12,4-16,3% der untersuchten Nukleotide ($7 \times 99 = 693$) wurden abweichend bzw. auch partiell abweichend von der jeweiligen Referenzsequenz oder gar nicht detektiert.

6.10 Auswertung der Daten

Für die statistische Auswertung der durchgeführten Vierfarb-Experimente war keine geeignete, kommerziell-erhältliche Software verfügbar. Eine manuelle Auswertung solcher Datenmengen stößt schnell an ihre Grenzen und ist nur eingeschränkt aussagekräftig. Deshalb wurde in Zusammenarbeit mit Aslihan Ay im Rahmen einer Diplomarbeit ein Programm etabliert, das die statistisch signifikante Verarbeitung der generierten Daten ermöglicht (Aslihan Ay, Fachhochschule Bingen, Januar 2006).

Um die APEX-Experimente auszuwerten, wurden folgende Analyseschritte speziell für die generierten Datensätze angepasst.

Einlesen der Daten

Nach der Quantifizierung (genepix 4.0) lag für jede Primer Extension-Reaktion eine Ausgabedatei, das so genannte *genepix result (gpr) file*, vor. Im Rahmen der Validierung des PR-Chips wurden also 96, für den RT-Chip 48 Dateien erstellt, die in die etablierte Software eingelesen wurden. Aus diesem umfangreichen Datensatz wurden nur bestimmte Werte ausgewählt und anschließend weiter bearbeitet: die Medianwerte der Signal- und Hintergrundintensitäten für die vier gemessenen Wellenlängen an jeder Position, sowie die den Werten zugrunde liegende Pixel-Anzahl.

Lokale Hintergrundkorrektur

Eine erste Korrektur der Werte erfolgt durch die Subtraktion des unmittelbaren Hintergrundes von der Intensität. Eine Angleichung verschiedener Spotgrößen, bedingt durch den Herstellungsprozess, wird erreicht mittels einer Multiplikation mit der jeweiligen Pixelanzahl.

Globale Hintergrundkorrektur

Die verwendeten Oligonukleotide bilden eine Vielzahl von Sequenzvarianten ab, im Einzelfall hybridisiert das zu untersuchende PCR-Produkt jedoch nur an einigen wenigen. Alle übrigen Primer werden nicht verlängert, in der Folge geben die Mehrzahl der Positionen auf dem Array kein Signal. Für die globale Hintergrundkorrektur wurden deshalb solche Positionen als Rauschen definiert, deren Intensitätswert kleiner ist als der Median aller Intensitäten, die mit der gleichen Wellenlänge ermittelt wurden. Eine Position entspricht einem *spot* auf dem Micoarray und repräsentiert die Sequenzen eines Oligonukleotids.

Filtern der Daten

Zusätzlich sollten alle Positionen ausgefiltert werden, von denen kein signifikantes Signal ausgeht. Hierzu galt es, ein ideales Qualitätskriterium zu definieren. Unter bestimmten Bedingungen ist die Basenzuordnung (auch in der Referenzsequenz) für eine Nukleotidposition nicht eindeutig. Das ist dann der Fall, wenn eine Probe verschiedene Virus-Populationen enthält, die sich an der entsprechenden Stelle in der Sequenz unterscheiden. In der Folge erhält man mehrere Signale, deren Intensitäten sich auf die entsprechenden Wellenlängen verteilen. Der Intensitätsschwellenwert muss so gewählt werden, dass auch solche Mischungen berücksichtigt werden.

Zunächst wurde jeweils die Summe der vier Intensitätswerte für eine Position (entsprechend den vier Wellenlängen, bei denen detektiert wurde) gebildet und der Median dieser Summen

errechnet. Von diesem Wert ausgehend, wurden verschiedene Filterkriterien getestet. Aus Diagramm 13 geht hervor, dass bei Anwendung unterschiedlicher Filterkriterien Auswirkungen auf die Anzahl der ausgefilterten Positionen und die der fehlerhaften Ergebnisse zu beobachten sind. Als Schwellenwerte wurden in aufsteigender Reihenfolge der Median aller Signalintensitäten, sowie je Median plus 0,25, 0,5, 0,75 und einmal Standardabweichung getestet, dargestellt auf der x-Achse. In Abhängigkeit davon ist die Anzahl der untersuchten Nukleotidpositionen aufgetragen, die nicht ausgefiltert wurden, also zur weiteren Auswertung herangezogen werden können (rote Linie) und die Anzahl der fehlerhaft bestimmten Positionen (blaue Linie). Durchgeführt wurde die Datenanalyse für ein Protease-spezifisches PCR-Produkt, die höchstmögliche Zahl der untersuchten Positionen liegt also bei 78. Im gezeigten Fall sind auch für alle untersuchten Positionen Sonden vorhanden, deren Signalintensität über dem Medianwert aller Signale liegt. Erhöht man diesen Schwellenwert, geht die Information für einige Stellen verloren. Umgekehrt verringert sich aber auch die Anzahl der fehlerhaft bestimmten Nukleotide. Berücksichtigt man zur weiteren Analyse nur diejenigen Signale, deren Intensität höher als der Median plus der zugehörigen Standardabweichung liegt, verbleiben nur noch 49 von 78 Nukleotidpositionen, diese sind dann allerdings alle völlig korrekt bestimmt.

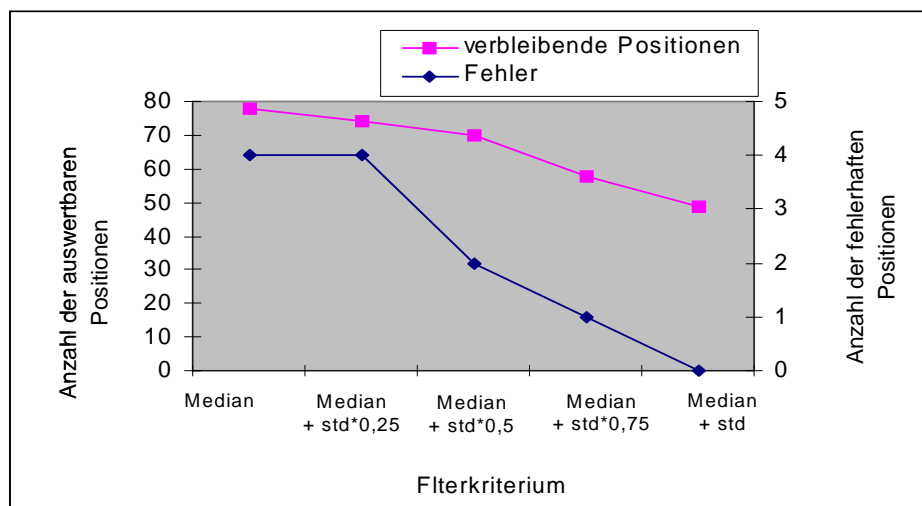


Diagramm 13: Vergleich verschiedener Filterkriterien

Je stringenter das Filterkriterium, desto höher die Anzahl der Positionen, die ausgefiltert werden, und desto niedriger die Fehlerrate.

Für die Validierung des PR- und RT-Arrays wurde als Schwellenwert der Median definiert. All diejenigen Positionen, für die die Summe unterhalb des Medians lag, wurden ausgefiltert, die übrigen als signifikant definiert. Dieses wenig stringente Filtern lässt sich deshalb vertreten, da auch nicht vorhandene, weil ausgefilterte Positionen als falsch gelten.

Möglicherweise falsch bestimmte Sequenzen hingegen lassen sich anhand der Referenzen identifizieren.

Normalisierung

Am Beispiel der SE-Kontrollen (Diagramm 4) wird deutlich, dass auch die Verlängerung von Primer gleicher Sequenz mit verschiedenen markierten ddNTPs zu unterschiedlich hohen Signalintensitäten führen. In der Auswertung wurden deshalb auch Normalisierungsschritte vorgenommen, die ungleiche Quantenausbeuten der Fluoreszenzfarbstoffe, aber auch das uneinheitliche Hybridisierungsverhalten unterschiedlicher Sonden berücksichtigen. Die Angleichung dieser technischen und experimentellen Parameter ermöglicht ein sinnvolles Gegenüberstellen verschiedener Datensätze.

Zunächst sollte das uneinheitliche Hybridisierungsverhalten verschiedener Oligonukleotide berücksichtigt werden, bedingt durch unterschiedliche Schmelztemperaturen oder die Fähigkeit zur Sekundärstrukturbildung. Dazu wurde jede Signalintensität durch die Summe der zugehörigen vier Intensitätswerte dividiert und somit jede Position auf den Wert 1 normiert. Dieser Rechenschritt erlaubt eine schnelle Beurteilung der Anteile am Gesamtsignal.

Des Weiteren sollten durch die Software Abweichungen verrechnet werden, die sich aus ungleichen Quantenausbeuten der Fluoreszenzfarbstoffe ergeben. Alle Intensitätswerte, die mit einer Wellenlänge korrelieren, wurden dafür jeweils der Größe nach sortiert. Einer der Farbstoffe (hier Cyanine5) wurde willkürlich als Standard definiert und die verbleibenden an diesen angepasst. Die Werte des Standards wurden in einem Punktdiagramm auf der x-Achse aufgetragen, und der jeweilige nächste Datensatz auf der y-Achse. Durch jeweils das 5%- und 95%-Quantil wird eine Regressionsgerade gelegt, die der linearen Geradengleichung entspricht. Als Ergebnis erhält man vier Geraden, deren Steigungen mittels multiplikativer Korrektur angeglichen werden können. In einem zweiten Rechenschritt werden auch die y-Achsenabschnitte aneinander angepasst. Dadurch erreicht man eine Normalisierung über die unterschiedlichen Fluoreszenzfarbstoffe.

Ausgabedatei

Es hat sich als sinnvoll erwiesen, zur besseren Übersicht zwei Ausgabedateien für einen eingelesenen und bearbeiteten Datensatz (*gpr file*), also für ein Experiment, zu erstellen:

Zum einen eine sehr detaillierte Liste aller Oligonukleotide, die ein statistisch signifikantes Signal liefern und somit zur Auswertung herangezogen werden können (*result file*). Für eine

abgefragte Nukleotidposition verbleiben im Einzelfall unterschiedlich viele Oligonukleotide in dieser Liste, abhängig von der ursprünglichen Anzahl auf dem Array und den Hybridisierungseigenschaften des zu untersuchenden PCR-Produktes. Für jedes dieser aufgeführten Oligonukleotide werden die beiden größten Signalanteile angegeben, vorausgesetzt die Werte liegen über einem Anteil von 0,3. Diese Grenze wurde willkürlich festgelegt und kann gegebenenfalls je nach Fragestellung variiert werden. Anteile unterhalb dieser Grenze werden der Übersicht wegen nicht in der Ausgabedatei angezeigt.

Die zweite Ausgabe beinhaltet die eindeutige Sequenzaussage für die einzelnen Nukleotidpositionen (*sequence file*). Signale, die von verschiedenen Oligonukleotiden stammen, die die gleiche Nukleotidposition abfragen, werden zusammengefasst. Mischungen werden beim derzeitigen Stand des Programms noch nicht automatisch erkannt, weshalb die Ausgabe der detaillierten Liste notwendig ist.

Sequenzvergleich

Zur Validierung des Microarrays wurden ausschließlich PCR-Produkte analysiert, für die bereits Sequenzen vorlagen. Ermittelt wurden diese Daten mittels Standard-Sequenzierung nach Sanger. Im Rahmen der Validierung war es wichtig, einen direkten Vergleich zwischen ermittelter Sequenz und der Referenzsequenz durchführen zu können. Die Software wurde daher so konzipiert, dass eine jeweilige Referenz eingelesen werden kann und dann ebenfalls in der Ausgabedatei erscheint.

manuelle Auswertung von Mischungen

Mischungen werden beim derzeitigen Stand des Programms noch nicht automatisch erkannt, aus diesem Grund ist es notwendig, das detaillierte *result file* auszugeben. Beide Ausgabedateien können mit einem Tabellenkalkulations-Programm geöffnet und weiter bearbeitet werden. Zunächst wurde im direkten Vergleich die ermittelte Sequenz pro Nukleotidposition mit der jeweiligen Referenz im *sequence file* überprüft. Bei Abweichungen konnte im *result file* auf die Information über die einzelnen Oligonukleotide zurückgegriffen werden. Im Einzelfall wurde dann entschieden, ob eine Mischung, also eine Verteilung der Signalintensitäten auf zwei Wellenlängen (detaillierte Erklärung unter 6.7.3), oder ein eindeutig abweichendes Ergebnis vorlag. Nach der Klassifizierung der Daten, wurde der prozentuale Anteil der „richtigen“, also übereinstimmenden und der „falschen“, das heißt der von der Referenz abweichenden Nukleotide an der Gesamtanzahl der untersuchten Positionen

bestimmt. In dieser Weise wurden alle Positionen in allen generierten Ausgabedateien ausgewertet.

6.11 HIV-assoziierte Gene (Pilotprojekt)

Bestimmte Ausprägungen oder Mutationen humaner Gene werden mit der Infektionswahrscheinlichkeit, Krankheitsprogression oder dem Therapieverlauf von HIV / AIDS in Verbindung gebracht. Anhand zweier Beispiele soll im Folgenden gezeigt werden, dass solche Mutationen ebenfalls mittels APEX-Technologie nachgewiesen werden können. Dies könnte eine sinnvolle Ergänzung des Microarray-basierten HIV-Resistenztests sein.

6.11.1 Multi Drug Receptor Gene 1 (MDR1)

Etablierung einer spezifischen PCR

Zunächst musste eine PCR etabliert werden, um die entsprechende Region innerhalb des *MDR1*-Gens aus humaner, genomischer DNA zu amplifizieren. Die Nukleotide 3434-3436 sind dabei von besonderem Interesse, für sie werden im Folgenden APEX-Sonden abgeleitet. Zur Etablierung eines PCR-Produktes, das diese Nukleotide abdeckt, wurden zunächst mit dem Programm *Primer Premier 5* Primer insgesamt fünf verschiedene Primer abgeleitet. Alle möglichen Primerkombinationen aus drei forward- und zwei reverse-Primern wurden getestet, Lage und Sequenz sind im Materialteil aufgeführt. Als Matrize wurden je 1µg eigener genomischer DNA eingesetzt.

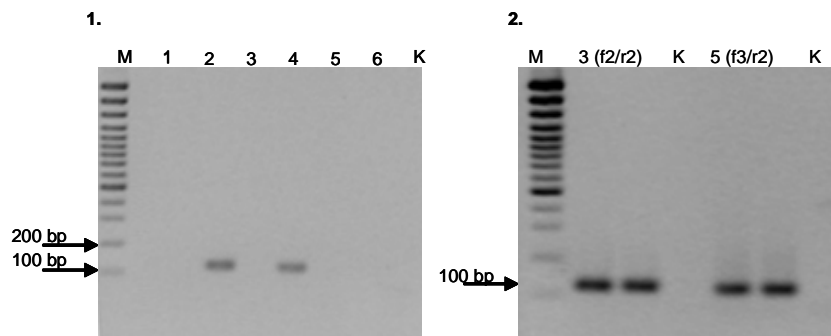


Abbildung 15: Etablierung einer MDR-spezifischen PCR

1. Kombination verschiedener *mdr*-spezifischer PCR-Primer:

Spur 1: MDR1 f1/r2 (208 bp)	Spur 2: MDR1 f1/r1 (278 bp)
Spur 3: MDR1 f2/r2 (116 bp)	Spur 4: MDR1 f2/r1 (186 bp)
Spur 5: MDR1 f3/r2 (114 bp)	Spur 6: MDR1 f3/r1 (184 bp)
 2. Reamplifikation der PCR-Produkte aus den Ansätzen 3 und 5
- K = Negativkontrollen der PCR-Reaktion; M = 100 bp DNA-Elektrophoresemarker

Abbildung 15.1 zeigt das Agarosegel. Ausschließlich die Kombinationen *mdr1* f2/r2 und f3/r2 (Spuren 3+5) führen zur Bildung eines PCR-Produktes (Abbildung 15.1). Aus diesen Ansätzen konnten in einer zweiten Reaktion die PCR-Produkte in höherer Konzentration reamplifiziert werden (Abbildung 15.2). Im Anschluss folgten die Reinigung der Amplifikate und die enzymatische Fragmentierung durch Uracil-N-Glykosylase (nicht gezeigt). Die so erhaltenen, verdauten Amplifikate wurden für die APEX-Reaktion eingesetzt.

MDR-spezifische Sonden

Als Grundlage diente die mRNA-Sequenz, die unter der GenBank Accession Number NM_000927 (Saji, *et al.* 2003) verfügbar ist. Für die Primer Extension wurden Sonden abgeleitet, die die Nukleotide 3434-3436 abfragen, also jeweils an der Position Minus 1 enden. Da im Gegensatz zu den HIV-Sequenzen die Bereiche strangauf- und -abwärts der zu untersuchenden Positionen konserviert vorliegen, reicht jeweils ein Oligonukleotid als Sonde pro Strang zur Detektion aus. Zur Abdeckung von beiden DNA-Strängen wurden also insgesamt sechs Sonden mit einer Länge von je 30 Basen plus 20 d(T) und C6-Aminomodifikation am 5'-Ende verwendet. Die Sonden sind analog den HIV-Primer bezeichnet mit dem Gennamen, der Orientierung (f oder r), der Nukleotidposition (3434-3436) und dem für den Wildtyp erwarteten Einbau.

MDR-spezifische APEX-Reaktion

Die Herstellung der Microarrays und die Primer Extension wurden wie im Methodenteil beschrieben durchgeführt, das Temperaturprofil entspricht dem des PR-Chips. Diagramm 14

zeigt für jede Sonde den Anteil am Gesamtsignal. Dieser kommt zustande, indem jeder Wert durch die Summe der vier zusammengehörigen Intensitäten pro Spot (für, Cy5, Cy3, Fluorescein und TexasRed) dividiert wird. In der Summe ergeben diese Anteile den Wert 1. Für jede abgefragte Position ist eine eindeutige Basenzuordnung möglich. Die Aussagen von Strang und Gegenstrang korrelieren und bilden je komplementäre Basenpaare (3434: G/C, 3435: C/G, 3436: T/A). Die Resultate konnten in sechs Wiederholungen mit beiden PCR-Produkten (MDR1 f2/r2 bzw. f3/r2) reproduziert werden. Die detektierte Sequenz entspricht dem Wildtyp und konnte durch Sequenzierung des Amplifikates verifiziert werden. Eine Sequenz mit veränderter Position 3435 war leider nicht verfügbar.

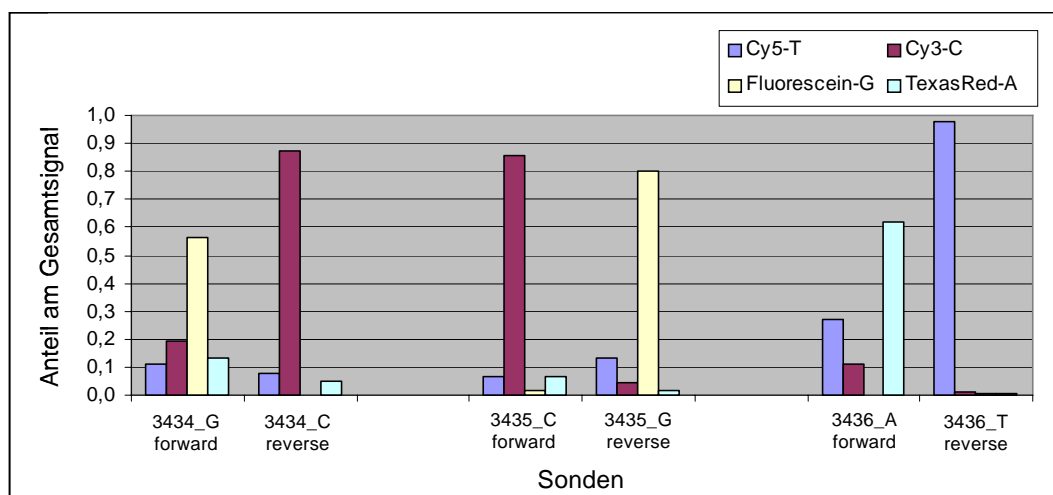


Diagramm 14: Nachweis der Nukleotide 3434-3436 im MDR1-Gen mittels APEX

Signalverhältnisse, die durch enzymatische Verlängerung der MDR1-spezifischen Sonden erreicht werden. Für jede Position lassen sich eindeutige Basenzuordnungen für Strang und Gegenstrang treffen: 3434: G/C; 3435: C/G; 3436: A/T

6.11.2 CCR5 Δ 32

Etablierung einer spezifischen PCR

Auch für die Δ 32-Mutation im humanen Chemokin-Rezeptor 5 wurde zunächst ein spezifisches PCR-System etabliert, basierend auf der linearen mRNA-Sequenz mit der GenBank Accession Number BC038398. Hierfür wurden verschiedene Primerkombinationen getestet, die den Bereich amplifizieren, den die entsprechende Deletion betrifft (554-585). Als Ziel-DNA für die Primer Extension wurde schließlich ein Amplifikat mit einer Länge von 184 Basenpaaren verwendet (451-635).

CCR5 Δ 32-spezifische Sonden

Da es sich bei dieser Mutation nicht um einen Einzelnukleotid-Austausch handelt, sondern um eine Deletion von 32 Basen, wurden Sonden abgeleitet, die folgende Kriterien erfüllen:

CCR5_f1_554_A/T: endet direkt vor Beginn der Mutation, sollte im WT mit A, bei Δ 32 mit T verlängert werden

CCR5_f1_wt584_C: liegt komplett im Δ 32 Bereich, sollte im WT mit C verlängert werden, bei Mutation kein Signal geben

CCR5-spezifische APEX-Reaktion

Die Herstellung der Microarrays und die Primer Extension-Reaktion folgt der im Methodenteil angegebenen Vorgehensweise. Das Temperaturprofil stimmt mit dem überein, das für den PR-Chip verwendet wurde (40°C / 65°C). Diagramm 15 zeigt die Anteile am auf den Wert 1 normalisierten Gesamtsignal für die Sonden CCR5_f1_554_A/T und CCR5_f1_wt584_C. Die eingebauten Nukleotide entsprechen denen einer Verlängerung komplementär zum Wildtyp. Die Ergebnisse konnten durch Sequenzierung verifiziert werden.

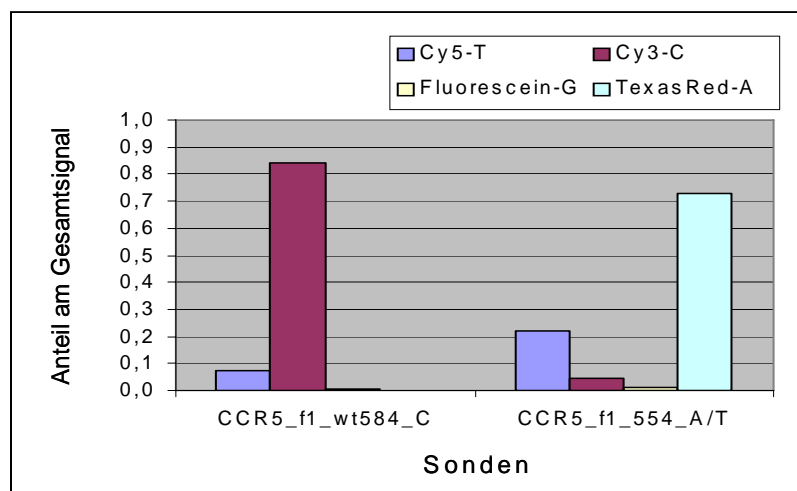


Diagramm 15: Signalintensitäten für die APEX-Reaktion zur Detektion der CCR5 Δ 32-Mutation

Der Nukleotideinbau entspricht dem für den Wildtyp erwarteten.

7 Diskussion

7.1 Optimierungsversuche

Innerhalb dieser Arbeit wurde ein genotypischer Resistenztest, basierend auf der APEX-Technologie, zum Nachweis der relevanten Mutationen innerhalb der HIV-Protease und – Reversen Transkriptase entwickelt. Dafür wurden ausgehend von den nach Kurg *et al.* (2000) beschriebenen Reaktionsbedingungen, verschiedene Parameter, die in Abhängigkeit untereinander stehen, optimiert und die APEX-Reaktion erfolgreich für den Nachweis dieser Resistenz-assoziierten Mutationen im HIV-Genom etabliert.

Um eine Hybridisierung und Verlängerungsreaktion der Sonden durch die Polymerase zu erreichen, muss die Schmelz- auf die Reaktionstemperatur abgestimmt werden. Während eine zu niedrige Temperatur unspezifische Hybridisierung und falsch-positive Ergebnisse fördert, führt eine zu hohe Temperatur zu einem vorzeitigen Dissoziieren der Ziel-DNA und einem Signalverlust. Das optimierte Temperaturprofil (6.3.1) konnte auf optimale Ergebnisse hinsichtlich Spezifität und Signalstärke abgestimmt werden. Aufgrund der hohen Heterogenität des HIV-Genoms variierte die Sondenlänge zwischen 25-30 Basen. Es konnte gezeigt werden, dass für die Reaktion auch bei variierender Schmelztemperatur, bedingt durch die Länge, stabile Ergebnisse erzielt werden.

Im Vergleich verschiedener Spottinglösungen für die Immobilisierung der Oligonukleotidsonden (50mere, einzelsträngig) an der Festphase des Microarrays, wurden bei Verwendung einer Epoxysilanoberfläche mit Natriumcarbonat-Puffer (pH 9) die eindeutig stärksten Signal- / Rausch-Verhältnisse erreicht (6.3.2 und 6.3.3). Dies korreliert mit Ergebnissen innerhalb der Arbeitsgruppe hinsichtlich der Herstellung von cDNA-Microarrays zur Untersuchung von Genexpressionsprofilen (Bauer, 2004). Die Kombination von Epoxysilan-beschichteten Glasträgern und Natriumcarbonat-Puffer führt auch bei der Immobilisierung längerer, doppelsträngiger Nukleinsäurefragmente (PCR-Produkte, 2kb) und anschließender Hybridisierung zu optimalen Signalausbeuten. Auch Lovmar *et al.* (2003) verwenden für ihren Array-basierten Primer Extension-Assay die gleichen Pufferbedingungen und Oberflächen.

Durch Verwendung von C6-Aminomodifikationen am 5'-Ende der Sonden konnte weiterhin die gemessenen Signalintensitäten nach enzymatischer Verlängerung durch die

DNA-Polymerase verbessert werden (6.3.4). Bei identischen Sequenzen, Konzentrationen und Kopplungsbedingungen auf dem Microarray kann im direkten Vergleich für Sonden mit 5'-Aminomodifikation eine höhere Signalstärke gemessen werden. Zurückzuführen ist dies auf die spezifische Anbindung der Oligonukleotide über die endständige Aminogruppe an die Epoxygruppe an der Oberfläche. Die Wahrscheinlichkeit einer Anbindung über das 5'-Ende wird erhöht, so dass das freie 3'-Ende der Prozessierung durch das Enzym zur Verfügung steht. Sterische Hinderung wurde durch den Einsatz von Oligo-d(T)-Spacer und durch Fragmentierung der Ziel-DNA minimiert. Dadurch konnten optimale Signal- / Hintergrund-Verhältnisse für die APEX-Reaktion erreicht werden.

7.2 Kontrollsonden

Zur Orientierung innerhalb des Arrays zum einen und der Überprüfung der kovalenten Anbindung an die Oberfläche zum anderen (6.4) wurden Kontrollsonden (PK) entwickelt und eingesetzt.

Innerhalb dieser Arbeit wurden weiterhin so genannte SE-Sonden etabliert, die eine Kontrolle der APEX-Reaktion unabhängig von Ziel-DNA und spezifischen Sonden ermöglicht. Besonders beim Einsatz neuer Oligonukleotide oder unbekannter Template-DNA ist die Verwendung unabhängiger Sonden von großem Vorteil.

Aus dem Ergebnis der Selbstverlängerung geht auch klar die Wichtigkeit der Datennormalisierung hervor, die bei der Entwicklung der Analyse-Software berücksichtigt wurde. Betrachtet man Diagramm 4, würde man erwarten, für die vier Extensionskontrollen annähernd gleich starke Signalintensitäten zu erhalten. Die Sonden unterscheiden sich lediglich in einer Nukleotidposition auf einer Gesamtlänge von 80 Basen. Da diese Position nicht innerhalb des Doppelstranges liegt und somit nicht zur Hybridisierung herangezogen wird, sollte die Einbaueffizienz gleich sein. Es konnte allerdings gezeigt werden, dass die Signalstärke in der Praxis reproduzierbar stark variiert. Die Ursachen für diese Abweichungen können in den unterschiedlichen Quantenausbeuten der Fluoreszenzfarbstoffe und in der Qualität der zur Detektion verwendeten Laser liegen. Obwohl die Normalisierung über andere Wege erreicht wurde, wäre auch eine Datennormalisierung über die SE-Oligonukleotide denkbar. Voraussetzung ist dann allerdings eine Immobilisierung in entsprechend hoher Anzahl auf dem Microarray.

7.3 Verwendung degenerierter Sonden

Als Grundlage für das Sondendesign zum Nachweis der Resistenz-assoziierten Codons innerhalb der HIV-PR und -RT lag ein sehr breites Sequenz-Variationsspektrum vor. Aufgrund dieser immensen Sequenzvariabilität mussten Strategien erarbeitet werden, um die Anzahl der Oligonukleotide auf dem Microarray zu verringern. Um mit möglichst geringem Sondenaufwand eine möglichst hohe Sequenzabdeckung zu erreichen, wurde der Einsatz degenerierter Primer etabliert. Eine Primersequenz gilt dann als degeneriert, wenn eine oder mehrere Positionen mehr als eine Base repräsentiert. Der Grad an Degeneration entspricht der Anzahl der Sequenzmöglichkeiten, die er enthält (Linhart & Shamir, 2002).

Über die Verwendung von degenerierten Sonden zur Primer Extension auf Microarrays lagen keine Literaturdaten vor hinsichtlich Anzahl und Lage der Modifikationen, sowie über den Einfluss auf die Einbaurrate der Polymerase und die daraus resultierende Signalstärke. Aus diesem Grund wurde im Rahmen dieser Arbeit für HIV ein Modellsystem für Sonden im Bereich der Protease-Sequenz konstruiert, um den Einfluss systematisch eingeführter, degenerierter Basen durchzuspielen. Es wurde gezeigt, dass der störende Einfluss von Mischungen am 5'-Ende deutlich geringer ist als am 3'-Ende des Primers, das zur Verlängerung durch die DNA-Polymerase dient. Aus den Versuchen geht hervor, dass mit einem Mindestabstand von fünf Nukleotiden zum 3'-Ende innerhalb eines Primers bis zu drei bis vier Positionen ersetzt werden können. Dies geht zwar im Vergleich zur unmodifizierten Sequenz mit einer Verringerung der Signalstärke einher, es werden aber trotzdem signifikante Signal- / Hintergrundverhältnisse erreicht. Das Austauschen von drei bis vier Basen innerhalb der Primersequenz durch degenerierte Basen führte zu Degenerationsgraden von sechs bis acht, da ausschließlich Mischungen von zwei Basen vorgenommen wurden ($K=G+T$, $M=A+C$, $R=A+G$, $W=A+T$, $Y=C+T$).

Lag eine Sequenz vor, in der eine variable Position drei oder alle vier möglichen Nukleotide umfasste, wurde sie durch ein Inosin ersetzt. Inosin zeichnet sich durch die Fähigkeit aus, mit allen anderen Basen Wasserstoffbrücken ausbilden zu können. Wie bei den „Wobble“-Modifikationen war der Einsatz allerdings nur in der Mitte oder am 5'-Ende der Primersequenz möglich. Es wurde gezeigt, dass der Austausch eines Nukleotids nahe des 3'-Endes, die Verlängerung durch die Polymerase stark beeinträchtigt (6.5.2).

Im Gegensatz zur Anwendung an der Festphase beschreibt die Literatur mehrfach den Einsatz degenerierter Primer zur Amplifikation von PCR-Produkten in Lösung, wenn die Sequenz der

Matrizen-DNA an einzelnen Positionen variiert oder evt. unbekannt ist. Dies wird häufig angewendet, wenn es um den Nachweis genetisch hochvariabler, viraler Sequenzen geht (z.B. Dovas *et al.*, 2004; Huang *et al.*, 2005; Kilpatrick *et al.*, 1996).

Im Bereich der Oligonukleotid-Array-Technologie dagegen findet sich eine Einsatzmöglichkeit in der Hybridisierung gegen variable Ziel-DNA (ohne enzymatische Prozessierung). Chizhikov *et al.* (2002) nutzen beispielsweise degenerierte Sonden zur Array-basierten Detektion klinisch relevanter Genotypen von Humanen Rotaviren (HRV), die Hauptverursacher von Gastroenteritis bei Kindern weltweit. Das RNA-Virus zeigt wie HIV einen hohen Grad an genetischer Diversität auf und lässt sich im Vergleich zu stabileren DNA-Viren ebenfalls nur schwer genotypisieren. Um diese Variabilität abzudecken, kommen Sonden zum Einsatz, die Modifikationen in Form von „Wobble“-Positionen oder Inosin als universeller Base tragen. Pro Oligonukleotid werden höchstens drei Positionen degeneriert. Da die Sonden ausschließlich zur Hybridisierung dienen, ist im Gegensatz zum APEX-Assay ein Mindestabstand zum 3'-Ende unerheblich. Denkbar ist hier, dass der störende Einfluss der Modifikation an oder nahe der Mittelposition am größten ist.

Neben der Verringerung der benötigten Sondenanzahl durch die Verwendung von „Wobble“-Positionen und Inosin, sind auch andere Möglichkeiten denkbar. Es existieren zahlreiche Nukleotidanaloga, die alternativ zu diesem Zweck eingesetzt werden könnten. 6H, 8H-3,4-dihydropyrimido[4,5-c][1,2]oxazin-7-one zum Beispiel bezeichnet ein Pyrimidin-Analogon (P), N⁶-methoxy-2,6-diaminopurine (K) kann alternativ zu Purinen eingebaut werden. In equimolarer Mischung wirken sie als universelle Base M (Hill *et al.*, 1998). Allerdings ist die Synthese von Oligonukleotiden mit diesen Modifikationen derzeit weitaus weniger verbreitet und sehr viel kostenintensiver als die Verwendung von Inosin.

Innerhalb dieser Arbeit konnte die Verwendung degenerierter Oligonukleotidsonden für die AEPX-Reaktion erfolgreich etabliert werden. Erst dadurch war es möglich, eine derart hoher Sequenzabdeckung technisch und finanziell zu realisieren.

7.4 PR / RT-Microarray - mögliche Fehlerquellen

Für die Validierung des PR- und RT-Microarrays wurden 94 bzw. 48 PCR-Produkte bekannter Sequenz analysiert (6.7 und 6.8). Die Amplifikate haben ihren Ursprung in Patientenmaterial und umfassen sowohl HIV Subtyp B- als auch nicht-Subtyp B-Proben. Die jeweilige Vergleichssequenz wurde durch herkömmliche Sequenzierung nach Sanger bestimmt. Da es sich bei den Proben bereits um RT-PCR-Produkte handelte, wurden diese zur Vorbereitung für die APEX-Reaktion unter Einbau von dUTP reamplifiziert. Sowohl die PCR als auch der darauf folgende enzymatische Verdau durch die Uracil-N-Glykosylase konnten für alle Proben erfolgreich durchgeführt werden.

Die innerhalb dieser Arbeit für den Microarray entwickelten Sonden berücksichtigen alle Sequenzvarianten, die mit einer Frequenz von >5% (PR) bzw. >10% (RT) in der Los Alamos HIV-Sequenz-Datenbank existierten. Weiterhin wurden solche Varianten einbezogen, für die eine Assoziation mit der Entwicklung von Resistenzen beschrieben ist, unabhängig von ihrer Häufigkeit. Die Anzahlen der verwendeten Sonden, die großenteils degenerierte Positionen enthielten, betragen 368 (PR, SONDENSATZ 1), 452 (PR, SONDENSATZ 2) und 523 (RT).

Mit einem ersten SONDENSATZ konnte für den PR-Chip insgesamt in Abhängigkeit davon, ob Mischungen als falsch oder richtig bewertet wurden, eine 92-94%ige Übereinstimmung mit den Vergleichssequenzen erreicht werden.

Für 2,5% der Positionen kann die Software aufgrund zu schwacher Signalintensitäten keine Sequenzzuweisung vornehmen. Bewertet man diese falsch-negativen Ergebnisse als neutral, erreichen die verbliebenen Resultate eine Übereinstimmung von bis zu 96,2%. Deutlich verbessern lässt sich die Detektionseffizienz durch die Ergänzung von Sonden. Das konnte für die Genotypisierung der Protease anhand von vier ausgewählten Proben mit zuvor unterdurchschnittlicher Übereinstimmungsrate gezeigt werden (6.7.4). Die Erhöhung lag bei durchschnittlich 5,5%.

Für den RT-Microarray wurden Übereinstimmungen über alle untersuchten Positionen von 86-90% im Vergleich zur Sanger-Sequenzierung erreicht. Der Anteil der falsch-negativen Ergebnisse liegt für die RT bei 4,9%, das ist knapp doppelt so hoch wie für die PR.

Auch unter optimierten Bedingungen traten bei der Analyse der RT-Sequenzen von Beginn an Schwierigkeiten auf. Obwohl der GC-Gehalt von PR- (39%) und RT-Sequenzen (38%)

vergleichbar ist, konnte die Verlängerung der Sonden nicht unter gleichen Reaktions-Bedingungen stattfinden. Aus diesem Grund musste das für die PR verwendete Temperatur-Profil weiter angepasst werden und auf mehrere Zyklen verlängert werden. Ursachen für die schlechtere Hybridisierung könnten möglicherweise Sekundärstrukturbildungen innerhalb der PCR-Produkte sein. Im Vergleich zur Protease können auch bei anderen alternativen Genotypisierungverfahren Sequenzen innerhalb der RT schlechter nachgewiesen werden (Tabelle 20). Für eine Weiterentwicklung des Microarrays könnte das Ergänzen weiterer Sonden das Ergebnis hinsichtlich der Spezifität verbessern.

Wie Diagramm 7 und Diagramm 8 zeigen, sind die Ergebnisse generell sehr stark positions-abhängig, das gilt sowohl für PR wie auch RT und ist vermutlich zurückzuführen auf Polymorphismen im Hybridisierungsbereich der Primer. Die HIV-1 Subtyp B Protease weist eine enorme genetische Variabilität von 47,5% auf, wobei viele der Resistenzmutationen als natürliche Polymorphismen in Therapie-naiven Patienten vorkommen (Kozal *et al.*, 1996). Das erschwert das Untersuchen dieser Sequenzen mittels Methoden, die ganz oder vollständig auf Hybridisierung gegen vorgefertigte Oligonukleotide beruht.

Aufgrund der hohen Variabilität wurden nicht alle möglichen Sequenzvarianten auf dem Array abgedeckt, dies schließt eine 100%ige Nachweiseffizienz aus. Es besteht eine bestimmte Wahrscheinlichkeit, dass die ZieldNA von den Sonden auf dem Array nicht erfasst wird und ein Hybridisierungsereignis ausbleibt. Daher korrelieren die erzielten Ergebnisse sehr gut mit den durch die verwendeten Oligonukleotide vorgegebenen Detektionslimits. Für die Protease konnte gezeigt werden, dass Ergänzung oder Ersetzen von Sonden die Fehlerrate verringert.

Zudem könnten in Zukunft Oligonukleotide eingesetzt werden, die den Nachweis von Insertionen besonders innerhalb der RT (Winters & Merigan, 2005) erlauben, die aufgrund der immensen Variabilität derzeit im Sondendesign nicht berücksichtigt sind. Von den Amplifikaten, die der Validierung zugrunde liegen, zeigt keines zusätzliche Nukleotide. Bei Untersuchung unbekannter Proben sollte dies aber bedacht und der Chip möglicherweise um die entsprechenden Sonden erweitert werden.

Die Möglichkeit zur Sekundärstrukturbildung von Oligonukleotiden wie sie auch bei PCR-Reaktionen in Lösung zu beobachten ist, ist in gleicher Weise für die Sonden an der Festphase des Microarrays gegeben. Im Einzelfall kann dies analog den SE-Sonden zu einer hier unbeabsichtigten Selbstverlängerung führen oder eine Prozessierung durch die Polymerase verhindern.

Eine weitere Fehlerquelle bergen die Vergleichssequenzen, die für die Chip-Validierung als „Referenzen“ herangezogen wurden. Diese Sequenzen wurden mit einem validierten kommerziellen Verfahren (ViroSeq) in der Routinediagnostik bestimmt. Zur Überprüfung ihrer Qualität wurde von sechs der 48 PCR-Produkte, die zur Validierung des RT-Microarrays analysiert wurden, in einem zweiten Labor eine weitere Sanger-Sequenzierung durchgeführt. Die Übereinstimmung dieser zwei Sequenzierungen (Virologie vs. DKFZ) lagen Amplifikat-abhängig zwischen 95,7% und 99,2% (Alignment nach Huang & Miller, 1991). Es zeigt sich, dass es sich bei denen zum Vergleich herangezogenen Sequenzen nicht um echte „Referenzen“ handelt, sondern dass auch diese eine gewisse Fehlerbelastung bergen.

Generell zeigen auch Resistenztests, die auf Standard-Sequenzierung beruhen, Probleme in der Reproduzierbarkeit (Korn *et al.*, 2003; Palmer *et al.*, 2005). Das zeigen auch Daten aus einem Ringversuch, durchgeführt 2003/2004 u.a. vom Nationalen Referenzzentrum für Retroviren in Erlangen. Dabei wurden insgesamt sechs definierte Proben (zwei davon mit Mischungen) zur Resistenztestung an 23 verschiedene Labore in Deutschland und Luxemburg verschickt, die verschiedene Genotypisierungs-Systeme verwenden (ViroSeq, TRUEGENE oder selbst entwickelte Systeme). Schwierigkeiten treten besonders dann auf, wenn die Sequenzen komplexe Mutationsmuster zeigen oder Mischungen vorliegen.

Auch für den APEX-Assay trifft eine Korrelation von Komplexität der vorhandenen Mutationen und Effizienz der Genotypisierung zu. Die Betrachtung der Validierungsdaten in diesem Zusammenhang stützt die These, dass für Proben mit weniger Mutationen weniger Positionen falsch bestimmt werden im Vergleich zu Proben, deren Sequenz stark vom WT abweicht. Eine Erklärung dafür wäre, dass die Oligonukleotide auf dem Microarray mit größerer Wahrscheinlichkeit komplementär zu einer weniger stark mutierten Ziel-DNA sind. Um eine statistisch eindeutige Aussage über den Zusammenhang treffen zu können, müsste allerdings eine größere Zahl an Proben untersucht werden.

Die für den Ringversuch zur Auswertung herangezogenen Datensätze zeigten weiterhin, dass Testsystem-abhängig nur 78-88% (ViroSeq), 16-33% (TRUEGENE) und 14-33% (eigene Systeme) der Labore alle Resistenz-relevanten Aminosäurepositionen in den beiden Proben mit Mischungen richtig bestimmen konnten. Diese Ergebnisse machen die Notwendigkeit der Entwicklung von neuen oder verbesserten Genotypisierungs-Assays deutlich, um die Qualität der Resistenztestung zu gewährleisten.

7.5 Detektion von Mischungen

Wie die Experimente zeigen, ist eine Detektion von Mischungen mit dem APEX-Chip möglich. Getestet wurden die Mischungen zweier PCR-Produkte (WT / mutierte Sequenz) in den Verhältnissen 100/0, 75/25, 50/50, 25/75 und 0/100. Der Unterschied der mutierten zur WT-Sequenz besteht nur in zwei Nukleotiden an den Positionen 28 (Codon 10) und 268 (Codon 90).

Jede der Nukleotidpositionen (28 und 268) wird von je drei Sonden abgefragt. Davon konnte allerdings nur jeweils eine pro Position zur Auswertung herangezogen werden. Die jeweils beiden anderen zeigten entweder die WT-Sequenz oder die Mutation an, aber keine Mischungen. Für diesen jeweils verbleibenden Sondentyp jedoch konnten die Verhältnisse für beide abgefragten Positionen qualitativ gut bestimmt werden. In jedem Fall war deutlich, ob eine Mischung vorlag und welche Sequenz in höherer Konzentration vorlag.

Die Möglichkeit einer Quantifizierung durch den APEX-Assay ist sehr stark Sonden-abhängig. Zudem unterscheiden sich die Sonden über den gesamten Array aufgrund verschiedener T_m und Sekundärstrukturbildungen zum Teil erheblich in ihrer Eignung zur enzymatischen Verlängerung. Es konnte gezeigt werden, dass unter gleichen Reaktionsbedingungen sehr unterschiedlich starke Signalintensitäten erreicht werden. Generell ist es nicht möglich, einen Array so auszulesen, dass keines der Signale im Sättigungsbereich liegt ohne einen Großteil der schwächeren zu verlieren. Ein direkter Rückschluss von anteiligen Signalintensitäten auf die Sequenzverteilung in einer untersuchten Mischung ist deshalb eher unsicher. Der APEX-Assays ist also für eine quantitative Auswertung eingeschränkt anwendbar, für die qualitative Bestimmung von Mischungen hingegen gut geeignet.

7.6 Datenanalyse

Der Umfang und die Komplexität der Datenmenge machten die manuelle Auswertung der HIV-Genotypisierungsdaten sehr aufwendig und wenig sinnvoll. Aus diesem Grund wurde in Zusammenarbeit mit Aslihan Ay im Rahmen einer Diplomarbeit eine computergestützte Analyseverfahren entwickelt, die spezifisch auf die besonderen Anforderungen der generierten Daten eingeht (6.10).

Für die Auswertung der Daten werden zunächst bestimmte Informationen aus dem Rohdatensatz (gpr file) ausgewählt. Die weitere Analyse basiert auf den Signalintensitäten der einzelnen Spots, also der Positionen auf dem Array, die durch das einmalige Aufsetzen der Spottingnadel entsteht. Hierfür wurde bewusst der Median der Intensität gewählt, der stabil gegenüber so genannten Ausreißerwerten ist und auf dem Mittelwert verzichtet. Solche Ausreißer können durch Unregelmäßigkeiten oder kleinste Verunreinigungen auf dem Array zustande kommen. Der globale Hintergrund hingegen trägt den Sonden Rechnung, die bei Einsatz einer jeweiligen Ziel-DNA nicht hybridisieren und nicht verlängert werden.

Durch das Filtern der Daten wird erreicht, dass nur signifikante Intensitätswerte in die weitere Analyse eingehen. Daher wurde ein geeigneter Schwellenwert ermittelt und festgelegt. Für die Auswertung des PR- und RT-Microarrays wurde nach dem Durchspielen verschiedener Möglichkeiten der Median als Grenzwert ausgewählt, um dem Vorkommen von Mischungen Rechnung zu tragen. Eine Mischung liegt dann vor, wenn sich die Gesamtintensität des Signals auf zwei oder mehrere Teilintensitäten verteilt, die mit verschiedenen Wellenlängen gemessen wurden. Solche Mischungen kommen dann zustande, wenn aus einer Patientenprobe, die unterschiedliche Viruspopulationen enthält, PCR-Produkte generiert werden, die sich an entsprechender Stelle unterscheiden.

Es wurde gezeigt, dass unter Verwendung weniger stringenter Filterkriterien, mehr falsch detektierte Positionen auftreten (falsch-positiv). Je nach Fragestellung kann es sinnvoll sein, diesen Wert zu variieren. Ein hoher Schwellenwert dagegen führt zu einer größeren Anzahl an ausgefilterten Signalintensitäten (falsch-negativ). Er nimmt somit einen Informationsverlust in Kauf, erhöht aber die Sicherheit in Bezug auf die ausgewerteten Daten. Umgekehrt lässt ein tieferer Grenzwert auch Mischungen zur weiteren Auswertung zu, was nicht zwangsweise ein Fehler in der detektierten Sequenz bedeutet. Das Risiko nicht signifikanter Intensitätswerte in der Ausgabedatei steigt hingegen.

Die Möglichkeit, im Rahmen der Chip-Validierung die ermittelten Daten mit Referenzdaten (aus Sanger-Sequenzierungen) vergleichen zu können, erlaubte für die durchgeführte Analyse einen weniger strengen Schwellenwert zum Filtern der Daten festzulegen. Falsch-positive Ergebnisse konnten als solche unmittelbar erkannt werden.

Wie schon unter 7.2 beschrieben, ist die Normalisierung der Daten unumgänglich, um eine Vergleichbarkeit der vier verwendeten Farbstoffe innerhalb eines Experiments zu erreichen. Durch die Angleichung der zum Teil erheblichen, durch unterschiedliche Hybridisierungseffizienzen der Sonden bedingten Intensitätsunterschiede wird ein Vergleich zwischen verschiedenen Microarrays ermöglicht. Um dies zu erreichen, werden alle Signale

anteilig am Wert 1 angegeben, was eine direkte Gegenüberstellung der Signal- / Hintergrundverhältnisse ermöglicht.

Die entwickelte Software ermöglicht eine schnelle und effiziente Datenanalyse. Allerdings muss derzeit bei Vorliegen einer Mischung stellenweise manuell ausgewertet werden. Maßgeblich für die Sequenzzuweisung ist das am stärksten repräsentierte Nukleotid, unabhängig vom Anteil des zweit- und drittgrößten Wertes. Der nächste Schritt wäre, neben den Basen A, C, G und T auch Mischungen wie R, G, Y usw. automatisch bestimmen zu können. Für die Weiterentwicklung wäre außerdem die Umsetzung der ermittelten Codons in Aminosäuren denkbar. Die graphische Darstellung der Ergebnisse in farblicher Unterscheidung könnte die Auswertung weiter vereinfachen. Vorstellen könnte man sich auch die Einbindung der Analyse in eine Datenbank mit weiteren Optionen zur Verarbeitung von zusätzlichen Patienteninformationen oder den Vergleich mit Ergebnissen früherer Resistenztests.

Parallel zur Etablierung des HIV-Resistenz-Tests in dieser Arbeit wurde ein geeignetes Auswertungsverfahren entwickelt, weil bereits bestehende Softwarelösungen die besonderen Anforderungen an die Datenanalyse nicht erfüllten.

Zur Auswertung von Daten, die durch APEX-Experimente generiert werden, entwickelte beispielsweise die Firma Asper Biotech (Tartu, Estland) ein kommerziell erhältliches Programm. Die *Genorama™ Genotyping Software 4.1* bietet benutzerfreundlich das Einlesen der Vierfarbbilder, die Zuordnung der Sondennamen und die Messung der Signalstärken. Zum Filtern der Daten ist die Definition eines Schwellenwerts möglich, die Normalisierung kann flexibel über die Signalverteilung in den einzelnen Kanälen, den maximalen Wert oder die schwächsten Signale vorgenommen werden. Zuletzt nimmt die Software die Zuweisung der Basen vor. Allerdings ist das Programm vornehmlich für den klassischen Nachweis von SNPs über je eine Sonde auf Strang und Gegenstrang konzipiert und erlaubt weiterhin nur die Verwendung von Spottingduplikaten (Benutzeranleitung, www.asperbio.com). Für die Auswertung der HIV-Genotypisierung mit einer Vielzahl an Sonden für eine abgefragte Position, die zumeist nur eine Orientierung abdeckten, war diese Software somit nicht geeignet.

Seit Anfang März dieses Jahres ist außerdem unter der Bezeichnung *MACGT* (multi-dimensional automated clustering genotyping tool) eine frei zugängliche Software zur Auswertung von Vierfarb-Primer Extension-Experimenten online verfügbar. Etwas flexibler

kann nun für jede Position ein ganzer SONDENSATZ analysiert werden, deren Aufbringung in frei wählbarer Replikatanzahl erlaubt ist, allerdings auch hier nur für Strang und Gegenstrang. Die Normalisierung wird über Positivkontrollen vorgenommen, das Programm beinhaltet im Weiteren die automatische Sequenzzuweisung und eine Visualisierung in 2-D oder 3-D-Darstellung (Walley *et al.*, 2006; www.mrl.ubc.ca/who/who_bios_scott_tebbutt.shtml).

7.7 Vergleich alternativer Methoden

Generell findet man in der Literatur zahlreiche Alternativ-Assays zur HIV-Genotypisierung. Aufgrund unterschiedlicher Ausgangsbedingungen ist es jedoch schwierig, einen direkten Vergleich zwischen den beschriebenen Methoden anzustellen. Die Unterschiede liegen z.B. in der Auswahl der nachzuweisenden Positionen innerhalb von PR und RT, Anzahl und Art der untersuchten Proben, Bewertung der Ergebnisse usw. Trotzdem soll im Folgenden ein Überblick über die bestehenden Systeme und ihre Grenzen, sowie ein Vergleich mit dem im Rahmen dieser Arbeit etablierten APEX-Assay gegeben werden.

Neben der herkömmlichen Sequenzierung nach Sanger als Standardmethode, auch erhältlich als kommerzielle Kits (z.B. TRUEGENE, Bayer; ViroSeq, Celera Diagnostics), existieren weitere Ansätze zur Durchführung von genotypischen Resistenztests in HIV. Wilson *et al* (2000) führen einen paarweisen Vergleich dreier Genotypisierungsmethoden, die kommerziell erhältlich sind, durch. Gegenübergestellt werden Line Probe Assay (LiPA, Innogenetics), HIV-GenotypR-System (Specialty Laboratories) und HIV PRT GeneChip (Affymetrix), die sich der Fragestellung in unterschiedlicher Weise nähern.

LiPA bezeichnet einen Nitrozellulose-Strip mit 33 immobilisierten Oligonukleotidsonden, der sieben Resistenz-assoziierte Codons der HIV-RT abfragt. Das Prinzip basiert auf Hybridisierung einer mit biotinylierten Primern amplifizierten cDNA gegen die Sonden. Der HIV PRT GeneChip umfasst in hoher Dichte 16 000 Sonden, die *in situ* an der Oberfläche synthetisiert werden (Lipshutz *et al.*, 1995). Nach Hybridisierung von fragmentierter, Fluoreszenz-markierter cDNA, kann er mittels konfokalen Laserscanners ausgelesen werden. In geringer Abwandlung entspricht das HIV-GenotypR-System der Standardsequenzierung.

Nach Analyse von 24 Patientenproben zeigt sich, dass nur 84% der Ergebnisse mit allen drei Assays übereinstimmend detektiert werden können. In Abhängigkeit davon, welche Methode im paarweisen Vergleich als Referenz definiert wird, liegen die Übereinstimmungswerte zwischen 84,1% und 90,2%. Es konnte gezeigt werden, dass keine Methode 100% sensitiv für

alle Mutationen in allen Quasispecies ist. Jeder Assay deckt Mutationen auf, die von den anderen jeweils nicht erfasst werden können. Unterschiedliche Quasispecies in einem Patienten in verschiedenen Anteilen verursachen die Fehlerrate von 16%.

Die Vorteile der direkten Sequenzierung liegen in der Möglichkeit, auch Insertionen und Deletionen zu detektieren, die unter Umständen auch mit Leserasterverschiebungen einhergehen. Der Affymetrix-Assay verringert durch starke Automatisierung das Risiko von Kontaminationen und Fehler in der Versuchsdurchführung, erfordert allerdings eine aufwendige Laborausstattung. Die Ergebnisse des LiPA-Ansatzes können relativ einfach ausgewertet werden durch Sichtbarwerden von Banden als farbiges Präzipitat auf den Streifen.

Beschränkte sich das getestete Prototyp-LiPA-System im Jahr 2000 noch auf die Analyse von sieben Codons der RT-Region, beschreiben Tsongalis *et al.* 2005 den Vergleich eines neuen LiPA-Assays, das acht Resistenz-assoziierte Codons in der HIV-Protease (*LiPA PR*) und elf Codons in Bereich der RT abfragt (*LiPA RT*) mit einem kommerziellen Sequenzierungssystem (TRUEGENE HIV1 Genotyping Kit, Bayer). Untersucht wurden 50 zufällig ausgewählte klinische Proben mit *LiPA PR*, 40 davon auch mit *LiPA RT*. Die Ergebnisse dieser Untersuchung werden differenziert bewertet in „übereinstimmend“, „partiell übereinstimmend“, „indeterminiert“ (durch LiPA) und „abweichend“. Sie tragen somit auch unklaren Situationen, die z.B. durch Mischungen hervorgerufen werden können, Rechnung. In exakter Übereinstimmung mit der Referenzmethode wurden 91,3% (PR) bzw. 88% (RT) der Resultate ermittelt. Die ermittelte Übereinstimmung der Ergebnisse korreliert in etwa mit derjenigen, die in dieser Arbeit durch die APEX-Methode erreicht wurde (94% PR und 89% RT). Als partiell konkordant wird ein Ergebnis dann eingestuft, wenn entweder in der einen oder anderen Methode eine Mischung vorliegt, also nur eine nicht eindeutige Sequenzzuweisung möglich ist, die nicht zwangsläufig falsch sein muss. Auch unter Berücksichtigung solcher Situationen kommt die Primer Extension-Reaktion im Vergleich mit einer Sanger-Sequenzierung zu ähnlichen Ergebnissen wie der LiPA-Assay. Der Anteil an partiell konkordanten Ergebnissen liegt mit LiPA bei 3% (PR) bzw. 5,9%, bei dem in der Arbeit beschriebenen APEX-Assay etwas niedriger bei 2% (PR) bzw. 3% (RT). Als indeterminiert sind solche Positionen definiert, die mit LiPA nicht bestimmt werden können. Das entspricht also denjenigen Positionen, die bei der Analyse der APEX-Daten aufgrund zu geringer Signalintensitäten ausgefiltert werden. Trifft das auf alle Sonden zu, die eine bestimmte Nukleotidposition abfragen, so kann an dieser Stelle keine Sequenzzuweisung

stattfinden. Die fehlende Information wird als falsches Ergebnis gewertet, ebenso wie eindeutig abweichende Sequenzen. Generell werden mit LiPA mehr Mischungen detektiert als mit der Referenzmethode, möglicherweise spricht das für eine höhere Sensitivität, was aber für die Patientenproben nicht zweifelsfrei nachgewiesen werden kann (Tsongalis *et al.* 2005).

Einen weiteren Chip zur HIV-Genotypisierung beschreiben Gonzales *et al.*, 2004. Es handelt sich dabei um einen hochauflösenden Oligonukleotid-Microarray mit einer Dichte von 69 000 Sonden pro Raster. Der nach dem Affymetrix-Verfahren hergestellte Chip fragt insgesamt 204 Resistenz-assoziierte Mutationen in PR (20 Codons) und RT (27 Codons, inklusive 2 Deletionen und 17 Insertionen), die in Gag gelegenen Proteaseschnittstellen (P24/P2, P2/P7, P7/P1, P1/P6), die Integrase und Positionen innerhalb eines pg41-Amplikon ab. Das Nachweisprinzip basiert auf der Hybridisierung amplifizierter, Fluoreszenz-markierter und fragmentierter cDNA gegen 20mer-Oligonukleotidsonden, die die DNA in jeweils beiden Orientierungen abdecken. Bei der Untersuchung von 99 Patientenproben (96% HIV1-Subtyp B) wurden übereinstimmende Ergebnisse mit Sequenzierungsdaten für die Schlüsselmutationen in PR und RT von 95% bzw. 92% erreicht, wobei keine Aussage über mögliche Mischungen getroffen wird. Für drei Codons, mit unterdurchschnittlicher Nachweiseffizienz konnten die Übereinstimmung durch Verlängerung der Sonden auf 25 Basen erhöht werden. Die Nachbesserung des Sondenlayouts führte auch für den APEX-Assay zu höheren Übereinstimmungswerten, wobei die Veränderung sowohl die Ausweitung der Oligonukleotidabdeckung auf den jeweiligen Gegenstrang als auch eine Verlängerung der Sonden bedeuten konnte (6.7.4).

Eine HIV-Genotypisierungsmethode, die wie APEX neben der Hybridisierung auch auf der Verlängerung von Oligonukleotidsonden durch eine Polymerase beruht, beschreiben Sugiura *et al.*, 2003 als „enzyme-linked Minisequencing Assay“ (*ELMA*).

Im 96well-Format legt der *Screening-Assay* sein Hauptaugenmerk auf eine möglichst einfache Durchführung, die keine teure Laborausstattung benötigt und evt. auch in Entwicklungsländern ihre Anwendung finden kann. Das Sondendesign beschränkt sich auf 6 Codons in der HIV-PR und 5 Codons für die RT und den Nachweis beschriebener Mutationen. Die Unterscheidung erfolgt hauptsächlich durch je nur zwei Sonden (WT / Mutation). Das Nachweisprinzip beruht auf dem Einbau biotinylierter Nukleotide, die Determinierung durch Streptavidin-HRP (horseradish-peroxidase) und die Sichtbarmachung

mit Hilfe des Substrates TMBBlue. Mit artifiziellen Mischungen klonierter Sequenzen konnten Minoritäten Sonden-abhängig bis 10% erkannt werden. Bei Untersuchung von 45 Patientenproben konnten 93% Übereinstimmung mit herkömmlichen Sequenzierungsdaten erreicht werden. Eine Ausnahme bildete Position 215 (RT), hier lag die Fehlerrate bei 31%. Für drei von 14 Proben konnte dies durch Polymorphismen in der Ziel-DNA begründet werden. Für die restlichen 11 falsch bestimmten Proben konnte keine Erklärung gefunden werden.

Eine weitere alternative Methode zur Genotypisierung bieten Assays, deren Funktionsprinzip auf der Ligation zweier Oligonukleotide bei vollständiger Hybridisierung gegen die zu untersuchende Ziel-DNA beruht (Landegren *et al.*, 1988). Der so genannte *Oligonukleotid Ligation Assay* (OLA) im 96well-Multititerplatten-Format beschränkt sich allerdings auf die Detektion einiger weniger Resistenz-assoziiierter Mutationen innerhalb der HIV-Protease. Bei Analyse von insgesamt 312 Triplets (nur PR) aus 54 zufällig ausgewählten Patientenproben mit unbekannter Vorbehandlung beschreiben Beck *et al.* (2002) eine Abweichung von 3,5% im Vergleich mit Sequenzdaten nach Sanger. Das Verfahren scheint gut geeignet zu sein für die Detektion von Virusvarianten, die in Mischungen als Minorität vorliegen. Sonden- und Positionsabhängig ist ein Nachweis schon ab einem Anteil von 5% möglich (Beck *et al.*, 2002; Edelstein *et al.*, 1998).

Bei Untersuchung von nicht-Subtyp B-Sequenzen stößt der Assay allerdings an seine Grenzen. Für den Nachweis ausgewählter Mutationen in Probenmaterial des Subtyps B und C, die Resistenzen gegen NRTIs (K65R, Q151M, M184V und T215Y/F) und NNRTIs (K103N, Y181C) bedingen, zeigt er sich im Vergleich mit Daten, die mit dem ViroSeq-Kit (Version 2.5) generiert wurden, als ungeeignet (Wallis *et al.*, 2005). Alle drei Literaturstellen führen falsch bestimmte Ergebnisse zurück auf eine schlechte Bindungseffizienz aufgrund von Polymorphismen nahe der Ligationsstelle der Sonden. Tabelle 20 zeigt noch einmal im Überblick die Übereinstimmung der verschiedenen HIV-Genotypisierungsmethoden mit der herkömmlichen Sanger-Sequenzierung.

Tabelle 20: Übereinstimmung verschiedener Genotypisierungsmethoden mit Standard-Sequenzierung

	LiPA	Hybridisierung	ELMA	OLA (PR)	APEX
Ref.	Tsongalis <i>et al.</i> 2005	Gonzales <i>et al.</i> , 2004	Sugiura <i>et al.</i> , 2003	Beck <i>et al.</i> , 2002	
PR	91-94%	95%	93%	96,5%	92-94%
RT	88-94%	92%		-	86-90%

Prozentuale Übereinstimmung der jeweiligen Ergebnisse mit Sequenzierungsdaten nach Sanger. Abgefragt wurden Resistenz-assoziierte Codons innerhalb der HIV-PR und -RT. Details siehe Text.

Generell war es mit keinem der Ansätze möglich, eine 100% Übereinstimmung mit den entsprechenden Referenzdaten zu erreichen. Da die Daten teilweise unterschiedlich interpretiert werden, hinsichtlich partiell übereinstimmender Ergebnisse (Mischungen) oder der Bewertung von Positionen, für die kein Ergebnis vorlag, ist der direkte Vergleich schwierig. Auch wird nicht immer deutlich, ob nur Subtyp B-Proben getestet wurden oder ob die Auswahl der Ziel-DNA willkürlich erfolgte. Für den APEX-Assay würde das beispielweise bedeuten, dass sich bei ausschließlicher Betrachtung von Subtyp B-Sequenzen die Übereinstimmung auf 87,3-90,5% (RT) bzw. 93,3-95,1% (PR) erhöhen würde.

Aus den gezeigten Literaturdaten geht in jedem Fall klar hervor, dass der in dieser Arbeit etablierte APEX -basierte Microarray hinsichtlich Spezifität und der Fähigkeit, nicht Subtyp B-Proben detektieren zu können, mit anderen, z.T. kommerziell erhältlichen HIV-Genotypisierungs-Assays vergleichbar ist.

7.8 Genotypisierung von nicht-Subtyp B-Proben

Die derzeit häufig in Routinelabors verwendeten und von der U.S. Food and Drug Administration (FDA) zugelassenen HIV1-Genotypisierungsassays sind das ViroSeq HIV1 Genotyping System (Celera Diagnostics) und das TRUEGENE HIV 1 Genotyping Kit (Bayer).

Beide Systeme sind durch Ergänzung von zusätzlichen Primern entgegen früheren Untersuchungen mittlerweile auch zur Analyse von nicht-Subtyp B-Proben geeignet, die vermehrt auch in Europa und den USA auftreten. Für den ViroSeq-Assay (Version 2.0) wird eine 98%ige Erfolgsrate bei Sequenzierung von 126 Proben unterschiedlicher geographischer Herkunft und Subtyp-Zugehörigkeit beschrieben (Eshleman *et al.*, 2004). Analog dazu

bestätigten Tong *et al.* (2005) den erfolgreichen Einsatz des TRUEGENE-Systems (Version 1 bzw. 1.5) für 99% von 208 sequenzierten Proben, ebenfalls aus verschiedenen geographischen Regionen stammend und größtenteils als nicht-Subtyp B eingeordnet. In der Literatur werden für beide Systeme eine sehr hohe, laborübergreifende Reproduzierbarkeit der Ergebnisse von >98% beschrieben (Eshleman *et al.*, 2004; Huang *et al.*, 2005).

Alternative Genotypisierungsassays hingegen, denen die Hybridisierung spezifischer Oligonukleotide zugrunde liegt, basieren oftmals auf der Sequenzgrundlage von HIV-1 Subtyp B und sind deshalb nicht uneingeschränkt zur Analyse abweichender Subtypen geeignet. Mit dem Oligonucleotide ligation assay (OLA) konnten Positions-abhängig sogar nur 0-72% der untersuchten Codons in Proben mit Subtyp C richtig identifiziert werden (Vahey *et al.*, 1999; Wallis *et al.*, 2005).

Der beschriebene PR-APEX-Assay wurde mit 18 nicht-Subtyp B-Proben HIV-infizierter Patienten getestet (13 CRF06_cpx, drei CRF02_AG, zwei nicht identifiziert). Mit dem derzeitigen Sondendesign konnten 91% Übereinstimmung mit den Referenzdaten aus der Sequenzierung nach Sanger erreicht werden (Aslihan Ay, Diplomarbeit an der Fachhochschule Bingen, 2006).

Unter den für die Validierung des PR-Microarrays zufällig ausgewählten Proben befanden sich auch nicht-Subtyp B-Sequenzen (6.9). Deutlich wird, dass Resistenz-assoziierte Mutationen mit dem derzeitigen SONDENSATZ eine höhere Nachweiseffizienz ausweisen (93,5-95,1% vs. 87,7-90,5%). Dies korreliert in etwa mit dem oben beschriebenen Ergebnis für die explizit nach Subtypzugehörigkeit ausgesuchten Proben aus Burkina Faso.

48 der 94 Proben wurde zur Validierung des RT-Microarray herangezogen, auch hier zeigen die Ergebnisse eine niedrigere Fehlerrate mit den zugehörigen Sequenzierungsdaten für die 41 Subtyp B- als für die sieben übrigen nicht-Subtyp B-Sequenzen (87,3-90,5% vs. 83,7-87,6%). Zurückzuführen ist dies möglicherweise auf schlechtere Bindungseffizienz der Sonden aufgrund nicht abgedeckter Polymorphismen. Der Anteil an Subtyp B-Sequenzen ist in der Stanford Datenbank deutlich überrepräsentiert. Daher könnte das Vorkommen bestimmter nicht-Subtyp B-Virusvarianten unter das gewählte Limit fallen und im Design der Oligonukleotide nicht berücksichtigt worden sein.

7.9 Kosten der Genotypisierung

Eine Genotypisierung wird unter bestimmten Bedingungen empfohlen, beispielsweise im Falle einer Umstellung der Behandlung infolge Therapieversagen. Die Übertragung resistenter Viren gilt inzwischen als bewiesen, deshalb sollte auch vor Einleitung der ersten Therapie eine Resistenzbestimmung durchgeführt werden (3.1.6).

Im Vergleich zur etwa doppelt so teuren Phänotypisierung liegen die Kosten für eine Genotypisierung je nach Labor noch zwischen 300-500 Euro. Neben dem Nutzen für den Patienten zahlt sich aus ökonomischer Sicht eine Genotypisierung vor Beginn von HAART dann aus, wenn das Versagen der ersten Therapie um 36% vermindert werden kann (Corzillius *et al.*, 2004). Auch Sax *et al.*, 2005 propagieren die genotypische Resistenzbestimmung als Routineuntersuchung für chronisch infizierte und Therapie-naive HIV-Patienten (auch) unter ökonomischen Gesichtspunkten.

Der ökonomische Nutzen würde dann umso größer, wenn sich die Kosten für solche Untersuchungen verringern. Steegen *et al.* (2005) beschreiben eine alternative RT-PCR mit hoher Sensitivität und anschließender Sequenzierung mit Kosten von 55 Euro/ Sequenzierung, was deutlich unter dem Preis kommerzieller Anbieter liegt.

Auch mit der APEX-Methode könnten die Kosten weiter reduziert werden. Unter der Voraussetzung einer ausreichenden Genauigkeit durch angepasstes Sondendesign, liegt der große Vorteil dieser Methode im kosten- und zeitsparenden Ablauf. Im Vergleich zur Herstellung von *high density arrays* (z.B. GeneChip, Affymetrix), deren Sonden *in situ* synthetisiert werden, ist das Aufbringen vorgefertigter Oligonukleotide auf Glasträger sehr schnell und in hohem Durchsatz zu bewerkstelligen. Aus handelsüblichen Mengen synthetisierter Primer lassen sich umgerechnet auf den einzelnen Microarray äußerst kostengünstig sehr große Stückzahlen erzeugen. Das aufgebrachte Volumen einer 50µM Lösung beträgt pro Spot etwa 1 nl. Die reinen Materialkosten für eine APEX-Reaktion liegen dann im Bereich von ca. 20-30 Euro. Ausgehend von der amplifizierte Ziel-DNA kann eine APEX-Untersuchung abhängig vom verwendeten Temperaturprofil in weniger als zwei (PR) bzw. vier Stunden (RT) ausgeführt werden. Durch die etablierte Software zur Auswertung kann schnell und einfach ein Ergebnis in Form von Sequenzzuweisungen für die relevanten Bereiche in PR und RT ausgegeben werden. Das Ergebnis einer kompletten Analyse kann in dringenden Fällen also bereits innerhalb eines Arbeitstages vorliegen. Durch die beschriebenen Zeit- und Kostenvorteile eignet sich APEX also sehr gut zur Routineuntersuchung mit hohem Probendurchsatz.

7.10 Genetischer Hintergrund (Pilotprojekt)

Neben Resistenz-assoziierten Mutationen im HIV-Genom üben auch Veränderungen in der DNA des Patienten einen Einfluss auf den Therapieerfolg und die Krankheitsprogression aus (Nelson, 1999, O'Brien & Nelson, 2004). Generell klassifiziert man solche Faktoren in Abhängigkeit davon, ob sie direkt auf die Replikation des Virus (z.B. Ko-Rezeptor-Gene) oder die antivirale Therapie (z.B. MDR1) einwirken oder aber die Ausstattung des Immunsystems betreffen (z.B. HLA).

Im Rahmen eines Pilotprojektes konnte an den Beispielen von CCR5 und des Nachweis der Position 3435 im MDR1-Gen gezeigt werden, dass es generell möglich ist, solche Positionen in humaner, genomischer DNA mittels APEX zu detektieren (6.11). Da das menschliche Genom im Vergleich zu den viralen Sequenzen wenig Variabilität aufweist, ist der Aufwand, der für die Etablierung der Sonden betrieben werden muss, ungleich geringer. Schwieriger könnte sich allerdings im Rahmen einer Validierung die Beschaffung geeigneten Referenzmaterials mit entsprechenden Genotypen für eine Validierung erweisen.

Die Literatur beschreibt eine Vielzahl solcher Gene, die mit dem Krankheitsverlauf in Verbindung gebracht werden. Ihre parallele Untersuchung auf dem gleichen Microarray, die mit geringem Mehraufwand durchgeführt werden kann, könnte daher künftig eine sinnvolle Ergänzung zur HIV-Genotypisierung sein.

8 Literaturverzeichnis

- Al Jabri, A.A.** (2001). HLA and in vitro susceptibility to HIV infection. *Mol Imm* **38**, 959-967.
- Arts, E.J., und Wainberg, M.A.** (1996). Mechanisms of nucleoside analog antiviral activity and resistance during human immunodeficiency virus reverse transcription. *Antimicrob Agents Chemother* **40**, 527-540.
- Ay, A.** (2006). Etablierung eines Verfahrens zur Auswertung für die Microarray-basierte Genotypisierung von HIV. Diplomarbeit an der Fachhochschule Bingen
- Bauer, A.** (2004). Genexpressionsanalysen in Karzinomen des Pankreas mittels DNA-Microarray-Technologie. Dissertation an der Universität Heidelberg
- Beck, I.A., Mahalanabis, M., Pepper, G., Wright, A., Hamilton, S., Langston, E., und Frenkel, L.M.** (2002). Rapid and sensitive oligonucleotide ligation assay for detection of mutations in human immunodeficiency virus type 1 associated with high-level resistance to protease inhibitors. *J Clin Microbiol* **40**, 1413-1419.
- Beerenwinkel, N., Lengauer, T., Daumer, M., Kaiser, R., Walter, H., Korn, K., Hoffmann, D., und Selbig, J.** (2003). Methods for optimizing antiviral combination therapies. *Bioinformatics* **19 Suppl 1**, i16-25.
- Beerenwinkel, N., Daumer, M., Oette, M., Korn, K., Hoffmann, D., Kaiser, R., Lengauer, T., Selbig, J., und Walter, H.** (2003). Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes. *Nucleic Acids Res* **31**, 3850-3855.
- Boden, D., Hurley, A., Zhang, L., Cao, Y., Guo, Y., Jones, E., Tsay, J., Ip, J., Farthing, C., Limoli, K., Parkin, N., und Markowitz, M.** (1999). HIV-1 drug resistance in newly infected individuals. *Jama* **282**, 1135-1141.
- Boucher, C.A., Cammack, N., Schipper, P., Schuurman, R., Rouse, P., Wainberg, M.A., und Cameron, J.M.** (1993). High-level resistance to (-) enantiomeric 2'-deoxy-3'-thiacytidine in vitro is due to one amino acid substitution in the catalytic site of human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother* **37**, 2231-2234.
- Brinkman, K., Smeitink, J.A., Romijn, J.A., und Reiss, P.** (1999). Mitochondrial toxicity induced by nucleoside-analogue reverse-transcriptase inhibitors is a key factor in the pathogenesis of antiretroviral-therapy-related lipodystrophy. *Lancet* **354**, 1112-1115.
- Cameron, D.W., Heath-Chiozzi, M., Danner, S., Cohen, C., Kravcik, S., Maurath, C., Sun, E., Henry, D., Rode, R., Potthoff, A., und Leonard, J.** (1998). Randomised placebo-controlled trial of ritonavir in advanced HIV-1 disease. The Advanced HIV Disease Ritonavir Study Group. *Lancet* **351**, 543-549.
- Cane, P.A.** (2005). Stability of transmitted drug-resistant HIV-1 species. *Curr Opin Infect Dis* **18**, 537-542.

- Carrington, M., Dean, M., Martin, M.P., und O'Brien, S.J.** (1999). Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences. *Hum Mol Genet* **8**, 1939-1945.
- Carrington, M., Nelson, G.W., Martin, M.P., Kissner, T., Vlahov, D., Goedert, J.J., Kaslow, R., Buchbinder, S., Hoots, K., und O'Brien, S.J.** (1999). HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283**, 1748-1752.
- Carrington, M., Nelson, G., und O'Brien, S.J.** (2001). Considering genetic profiles in functional studies of immune responsiveness to HIV-1. *Immunol Lett* **79**, 131-140.
- Carrington, M., und O'Brien, S.J.** (2003). The influence of HLA genotype on AIDS. *Annu Rev Med* **54**, 535-551.
- Chang, Y.C., und Tying, S.K.** (2004). Therapy of HIV infection. *Dermatol Ther* **17**, 449-464.
- Chen, J., Iannone, M.A., Li, M.S., Taylor, J.D., Rivers, P., Nelsen, A.J., Slentz-Kesler, K.A., Roses, A., und Weiner, M.P.** (2000). A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. *Genome Res* **10**, 549-557.
- Chizhikov, V., Wagner, M., Ivshina, A., Hoshino, Y., Kapikian, A.Z., und Chumakov, K.** (2002). Detection and genotyping of human group A rotaviruses by oligonucleotide microarray hybridization. *J Clin Microbiol* **40**, 2398-2407.
- Clavel, F., Race, E., und Mammano, F.** (2000). HIV drug resistance and viral fitness. *Adv Pharmacol* **49**, 41-66.
- Condra, J.H., Schleif, W.A., Blahy, O.M., Gabryelski, L.J., Graham, D.J., Quintero, J.C., Rhodes, A., Robbins, H.L., Roth, E., Shivaprakash, M., und et al.** (1995). In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374**, 569-571.
- Condra, J.H., Holder, D.J., Schleif, W.A., Blahy, O.M., Danovich, R.M., Gabryelski, L.J., Graham, D.J., Laird, D., Quintero, J.C., Rhodes, A., Robbins, H.L., Roth, E., Shivaprakash, M., Yang, T., Chodakewitz, J.A., Deutsch, P.J., Leavitt, R.Y., Massari, F.E., Mellors, J.W., Squires, K.E., Steigbigel, R.T., Teppler, H., und Emini, E.A.** (1996). Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor. *J Virol* **70**, 8270-8276.
- Condra, J.H.** (1998). Resistance to HIV protease inhibitors. *Haemophilia* **4**, 610-615.
- Corzillius, M., Muhlberger, N., Sroczynski, G., Jaeger, H., Wasem, J., und Siebert, U.** (2004). Cost effectiveness analysis of routine use of genotypic antiretroviral resistance testing after failure of antiretroviral treatment for HIV. *Antivir Ther* **9**, 27-36.
- D'Aquila, R.T.** (1995). HIV-1 chemotherapy and drug resistance. *Clin Diagn Virol* **3**, 299-316.

- De Mendoza, D., Gallego, O., und Soriano, V.** (2002). Mechanisms of resistance to antiretroviral drugs--clinical implications. *AIDS Rev* **4**, 64-82.
- Den Uyl, C., van der Horst-Bruinsma, I.E., und van Agtmael, M.** (2004). Progression of HIV to AIDS: a protective role for HLA-B27?. *AIDS Rev* **6**, 89-96.
- De Oliveira, T., Deforche, K., Cassol, S., Salminen, M., Paraskevis, D., Seebregts, C., Snoeck, J., van Rensburg, E.J., Wensing, A.M., van de Vijver, D.A., Boucher, C.A., Camacho, R., und Vandamme, A.M.** (2005). An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* **21**, 3797-3800.
- Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Smith, M.W., Allikmets, R., Goedert, J.J., Buchbinder, S.P., Vittinghoff, E., Gomperts, E., Donfield, S., Vlahov, D., Kaslow, R., Saah, A., Rinaldo, C., Detels, R., und O'Brien, S.J.** (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CKR5* structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* **273**, 1856-1862.
- Demeter, L.M., D'Aquila, R., Weislow, O., Lorenzo, E., Erice, A., Fitzgibbon, J., Shafer, R., Richman, D., Howard, T.M., Zhao, Y., Fisher, E., Huang, D., Mayers, D., Sylvester, S., Arens, M., Sannerud, K., Rasheed, S., Johnson, V., Kuritzkes, D., Reichelderfer, P., und Japour, A.** (1998). Interlaboratory concordance of DNA sequence analysis to detect reverse transcriptase mutations in HIV-1 proviral DNA. ACTG Sequencing Working Group. *AIDS Clinical Trials Group. J Virol Methods* **75**, 93-104.
- Demeter, L., und Haubrich, R.** (2001). International perspectives on antiretroviral resistance. Phenotypic and genotypic resistance assays: methodology, reliability, and interpretations. *J Acquir Immune Defic Syndr* **26 Suppl 1**, S3-9.
- Dovas, C.I., Efthimiou, K., und Katis, N.I.** (2004). Generic detection and differentiation of tobamoviruses by a spot nested RT-PCR-RFLP using dI-containing primers along with homologous dG-containing primers. *J Virol Methods* **117**, 137-144.
- Durant, J., Clevenbergh, P., Halfon, P., Delgiudice, P., Porsin, S., Simonet, P., Montagne, N., Boucher, C.A., Schapiro, J.M., und Dellamonica, P.** (1999). Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet* **353**, 2195-2199.
- Duwe, S., Brunn, M., Altmann, D., Hamouda, O., Schmidt, B., Walter, H., Pauli, G., und Kucherer, C.** (2001). Frequency of genotypic and phenotypic drug-resistant HIV-1 among therapy-naive patients of the German Seroconverter Study. *J Acquir Immune Defic Syndr* **26**, 266-273.
- Edelstein, R.E., Nickerson, D.A., Tobe, V.O., Manns-Arcuino, L.A., und Frenkel, L.M.** (1998). Oligonucleotide ligation assay for detecting mutations in the human immunodeficiency virus type 1 *pol* gene that are associated with resistance to zidovudine, didanosine, and lamivudine. *J Clin Microbiol* **36**, 569-572.

- Egger, M., Hirschel, B., Francioli, P., Sudre, P., Wirz, M., Flepp, M., Rickenbach, M., Malinverni, R., Vernazza, P., und Battegay, M.** (1997). Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. *Swiss HIV Cohort Study. Bmj* **315**, 1194-1199.
- Egger, M., May, M., Chene, G., Phillips, A.N., Ledergerber, B., Dabis, F., Costagliola, D., D'Arminio Monforte, A., de Wolf, F., Reiss, P., Lundgren, J.D., Justice, A.C., Staszewski, S., Leport, C., Hogg, R.S., Sabin, C.A., Gill, M.J., Salzberger, B., und Sterne, J.A.** (2002). Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet* **360**, 119-129.
- Erali, M., Page, S., Reimer, L.G., und Hillyard, D.R.** (2001). Human immunodeficiency virus type 1 drug resistance testing: a comparison of three sequence-based methods. *J Clin Microbiol* **39**, 2157-2165.
- Erdogan, F., Kirchner, R., Mann, W., Ropers, H.H., und Nuber, U.A.** (2001). Detection of mitochondrial single nucleotide polymorphisms using a primer elongation reaction on oligonucleotide microarrays. *Nucleic Acids Res* **29**, E36.
- Ermantraut, E.** (1999). Herstellung eines biomolekularen Arrays - eine technologische Herausforderung. *Medizinische Genetik* **11**, 6-11.
- Eshleman, S.H., Hackett, J., Jr., Swanson, P., Cunningham, S.P., Drews, B., Brennan, C., Devare, S.G., Zekeng, L., Kaptue, L., und Marlowe, N.** (2004). Performance of the Celera Diagnostics ViroSeq HIV-1 Genotyping System for sequence-based analysis of diverse human immunodeficiency virus type 1 strains. *J Clin Microbiol* **42**, 2711-2717.
- Fellay, J., Marzolini, C., Meaden, E.R., Back, D.J., Buclin, T., Chave, J.P., Decosterd, L.A., Furrer, H., Opravil, M., Pantaleo, G., Retelska, D., Ruiz, L., Schinkel, A.H., Vernazza, P., Eap, C.B., und Telenti, A.** (2002). Response to antiretroviral treatment in HIV-1-infected individuals with allelic variants of the multidrug resistance transporter 1: a pharmacogenetics study. *Lancet* **359**, 30-36.
- Fellenberg, K., Hauser, N.C., Brors, B., Neutzner, A., Hoheisel, J.D., und Vingron, M.** (2001). Correspondence analysis applied to microarray data. *Proc Natl Acad Sci U S A* **98**, 10781-10786.
- Garcia-Lerma, J.G., und Heneine, W.** (2001). Resistance of human immunodeficiency virus type 1 to reverse transcriptase and protease inhibitors: genotypic and phenotypic testing. *J Clin Virol* **21**, 197-212.
- Gonzalez, R., Masquelier, B., Fleury, H., Lacroix, B., Troesch, A., Vernet, G., und Telles, J.N.** (2004). Detection of human immunodeficiency virus type 1 antiretroviral resistance mutations by high-density DNA probe arrays. *J Clin Microbiol* **42**, 2907-2912.
- Goudsmit, J., De Ronde, A., Ho, D.D., und Perelson, A.S.** (1996). Human immunodeficiency virus fitness in vivo: calculations based on a single zidovudine resistance mutation at codon 215 of reverse transcriptase. *J Virol* **70**, 5662-5664.

- Greenberg, M.L., und Cammack, N.** (2004). Resistance to enfuvirtide, the first HIV fusion inhibitor. *J Antimicrob Chemother* **54**, 333-340.
- Gut, I.G.** (2001). Automation in genotyping of single nucleotide polymorphisms. *Hum Mutat* **17**, 475-492.
- Hacia, J.G., und Collins, F.S.** (1999). Mutational analysis using oligonucleotide microarrays. *J Med Genet* **36**, 730-736.
- Hacia, J.G.** (1999). Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet* **21**, 42-47.
- Hammer, S.M., Squires, K.E., Hughes, M.D., Grimes, J.M., Demeter, L.M., Currier, J.S., Eron, J.J., Jr., Feinberg, J.E., Balfour, H.H., Jr., Deyton, L.R., Chodakewitz, J.A., und Fischl, M.A.** (1997). A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* **337**, 725-733.
- Haubrich, R.H.** (2004). Resistance and replication capacity assays: clinical utility and interpretation. *Top HIV Med* **12**, 52-56.
- Hertogs, K., de Bethune, M.P., Miller, V., Ivens, T., Schel, P., Van Cauwenberge, A., Van Den Eynde, C., Van Gerwen, V., Azijn, H., Van Houtte, M., Peeters, F., Staszewski, S., Conant, M., Bloor, S., Kemp, S., Larder, B., und Pauwels, R.** (1998). A rapid method for simultaneous detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. *Antimicrob Agents Chemother* **42**, 269-276.
- Hill, F., Loakes, D., und Brown, D.M.** (1998). Polymerase recognition of synthetic oligodeoxyribonucleotides incorporating degenerate pyrimidine and purine bases. *Proc Natl Acad Sci U S A* **95**, 4258-4263.
- Hirsch, M.S., Conway, B., D'Aquila, R.T., Johnson, V.A., Brun-Vezinet, F., Clotet, B., Demeter, L.M., Hammer, S.M., Jacobsen, D.M., Kuritzkes, D.R., Loveday, C., Mellors, J.W., Vella, S., und Richman, D.D.** (1998). Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. International AIDS Society--USA Panel. *Jama* **279**, 1984-1991.
- Huang, Z., und Buckwold, V.E.** (2005). A TaqMan PCR assay using degenerate primers for the quantitative detection of woodchuck hepatitis virus DNA of multiple genotypes. *Mol Cell Probes* **19**, 282-289.
- Huang, D.D., Bremer, J.W., Brambilla, D.J., Palumbo, P.E., Aldrovandi, G., Eshleman, S., Brown, C., Fiscus, S., Frenkel, L., Hamdan, H., Hart, S., Kovacs, A., Krogstad, P., LaRussa, P., Sullivan, J., Weinberg, A., und Zhao, Y.Q.** (2005). Model for assessment of proficiency of human immunodeficiency virus type 1 sequencing-based genotypic antiretroviral assays. *J Clin Microbiol* **43**, 3963-3970.

- Huang, X., und Miller, W.** (1991). A time-efficient, linear-space local similarity algorithm. *Adv Appl Math* **12**, 33-357.
- Kaplan, A.H., Michael, S.F., Wehbie, R.S., Knigge, M.F., Paul, D.A., Everitt, L., Kempf, D.J., Norbeck, D.W., Erickson, J.W., und Swanstrom, R.** (1994). Selection of multiple human immunodeficiency virus type 1 variants that encode viral proteases with decreased sensitivity to an inhibitor of the viral protease. *Proc Natl Acad Sci U S A* **91**, 5597-5601.
- Kilpatrick, D.R., Nottay, B., Yang, C.F., Yang, S.J., Mulders, M.N., Holloway, B.P., Pallansch, M.A., und Kew, O.M.** (1996). Group-specific identification of polioviruses by PCR using primers containing mixed-base or deoxyinosine residue at positions of codon degeneracy. *J Clin Microbiol* **34**, 2990-2996.
- Korn, K., Reil, H., Walter, H., und Schmidt, B.** (2003). Quality control trial for human immunodeficiency virus type 1 drug resistance testing using clinical samples reveals problems with detecting minority species and interpretation of test results. *J Clin Microbiol* **41**, 3559-3565.
- Kozal, M.J., Shah, N., Shen, N., Yang, R., Fucini, R., Merigan, T.C., Richman, D.D., Morris, D., Hubbell, E., Chee, M., und Gingeras, T.R.** (1996). Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat Med* **2**, 753-759.
- Kurg, A., Tonisson, N., Georgiou, I., Shumaker, J., Tollett, J., und Metspalu, A.** (2000). Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. *Genet Test* **4**, 1-7.
- Lalezari, J.P., Henry, K., O'Hearn, M., Montaner, J.S., Piliero, P.J., Trottier, B., Walmsley, S., Cohen, C., Kuritzkes, D.R., Eron, J.J., Jr., Chung, J., DeMasi, R., Donatucci, L., Drobnes, C., Delehanty, J., und Salgo, M.** (2003). Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America. *N Engl J Med* **348**, 2175-2185.
- Landegren, U., Kaiser, R., Sanders, J., und Hood, L.** (1988). A ligase-mediated gene detection technique. *Science* **241**, 1077-1080.
- Larder, B.A., und Kemp, S.D.** (1989). Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* **246**, 1155-1158.
- Lazzarin, A.** (2005). Enfuvirtide: the first HIV fusion inhibitor. *Expert Opin Pharmacother* **6**, 453-464.
- Libert, F., Cochaux, P., Beckman, G., Samson, M., Aksenova, M., Cao, A., Czeizel, A., Claustres, M., de la Rua, C., Ferrari, M., Ferrec, C., Glover, G., Grinde, B., Guran, S., Kucinkas, V., Lavinha, J., Mercier, B., Ogur, G., Peltonen, L., Rosatelli, C., Schwartz, M., Spitsyn, V., Timar, L., Beckman, L., Parmentier, M., und Vassart, G.** (1998). The *deltacr5* mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum Mol Genet* **7**, 399-406.

- Linhart, C., und Shamir, R.** (2002). The degenerate primer design problem. *Bioinformatics* **18 Suppl 1**, S172-181.
- Lipshutz, R.J., Morris, D., Chee, M., Hubbell, E., Kozal, M.J., Shah, N., Shen, N., Yang, R., und Fodor, S.P.** (1995). Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* **19**, 442-447.
- Little, S.J., Daar, E.S., D'Aquila, R.T., Keiser, P.H., Connick, E., Whitcomb, J.M., Hellmann, N.S., Petropoulos, C.J., Sutton, L., Pitt, J.A., Rosenberg, E.S., Koup, R.A., Walker, B.D., und Richman, D.D.** (1999). Reduced antiretroviral drug susceptibility among patients with primary HIV infection. *Jama* **282**, 1142-1149.
- Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A., und Landau, N.R.** (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367-377.
- Lovmar, L., Fock, C., Espinoza, F., Bucardo, F., Syvanen, A.C., und Bondeson, K.** (2003). Microarrays for genotyping human group A rotavirus by multiplex capture and type-specific primer extension. *J Clin Microbiol* **41**, 5153-5158.
- Mammano, F., Trouplin, V., Zennou, V., und Clavel, F.** (2000). Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drug. *J Virol* **74**, 8524-8531.
- Martin, M.P., Gao, X., Lee, J.H., Nelson, G.W., Detels, R., Goedert, J.J., Buchbinder, S., Hoots, K., Vlahov, D., Trowsdale, J., Wilson, M., O'Brien, S.J., und Carrington, M.** (2002). Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet* **31**, 429-434.
- Mellors, J.W., Munoz, A., Giorgi, J.V., Margolick, J.B., Tassoni, C.J., Gupta, P., Kingsley, L.A., Todd, J.A., Saah, A.J., Detels, R., Phair, J.P., und Rinaldo, C.R., Jr.** (1997). Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann Intern Med* **126**, 946-954.
- Menendez-Arias, L.** (2002). Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends Pharmacol Sci* **23**, 381-388.
- Meyer, P.R., Matsuura, S.E., Schinazi, R.F., So, A.G., und Scott, W.A.** (2000). Differential removal of thymidine nucleotide analogues from blocked DNA chains by human immunodeficiency virus reverse transcriptase in the presence of physiological concentrations of 2'-deoxynucleoside triphosphates. *Antimicrob Agents Chemother* **44**, 3465-3472.
- Moutouh, L., Corbeil, J., und Richman, D.D.** (1996). Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc Natl Acad Sci U S A* **93**, 6106-6111.
- Nelson, M.L.** (1999). Host genetic influences on HIV-1 pathogenesis. *Curr Opin Immunol* **11**, 466-474.

- Nielsen, P.E., Egholm, M., Berg, R.H., und Buchardt, O.** (1991). Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254**, 1497-1500.
- Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., und Boucher, C.A.** (1999). Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. *Aids* **13**, 2349-2359.
- Noordewier, M.O., und Warren, P.V.** (2001). Gene expression microarrays and the integration of biological knowledge. *Trends Biotechnol* **19**, 412-415.
- O'Brien, S.J., und Nelson, G.W.** (2004). Human genes that limit AIDS. *Nat Genet* **36**, 565-574.
- Palella, F.J., Jr., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J., und Holmberg, S.D.** (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* **338**, 853-860.
- Palmer, S., Kearney, M., Maldarelli, F., Halvas, E.K., Bixby, C.J., Bazmi, H., Rock, D., Falloon, J., Davey, R.T., Jr., Dewar, R.L., Metcalf, J.A., Hammer, S., Mellors, J.W., und Coffin, J.M.** (2005). Multiple, Linked Immunodeficiency Virus Type 1 Drug Resistance Mutations in Treatment-Experienced Patients Are Missed by Standard Genotype Analysis. *J Clin Microbiol* **43**, 406-413.
- Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., und Syvanen, A.C.** (1997). Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res* **7**, 606-614.
- Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., und Ho, D.D.** (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582-1586.
- Petropoulos, C.J., Parkin, N.T., Limoli, K.L., Lie, Y.S., Wrin, T., Huang, W., Tian, H., Smith, D., Winslow, G.A., Capon, D.J., und Whitcomb, J.M.** (2000). A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother* **44**, 920-928.
- Preston, B.D., Poiesz, B.J., und Loeb, L.A.** (1988). Fidelity of HIV-1 reverse transcriptase. *Science* **242**, 1168-1171.
- Raboud, J.M., Montaner, J.S., Conway, B., Rae, S., Reiss, P., Vella, S., Cooper, D., Lange, J., Harris, M., Wainberg, M.A., Robinson, P., Myers, M., und Hall, D.** (1998). Suppression of plasma viral load below 20 copies/ml is required to achieve a long-term response to therapy. *Aids* **12**, 1619-1624.
- Richman, D.D.** (2004). Benefits and limitations of testing for resistance to HIV drugs. *J Antimicrob Chemother* **53**, 555-557.

- Roberts, J.D., Bebenek, K., und Kunkel, T.A.** (1988). The accuracy of reverse transcriptase from HIV-1. *Science* **242**, 1171-1173.
- Sanger, F., Nicklen, S., und Coulson, A.R.** (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.
- Sarmati, L., Nicastrì, E., Parisi, S.G., d'Ettore, G., Mancino, G., Narciso, P., Vullo, V., und Andreoni, M.** (2002). Discordance between genotypic and phenotypic drug resistance profiles in human immunodeficiency virus type 1 strains isolated from peripheral blood mononuclear cells. *J Clin Microbiol* **40**, 335-340.
- Shafer, R.W.** (2002). Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin Microbiol Rev* **15**, 247-277.
- Shumaker, J.M., Tollet, J.J., Filbin, K.J., Montague-Smith, M.P., und Pirrung, M.C.** (2001). APEX disease gene resequencing: mutations in exon 7 of the p53 tumor suppressor gene. *Bioorg Med Chem* **9**, 2269-2278.
- Syvänen, A.C.** (1999). From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat* **13**, 1-10.
- Syvänen, A.C.** (2001). Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet* **2**, 930-942.
- Southern, E.M.** (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503-517.
- Steege, K., Demecheleer, E., De Cabooter, N., Nges, D., Temmerman, M., Ndumbe, P., Mandaliya, K., Plum, J., und Verhofstede, C.** (2006). A sensitive in-house RT-PCR genotyping system for combined detection of plasma HIV-1 and assessment of drug resistance. *J Virol Methods* **133**, 137-145.
- Sturmer, M., Berger, A., und Preiser, W.** (2004). HIV-1 genotyping: comparison of two commercially available assays. *Expert Rev Mol Diagn* **4**, 281-291.
- Sugiura, W., Shimada, K., Matsuda, M., Chiba, T., Myint, L., Okano, A., und Yamada, K.** (2003). Novel enzyme-linked minisequence assay for genotypic analysis of human immunodeficiency virus type 1 drug resistance. *J Clin Microbiol* **41**, 4971-4979.
- Tong, C.Y., Mullen, J., Kulasegaram, R., De Ruiter, A., O'Shea, S., und Chrystie, I.L.** (2005). Genotyping of B and non-B subtypes of human immunodeficiency virus type 1. *J Clin Microbiol* **43**, 4623-4627.
- Tonisson, N., Kurg, A., Kaasik, K., Lohmussaar, E., und Metspalu, A.** (2000). Unravelling genetic data by arrayed primer extension. *Clin Chem Lab Med* **38**, 165-170.
- Tsongalis, G.J., Gleeson, T., Rodina, M., Anamani, D., Ross, J., Joannis, I., Tanimoto, L., und Ziermann, R.** (2005). Comparative performance evaluation of the HIV-1 LiPA protease and reverse transcriptase resistance assay on clinical isolates. *J Clin Virol* **34**, 268-271.

- Vahey, M., Nau, M.E., Barrick, S., Cooley, J.D., Sawyer, R., Sleeker, A.A., Vickerman, P., Bloor, S., Larder, B., Michael, N.L., und Wegner, S.A.** (1999). Performance of the Affymetrix GeneChip HIV PRT 440 platform for antiretroviral drug resistance genotyping of human immunodeficiency virus type 1 clades and viral isolates with length polymorphisms. *J Clin Microbiol* **37**, 2533-2537.
- Vandamme, A.M., Sonnerborg, A., Ait-Khaled, M., Albert, J., Asjo, B., Bacheler, L., Banhegyi, D., Boucher, C., Brun-Vezinet, F., Camacho, R., Clevenbergh, P., Clumeck, N., Dedes, N., De Luca, A., Doerr, H.W., Faudon, J.L., Gatti, G., Gerstoft, J., Hall, W.W., Hatzakis, A., Hellmann, N., Horban, A., Lundgren, J.D., Kempf, D., Miller, M., Miller, V., Myers, T.W., Nielsen, C., Opravil, M., Palmisano, L., Perno, C.F., Phillips, A., Pillay, D., Pumarola, T., Ruiz, L., Salminen, M., Schapiro, J., Schmidt, B., Schmit, J.C., Schuurman, R., Shulse, E., Soriano, V., Staszewski, S., Vella, S., Youle, M., Ziermann, R., und Perrin, L.** (2004). Updated European recommendations for the clinical use of HIV drug resistance testing. *Antivir Ther* **9**, 829-848.
- Walley, D.C., Tripp, B.W., Song, Y.C., Walley, K.R., und Tebbutt, S.J.** (2006). MACGT: multi-dimensional automated clustering genotyping tool for analysis of microarray-based mini-sequencing data. *Bioinformatics* **22**, 1147-1149.
- Wallis, C.L., Mahomed, I., Morris, L., Chidarikire, T., Stevens, G., Rekhviashvili, N., und Stevens, W.** (2005). Evaluation of an oligonucleotide ligation assay for detection of mutations in HIV-1 subtype C individuals who have high level resistance to nucleoside reverse transcriptase inhibitors and non-nucleoside reverse transcriptase inhibitors. *J Virol Methods* **125**, 99-109.
- Wegner, S.A., Wallace, M.R., Aronson, N.E., Tasker, S.A., Blazes, D.L., Tamminga, C., Fraser, S., Dolan, M.J., Stephan, K.T., Michael, N.L., Jagodzinski, L.L., Vahey, M.T., Gilcrest, J.L., Tracy, L., Milazzo, M.J., Murphy, D.J., McKenna, P., Hertogs, K., Rinehart, A., Larder, B., und Birx, D.L.** (2004). Long-term efficacy of routine access to antiretroviral-resistance testing in HIV type 1-infected patients: results of the clinical efficacy of resistance testing trial. *Clin Infect Dis* **38**, 723-730.
- Weiler, J., Gausepohl, H., Hauser, N., Jensen, O.N., und Hoheisel, J.D.** (1997). Hybridisation based DNA screening on peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acids Res* **25**, 2792-2799.
- Weinstein, M.C., Goldie, S.J., Losina, E., Cohen, C.J., Baxter, J.D., Zhang, H., Kimmel, A.D., und Freedberg, K.A.** (2001). Use of genotypic resistance testing to guide hiv therapy: clinical impact and cost-effectiveness. *Ann Intern Med* **134**, 440-450.
- Wilson, J.W., Bean, P., Robins, T., Graziano, F., und Persing, D.H.** (2000). Comparative evaluation of three human immunodeficiency virus genotyping systems: the HIV-GenotypR method, the HIV PRT GeneChip assay, and the HIV-1 RT line probe assay. *J Clin Microbiol* **38**, 3022-3028.
- Wilson, J.W.** (2003). Update on antiretroviral drug resistance testing: combining laboratory technology with patient care. *AIDS Read* **13**, 25-30, 35-28.

Winters, M.A., und Merigan, T.C. (2005). Insertions in the human immunodeficiency virus type 1 protease and reverse transcriptase genes: clinical impact and molecular mechanisms. *Antimicrob Agents Chemother* **49**, 2575-2582.

Yerly, S., Kaiser, L., Race, E., Bru, J.P., Clavel, F., und Perrin, L. (1999). Transmission of antiretroviral-drug-resistant HIV-1 variants. *Lancet* **354**, 729-733.

9 Eigene Publikationen

Schanné, M., Bodem, J., Kräusslich, H.G., Jacob, A. & Hoheisel, J.D. (2005). Microarray-basierte Resistenzbestimmung in HIV. *LaborMedizin & Diagnostik* **01.05**, 19.

Brems, S., Guilbruide, L., Gundlesdodjir-Planck, D., Busold, C., Luu, V.-D., **Schanné, M.**, Hoheisel, J.D. & Clayton, C. (2005). The transcriptomes of *Trypanosoma brucei* Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. *Mol. Biochem. Parasit.* **139**, 163-172.

Beier, V., Mund, C., **Schanné, M.**, Lyko, F. & Hoheisel, J.D. (2005). Characterisation of genomic methylation patterns in tumours. *Exocrine Pancreas Cancer* (Gress, T.M., Neoptolemos, J.P., Lemoine, N.R. & Real, F.X., eds.), Solvay, 252-258.

10 Abkürzungen

A	Adenin
AS	Aminosäure
ATP	Adenosintri-phosphat
bp	Basenpaare
C	Cytosin
cDNA	copyDNA
CTP	Cytosintri-phosphat
Cy3	Cyanine3
Cy5	Cyanine5
dATP	Desoxyadenosintri-phosphat
dCTP	Desoxycytosintri-phosphat
ddATP	Didesoxyadenosintri-phosphat
ddCTP	Didesoxycytosintri-phosphat
ddGTP	Didesoxyguanosintri-phosphat
ddUTP	Didesoxyuridintri-phosphat
ddNTP	Didesoxynukleosidtri-phosphat
dest	destilliert
dGTP	Desoxyguanosintri-phosphat
DNA	Desoxyribonukleinsäure(-acid)
dNTP	Desoxynukleosidtri-phosphat
DTT	Dithiothreitol
dTTP	Desoxythymidintri-phosphat
dUTP	Desoxyuridintri-phosphat
EDTA	Ethylendiamintetraessigsäure
EtBr	Ethidiumbromid
EtOH	Ethanol
f	forward (bezüglich der Orientierung einer DNA-Sequenz)
G	Guanin
gp	Glykoprotein
GTP	Guanosintri-phosphat
IUB	International Union of Biochemistry
mRNA	Boten-RNA

NaOH	Natronlauge
OD	optische Dichte
Oligo(dT)-Spacer	Oligodesoxythymidinspacer
PCR	Polymerase-Kettenreaktion (Polymerase chain reaction)
PR	Protease
r	reverse (bezüglich der Orientierung einer DNA-Sequenz)
RNA	Ribonukleinsäure(-acid)
rpm	Umdrehungen pro Minute
RT	Reverse Transkriptase (je nach Kontext)
RT	Raumtemperatur (je nach Kontext)
SDS	Natrium-Dodecylsulfat
SSC	Saline-Sodium-Citrat
T	Thymin
TAE	Tris-Acetat-EDTA
TE	Tris-Acetat
T _m	Schmelztemperatur (melting temperature)
Tris	Tris(Hydroxymethyl)aminomethan
U	Uracil
UNG	Uracil-N-Glykosylase
UTP	Uridintriphosphat
UV	Ultraviolett
v/v	Volumenprozent
WT	Wildtyp
w/v	Gewichtsprozent

Außerdem wurden die üblichen Abkürzungen physikalischer und chemischer Einheiten verwendet.

11 Anhang

Alle aufgeführten Oligonukleotide haben, soweit nicht anders gekennzeichnet, einen 5'-C6-Aminolink und darauf folgenden 5'-Oligod(T)-Spacer, so dass die Gesamtlänge jeweils 50 Basen beträgt.

Oligonukleotide für die APEX-Optimierung

Name	Sequenz 5'-3'	Länge	5'-Aminolink
PR_AS60_r1_60_C	ATCATCTGCTCCTGTATCTAATAGAGCTTC	30	
PR_AS60_r1a1_60_C	ATCATCTGCTCCTGTATCTAATAGAGCTTC	30	x
PR_AS60_r2_60_C	CTGCTCCTGTATCTAATAGAGCTTC	25	
PR_AS60_r2a1_60_C	CTGCTCCTGTATCTAATAGAGCTTC	25	x
PR_AS60_r3_60_C	TGCTCCTGTATCTAATAGAGCTTC	24	
PR_AS60_r3a1_60_C	TGCTCCTGTATCTAATAGAGCTTC	24	x
PR_AS60_r4_60_C	GCTCCTGTATCTAATAGAGCTTC	23	
PR_AS60_r4a1_60_C	GCTCCTGTATCTAATAGAGCTTC	23	x
PR_AS60_r5_60_C	CTCCTGTATCTAATAGAGCTTC	22	
PR_AS60_r5a1_60_C	CTCCTGTATCTAATAGAGCTTC	22	x
PR_AS60_r6_60_C	TCCTGTATCTAATAGAGCTTC	21	
PR_AS60_r6a1_60_C	TCCTGTATCTAATAGAGCTTC	21	x
PR_AS60_r7_60_C	CCTGTATCTAATAGAGCTTC	20	
PR_AS60_r7a1_60_C	CCTGTATCTAATAGAGCTTC	20	x

Oligonukleotide für den Protease-Chip

Die Position des Oligonukleotids setzt sich zusammen aus zwei Spalten (für das Triplet und für das Nukleotid innerhalb dieses Triplets). Wie im Ergebnisteil beschrieben, wurden zwei verschiedene SONDENSÄTZE zur Herstellung des PR-Microarrays verwendet (1. und 2. Satz).

Position	Name	Sequenz 5'-3'	Länge	1. Satz	2. Satz
10	28 PR_AS10_f1(20)_28_C	TCACTCTTTGGCAGCGACCC	20	x	
	PR_AS10_f1(30)_28_C	CCTCAGATCACTCTTTGGCAGCGACCC	27	x	x
	PR_AS10_f1_29_T	CACTCTTTGGCAGCGACCCC	20	x	x
29	PR_AS10_f2(28)_29_T	CCTCARATCACTCTTTGGCARCGACCCC	28	x	x
	PR_AS10_f3(28)_29_T	CCTCARATCACTCTTTGGCARCGACCCA	28	x	x
	PR_AS10_f4(28)_29_T	CCTCARATCACTCTTTGGCARCGACCCA	28	x	x
	PR_AS10_f5(28)_29_T	CCTCARATCACTCTTTGGCARCGACCCA	28	x	x
	PR_AS10_f1_30_C	ACTCTTTGGCAGCGACCCCT	20	x	x
30	PR_AS10_f2(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCCT	29	x	x
	PR_AS10_f3(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCAT	29	x	x
	PR_AS10_f4(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCCT	29	x	x
	PR_AS10_f5(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCGT	29	x	x
	PR_AS10_f6(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCCG	29	x	x
	PR_AS10_f7(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCAG	29	x	x
	PR_AS10_f8(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCCTG	29	x	x
	PR_AS10_f9(29)_30_C	CCTCARATCACTCTTTGGCARCGACCCGG	29	x	x
	20	58 PR_AS20_r1(30)_58_T	CATCTGCTCCTGTATCTAATAGAGCTTCCT	30	x

		PR_AS20_r2(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCCT	30	x	x
		PR_AS20_r3(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCCA	30	x	x
		PR_AS20_r4(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCCC	30	x	x
		PR_AS20_r5(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCTT	30	x	x
		PR_AS20_r6(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCTA	30	x	x
		PR_AS20_r7(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCTC	30	x	x
		PR_AS20_r8(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCCG	30	x	x
		PR_AS20_r9(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCTG	30	x	x
		PR_AS20_r10(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCAT	30	x	x
		PR_AS20_r11(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCAA	30	x	x
		PR_AS20_r12(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCAC	30	x	x
		PR_AS20_r13(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCAG	30	x	x
		PR_AS20_r14(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCGT	30	x	x
		PR_AS20_r15(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCGA	30	x	x
		PR_AS20_r16(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCGG	30	x	x
		PR_AS20_r17(30)_58_T	CATCTGCTCCTGTRTCTAATAGAGCTTCGC	30	x	x
59		PR_AS20_r1(20)_59_T	CTGTATCTAATAGAGCTTCC	20	x	x
		PR_AS20_r1(30)_59_T	TCATCTGCTCCTGTATCTAATAGAGCTTCC	30	x	x
		PR_AS20_r2(30)_59_T	TCATCTGCTCCTGTRTCTAATAGAGCTTCC	30	x	x
		PR_AS20_r3(30)_59_T	TCATCTGCTCCTGTRTCTAATAGAGCTTCT	30	x	x
		PR_AS20_r4(30)_59_T	TCATCTGCTCCTGTRTCTAATAGAGCTTCA	30	x	x
		PR_AS20_r5(30)_59_T	TCATCTGCTCCTGTRTCTAATAGAGCTTCG	30	x	x
60		PR_AS20_r1(20)_60_C	CCTGTATCTAATAGAGCTTC	20	x	x
		PR_AS20_r1(30)_60_C	ATCATCTGCTCCTGTATCTAATAGAGCTTC	30	x	x
		PR_AS20_r2(30)_60_C	ATCATCTGCTCCTGTRTCTAATAGAGCTTC	30	x	x
23	67	PR_AS23_r1_67_G	ATACTGTATCATCTGCTCCTGTATCTAATA	30	x	
		PR_AS23_r2_67_G	ATACTGTATCATCTGCTCCTGTGTCTAATA	30	x	x
		PR_AS23_f1_67_C	ATAAAGATAGGGGGGCARITAAAGGAAGCT	30		x
		PR_AS23_f2_67_C	ATAAAGATAGGGGGGCARITAAAGAAGCT	30		x
		PR_AS23_f3_67_C	ATAAAGATAGGGGGGCARITAAATGGAAGCT	30		x
		PR_AS23_f4_67_C	ATAAAGATAGGGGGGCARITAAATAGAAGCT	30		x
		PR_AS23_f5_67_C	ATAAAGATAGGGGGGCARITAAAGGAAGCT	30		x
		PR_AS23_f6_67_C	ATAAAGATAGGGGGGCARITAAAGAGAAGCT	30		x
	68	PR_AS23_r1_68_A	AATACTGTATCATCTGCTCCTGTATCTAAT	30	x	x
		PR_AS23_r2_68_A	AATACTGTATCATCTGCTCCTGTGTCTAAT	30	x	x
	69	PR_AS23_r1_69_T	TAATACTGTATCATCTGCTCCTGTATCTAA	30	x	x
		PR_AS23_r2_69_T	TAATACTGTATCATCTGCTCCTGTGTCTAA	30	x	x
24	70	PR_AS24_r_70_A	ATCATCTGCTCCTGTATCTA	20	x	x
		PR_AS24_r2(29)_70_A	CTAATACTGTATCATCTGCTCTGTATCTA	29	x	x
		PR_AS24_r3(29)_70_A	CTAATACTGTATCATCTGCTCTGTGTCTA	29	x	x
	71	PR_AS24_r_71_A	TATCATCTGCTCCTGTATCT	20	x	x
		PR_AS24_r2(30)_71_A	TCTAATACTGTATCATCTGCTCCTGTATCT	30	x	x
		PR_AS24_r3(30)_71_A	TCTAATACTGTATCATCTGCTCCTGTGTCT	30	x	x
	72	PR_AS24_r_72_T	GTATCATCTGCTCCTGTATC	20	x	x
		PR_AS24_r2_72_T	GTATCATCTGCTCCTGTGTCT	20	x	x
		PR_AS24_r2(30)_72_T	TTCTAATACTGTATCATCTGCTCCTGTATC	30	x	x
		PR_AS24_r3(30)_72_T	TTCTAATACTGTATCATCTGCTCCTGTGTCT	30	x	x
30	88	PR_AS30_f_88_G	TATTAGATACAGGAGCAGAT	20	x	x
		PR_AS30_f(30)_88_G	GAAGCTCTATTAGATACAGGAGCAGAT	27	x	x
		PR_AS30_f2(27)_88_G	GAAGCTCTATTAGAYACAGGAGCAGAT	27	x	x
		PR_AS30_r1_88_C	AAAYTCATITCTTCTAATACTGTAT	25		x
		PR_AS30_r2_88_C	AAAYTTATITCTTCTAATACTGTAT	25		x
		PR_AS30_r3_88_C	AAAYCCATITCTTCTAATACTGTAT	25		x
		PR_AS30_r4_88_C	AAAYCTATITCTTCTAATACTGTAT	25		x
	89	PR_AS30_f_89_A	ATTAGATACAGGAGCAGATG	20	x	x
		PR_AS30_f2(28)_89_A	GAAGCTCTATTAGAYACAGGAGCAGATG	28	x	x
		PR_AS30_f3(28)_89_A	GAAGCTCTATTAGAYACAGGAGCAGATA	28	x	x
	90	PR_AS30_f_90_T	TTAGATACAGGAGCAGATGA	20	x	x
		PR_AS30_f2(29)_90_T	GAAGCTCTATTAGAYACAGGAGCAGATGA	29	x	x
		PR_AS30_f3(29)_90_T	GAAGCTCTATTAGAYACAGGAGCAGATAA	29	x	x
32	94	PR_AS32_f_94_G	ATACAGGAGCAGATGATACA	20	x	x
		PR_AS32_f2(30)_94_G	GCTCTATTAGAYACAGGAGCAGATGATACA	30	x	x
	95	PR_AS32_f_95_T	TACAGGAGCAGATGATACAG	20	x	x
		PR_AS32_f2(30)_95_T	CTCTATTAGAYACAGGAGCAGATGATACAG	30	x	x
		PR_AS32_f3(30)_95_T	CTCTATTAGAYACAGGAGCAGATGATACAA	30	x	x
	96	PR_AS32_f_96_A	ACAGGAGCAGATGATACAGT	20	x	x
		PR_AS32_f2(30)_96_A	TCTATTAGAYACAGGAGCAGATGATACAGT	30	x	x

		PR_AS32_f3(30)_96_A	TCTATTAGAYACAGGAGCAGATGATACAAT	30	x	x
33	97	PR_AS33_f_97_T	CAGGAGCAGATGATACAGTA	20	x	x
		PR_AS33_f(30)_97_T	CTATTAGAYACAGGAGCAGATGATACAGTA	30	x	x
	98	PR_AS33_f_98_T	AGGAGCAGATGATACAGTAT	20	x	x
		PR_AS33_f(30)_98_T	TATTAGAYACAGGAGCAGATGATACAGTAT	30	x	
		PR_AS33_f2(30)_98_T	TATTAGAYACAGGAGCAGATGATACAGTAG	30	x	x
		PR_AS33_f3(30)_98_T	TATTAGAYACAGGAGCAGATGATACAGTAA	30	x	x
	99	PR_AS33_f_99_A	GGAGCAGATGATACAGTATT	20	x	x
		PR_AS33_f2(30)_99_A	ATTAGAYACAGGAGCAGATGATACAGTATT	30	x	x
		PR_AS33_f3(30)_99_A	ATTAGAYACAGGAGCAGATGATACAGTAGT	30	x	x
		PR_AS33_f4(30)_99_A	ATTAGAYACAGGAGCAGATGATACAGTAAT	30	x	x
36	106	PR_AS36_f_106_A	ATGATACAGTATTAGAAGAA	20	x	x
		PR_AS36_f2(30)_106_A	ACAGGAGCAGATGATACAGTATTAGAAGAA	30	x	x
		PR_AS36_f3(30)_106_A	ACAGGAGCAGATGATACAGTATTAGAAGAT	30	x	x
		PR_AS36_f4(30)_106_A	ACAGGAGCAGATGATACAGTATTAGAAGAC	30	x	x
	107	PR_AS36_f2(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGAAA	30	x	x
		PR_AS36_f3(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGATA	30	x	x
		PR_AS36_f4(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGACA	30	x	x
		PR_AS36_f5(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGAAG	30	x	x
		PR_AS36_f6(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGATG	30	x	x
		PR_AS36_f7(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGACG	30	x	x
		PR_AS36_f8(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGAAT	30	x	x
		PR_AS36_f9(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGATT	30	x	x
		PR_AS36_f10(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGACT	30	x	x
		PR_AS36_f11(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGAAC	30	x	x
		PR_AS36_f12(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGATC	30	x	x
		PR_AS36_f13(30)_107_T	CAGGAGCAGATGATACAGTATTAGAAGACC	30	x	x
	108	PR_AS36_f_108_G	GATACAGTATTAGAAGAAAT	20	x	x
		PR_AS36_f(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAAAT	30	x	x
		PR_AS36_f2(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATAT	30	x	x
		PR_AS36_f3(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACAT	30	x	x
		PR_AS36_f4(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAAGT	30	x	x
		PR_AS36_f5(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATGT	30	x	x
		PR_AS36_f6(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACGT	30	x	
		PR_AS36_f7(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAATT	30	x	x
		PR_AS36_f8(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATTT	30	x	x
		PR_AS36_f9(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACTT	30	x	x
		PR_AS36_f10(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAACT	30	x	x
		PR_AS36_f11(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATCT	30	x	x
		PR_AS36_f12(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACCT	30	x	x
		PR_AS36_f13(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAAAC	30	x	x
		PR_AS36_f14(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATAC	30	x	x
		PR_AS36_f15(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACAC	30	x	x
		PR_AS36_f16(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAAGC	30	x	x
		PR_AS36_f17(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATGC	30	x	x
		PR_AS36_f18(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACGC	30	x	x
		PR_AS36_f19(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAATC	30	x	x
		PR_AS36_f20(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATTC	30	x	x
		PR_AS36_f21(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACTC	30	x	x
		PR_AS36_f22(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGAACC	30	x	x
		PR_AS36_f23(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGATCC	30	x	x
		PR_AS36_f24(30)_108_G	AGGAGCAGATGATACAGTATTAGAAGACCC	30	x	x
		PR_AS36_r1_108_C	ATTTTTGGTTTTCCATYTYCCTGGCAAATT	29		x
		PR_AS36_r2_108_C	ATTTTTGGTTTTCCATYTYCCTGGCAAATC	29		x
		PR_AS36_r3_108_C	ATTTTTGGTTTTCCATYTYCCTGGCAAACT	29		x
		PR_AS36_r4_108_C	ATTTTTGGTTTTCCATYTYCCTGGCAAACC	29		x
		PR_AS36_r5_108_C	ATTTTTGGTTTTCCATYTYCCTGGTAAATT	29		x
		PR_AS36_r6_108_C	ATTTTTGGTTTTCCATYTYCCTGGTAAATC	29		x
		PR_AS36_r7_108_C	ATTTTTGGTTTTCCATYTYCCTGGTAAACT	29		x
		PR_AS36_r8_108_C	ATTTTTGGTTTTCCATYTYCCTGGTAAACC	29		x
46	136	PR_AS46_r_136_T	ACCTCCAATTCCTCCCTATCA	20	x	x
		PR_AS46_r2_136_T	ACCTCCAATTCCTCCCTATTA	20	x	x
		PR_AS46_r3_136_T	ACCTCCAATTCCTCCCTATAA	20	x	x
		PR_AS46_r4_136_T	ACCTCCAATTCCTCCCTATGA	20	x	x
	137	PR_AS46_r_137_A	AACCTCCAATTCCTCCCTATC	20	x	x
		PR_AS46_r2_137_A	AACCTCCAATTCCTCCCTATT	20	x	x
		PR_AS46_r3_137_A	AACCTCCAATTCCTCCCTATA	20	x	x
		PR_AS46_r4_137_A	AACCTCCAATTCCTCCCTATG	20	x	x

	138	PR_AS46_r_138_C	AAACCTCCAATTCCCCCTAT	20	x	x
47	139	PR_AS47_r_139_T	AAAACCTCCAATTCCCCCCA	20	x	x
		PR_AS47_r(30)_139_T	TTACTTTGAYAAAACCTCCAATTCCCCCCA	30	x	x
		PR_AS47_r2(30)_139_T	TTACTTTGAYAAAACCTCCAATTCCCCCTA	30	x	x
		PR_AS47_f1_139_A	TTTRCCAGGRAGATGGAAACCAAAAATG	28		x
		PR_AS47_f2_139_A	TTTRCCAGGRAGATGGAAACCAAAAATA	28		x
		PR_AS47_f3_139_A	TTTRCCAGGAAAATGGAAACCAAAAATG	28		x
		PR_AS47_f4_139_A	TTTRCCAGGAAAATGGAAACCAAAAATA	28		x
	140	PR_AS47_r_140_A	TAAAACCTCCAATTCCCCCT	20	x	x
		PR_AS47_r2_140_A	CAAAACCTCCAATTCCCCCT	20	x	x
		PR_AS47_r(29)_140_A	TTACTTTGAYAAAACCTCCAATTCCCCCC	29	x	x
		PR_AS47_r2(29)_140_A	TTACTTTGAYAAAACCTCCAATTCCCCCT	29	x	x
	141	PR_AS47_r3(28)_141_T	TTACTTTGAYAAAACCTCCAATTCCCCC	28	x	x
48	142	PR_AS48_r_142_C	GATAAAACCTCCAATTCCCC	20	x	x
		PR_AS48_r2_142_C	GACAAAACCTCCAATTCCCC	20	x	x
		PR_AS48_r2_142_C	GTYYTACTTTGAYAAAACCTCCAATTCCCA	30	x	x
	143	PR_AS48_r_143_C	TGATAAAACCTCCAATTCCC	20	x	x
		PR_AS48_r2(30)_143_C	TGTYTTACTTTGAYAAAACCTCCAATTCCC	30	x	x
	144	PR_AS48_r_144_C	TTGATAAAACCTCCAATTCC	20	x	x
		PR_AS48_r2(29)_144_C	TGTYTTACTTTGAYAAAACCTCCAATTCC	29	x	x
50	148	PR_AS50_f_148_A	AACCAAAAATGATAGGGGGA	20	x	x
		PR_AS50_f2(27)_148_A	ARATGGAAACCAAAAATGATAGGGGGA	27	x	x
		PR_AS50_f3(27)_148_A	ARATGGAAACCAAAAATAATAGGGGGA	27	x	x
	149	PR_AS50_f_149_T	ACCAAAAATGATAGGGGGAA	20	x	x
		PR_AS50_f2(28)_149_T	ARATGGAAACCAAAAATRATAGGGGGAA	28	x	x
		PR_AS50_f3(28)_149_T	ARATGGAAACCAAAAATRATAGGGGGAG	28	x	x
		PR_AS50_f4(28)_149_T	ARATGGAAACCAAAAATRATAGGGGGAC	28	x	x
	150	PR_AS50_f_150_T	CCAAAAATGATAGGGGGAAT	20	x	x
		PR_AS50_f2(29)_150_T	ARATGGAAACCAAAAATRATAGGGGGAAT	29	x	x
		PR_AS50_f3(29)_150_T	ARATGGAAACCAAAAATRATAGGGGGAGT	29	x	x
		PR_AS50_f4(29)_150_T	ARATGGAAACCAAAAATRATAGGGGGACT	29	x	x
53	157	PR_AS53_f_157_T	TGATAGGGGGAATTGGAGGT	20	x	x
		PR_AS53_f2(30)_157_T	AAACCAAAAATRATAGGGGGAATTGGAGGT	30	x	x
	158	PR_AS53_f_158_T	GATAGGGGGAATTGGAGGTT	20	x	x
		PR_AS53_f2(30)_158_T	AACCAAAAATRATAGGGGGAATTGGAGGTT	30	x	x
		PR_AS53_f3(30)_158_T	AACCAAAAATRATAGGGGGAATTGGAGGTC	30	x	x
		PR_AS53_f4(30)_158_T	AACCAAAAATRATAGGGGGAATTGGAGGTA	30	x	x
	159	PR_AS53_f_159_T	ATAGGGGGAATTGGAGGTTT	20	x	x
		PR_AS53_f2(30)_159_T	ACCAAAAATRATAGGGGGAATTGGAGGTTT	30	x	x
		PR_AS53_f3(30)_159_T	ACCAAAAATRATAGGGGGAATTGGAGGTTA	30	x	x
		PR_AS53_f4(30)_159_T	ACCAAAAATRATAGGGGGAATTGGAGGTCT	30	x	x
		PR_AS53_f5(30)_159_T	ACCAAAAATRATAGGGGGAATTGGAGGTCA	30	x	x
		PR_AS53_f6(30)_159_T	ACCAAAAATRATAGGGGGAATTGGAGGTAT	30	x	x
54	160	PR_AS54_f_160_A	TAGGGGGAATTGGAGGTTTT	20	x	x
		PR_AS54_f2(30)_160_A	CCAAAAATRATAGGGGGAATTGGAGGTTTT	30	x	x
	161	PR_AS54_f_161_T	AGGGGGAATTGGAGGTTTTA	20	x	x
		PR_AS54_f2_161_T	AGGGGGAATTGGAGGTTTTTG	20	x	x
		PR_AS54_f3(30)_161_T	CAAAAATRATAGGGGGAATTGGAGGTTTTTC	30	x	x
	162	PR_AS54_f_162_C	GGGGGAATTGGAGGTTTTAT	20	x	x
		PR_AS54_f2_162_C	GGGGGAATTGGAGGTTTTTGT	20	x	x
		PR_AS54_f3(30)_162_C	AAAAATRATAGGGGGAATTGGAGGTTTTTCT	30	x	x
		PR_AS54_f5(30)_162_C	AAAAATRATAGGGGGAATTGGAGGTTTTTGA	30	x	x
		PR_AS54_f6(30)_162_C	AAAAATRATAGGGGGAATTGGAGGTTTTTCA	30	x	x
60	178	PR_AS60_f_178_G	TTATCAAAGTAAAACAGTAT	20	x	x
		PR_AS60_f2(30)_178_G	ATTGGAGGTTTTTRTCAAAGTAARACAGTAT	30	x	x
		PR_AS60_f3(30)_178_G	ATTGGAGGTTTTTRTCAAAGTAARACAATAT	30	x	x
		PR_AS60_r1_178_C	CCGCAGATTTCTAYIAGTATCTGAT	25		x
		PR_AS60_r2_178_C	CCGCAGATTTCTAYIAGTATTTGAT	25		x
		PR_AS60_r3_178_C	CCGCAGATTTCTAYIAGTACCTGAT	25		x
		PR_AS60_r4_178_C	CCGCAGATTTCTAYIAGTACTTGAT	25		x
		PR_AS60_r5_178_C	CCGCAGATTTCTAYIGGTATCTGAT	25		x
		PR_AS60_r6_178_C	CCGCAGATTTCTAYIGGTATTTGAT	25		x
		PR_AS60_r7_178_C	CCGCAGATTTCTAYIGGTACTTGAT	25		x
		PR_AS60_r8_178_C	CCGCAGATTTCTAYIGGTACTTGAT	25		x
	179	PR_AS60_f_179_A	TATCAAAGTAAAACAGTATG	20	x	x

	PR_AS60_f2_179_A	TRTCAAAGTAARACAATATG	20	x	x
	PR_AS60_f3(30)_179_A	TTGGAGGTTTTTRTCAAAGTAARACAATATG	30	x	x
	PR_AS60_f4(30)_179_A	TTGGAGGTTTTTRTCAAAGTAARACAGTATG	30	x	x
180	PR_AS60_f_180_T	ATCAAAGTAAAAACAGTATGA	20	x	x
	PR_AS60_f2(30)_180_T	TGGAGGTTTTTRTCAAAGTAAGACAGTATGA	30	x	x
	PR_AS60_f3(30)_180_T	TGGAGGTTTTTRTCAAAGTAARACAGTATGA	30	x	x
	PR_AS60_f4(30)_180_T	TGGAGGTTTTTRTCAAAGTAARACAATATGA	30	x	x
62	184 PR_AS62_f1(30)_184_A	GGTTTTTRTCAAAGTAARACAGTATGATCAG	30	x	x
	PR_AS62_f2(30)_184_A	GGTTTTTRTCAAAGTAARACAGTATGATCAA	30	x	x
	PR_AS62_f3(30)_184_A	GGTTTTTRTCAAAGTAARACAATATGATCAG	30	x	x
	PR_AS62_f4(30)_184_A	GGTTTTTRTCAAAGTAARACAATATGATCAA	30	x	x
	PR_AS62_r1_184_T	CTTTATGTCCRCARATTTCTATGAGTA	27		x
	PR_AS62_r2_184_T	CTTTATGTCCRCARATTTCTATGAGTA	27		x
	PR_AS62_r3_184_T	CTTTATGTCCRCARATTTCTATAAGTA	27		x
	PR_AS62_r4_184_T	CTTTATGTCCRCARATTTCTATAAGTA	27		x
	PR_AS62_r5_184_T	CTTTATGTCCRCARATTTCTATTAGTA	27		x
	PR_AS62_r6_184_T	CTTTATGTCCRCARATTTCTATTAGTA	27		x
	PR_AS62_r7_184_T	CTTTATGTCCRCARATTTCTAAGAGTA	27		x
	PR_AS62_r8_184_T	CTTTATGTCCRCARATTTCTAAGAGTA	27		x
	PR_AS62_r9_184_T	CTTTATGTCCRCARATTTCTAAAAGTA	27		x
	PR_AS62_r10_184_T	CTTTATGTCCRCARATTTCTAAAAGTA	27		x
	PR_AS62_r11_184_T	CTTTATGTCCRCARATTTCTAATAGTA	27		x
	PR_AS62_r12_184_T	CTTTATGTCCRCARATTTCTAATAGTA	27		x
185	PR_AS62_f1(30)_185_T	GTTTTTRTCAAAGTAARACARTATGATCAGG	30	x	x
	PR_AS62_f2(30)_185_T	GTTTTTRTCAAAGTAARACARTATGATCAAA	30	x	x
	PR_AS62_f3(30)_185_T	GTTTTTRTCAAAGTAARACARTATGATCAGA	30	x	x
	PR_AS62_f4(30)_185_T	GTTTTTRTCAAAGTAARACARTATGATCAAG	30	x	x
	PR_AS62_r1_185_T	CTTTATGTCCRCARATTTCTATGAGT	26		x
	PR_AS62_r2_185_T	CTTTATGTCCRCARATTTCTATGAGT	26		x
	PR_AS62_r3_185_T	CTTTATGTCCRCARATTTCTATAAGT	26		x
	PR_AS62_r4_185_T	CTTTATGTCCRCARATTTCTATAAGT	26		x
	PR_AS62_r5_185_T	CTTTATGTCCRCARATTTCTATTAGT	26		x
	PR_AS62_r6_185_T	CTTTATGTCCRCARATTTCTATTAGT	26		x
	PR_AS62_r7_185_T	CTTTATGTCCRCARATTTCTAAGAGT	26		x
	PR_AS62_r8_185_T	CTTTATGTCCRCARATTTCTAAGAGT	26		x
	PR_AS62_r9_185_T	CTTTATGTCCRCARATTTCTAAAAGT	26		x
	PR_AS62_r10_185_T	CTTTATGTCCRCARATTTCTAAAAGT	26		x
	PR_AS62_r11_185_T	CTTTATGTCCRCARATTTCTAATAGT	26		x
	PR_AS62_r12_185_T	CTTTATGTCCRCARATTTCTAATAGT	26		x
186	PR_AS62_f1(30)_186_A	TTTTTRTCAAAGTAARACARTATGATCAGGT	30	x	x
	PR_AS62_f2(30)_186_A	TTTTTRTCAAAGTAARACARTATGATCAAAT	30	x	x
	PR_AS62_f3(30)_186_A	TTTTTRTCAAAGTAARACARTATGATCAGAT	30	x	x
	PR_AS62_f4(30)_186_A	TTTTTRTCAAAGTAARACARTATGATCAAGT	30	x	x
63	187 PR_AS63_f2(26)_187_C	TCAAAGTAARACARTATGATCAGATA	26	x	x
	PR_AS63_f3(26)_187_C	TCAAAGTAARACARTATGATCAGGTA	26	x	x
	PR_AS63_f4(26)_187_C	TCAAAGTAARACARTATGATCAAGTA	26	x	x
	PR_AS63_f5(26)_187_C	TCAAAGTAARACARTATGATCAAATA	26	x	x
188	PR_AS63_f_188_T	AAGACAGTATGATCAGATAC	20	x	x
	PR_AS63_f2_188_T	AARACARTATGATCAAGTAC	20	x	x
	PR_AS63_f3_188_T	AARACARTATGATCAAATAC	20	x	x
	PR_AS63_f4_188_T	AARACARTATGATCAGGTAC	20	x	x
189	PR_AS63_f2(28)_189_C	TCAAAGTAARACARTATGATCAGATACT	28	x	
	PR_AS63_r1_189_G	AGCTTTWTKTCCGCAGATTTCTAT	24		x
	PR_AS63_r2_189_G	AGCTTTWTKTCCGCAGATTTCTAC	24		x
	PR_AS63_r3_189_G	AGCTTTWTKTCCGCAAATTTCTAT	24		x
	PR_AS63_r4_189_G	AGCTTTWTKTCCGCAAATTTCTAC	24		x
	PR_AS63_r5_189_G	AGCTTTWTKKCCGCAGATTTCTAT	24		x
	PR_AS63_r6_189_G	AGCTTTWTKKCCGCAGATTTCTAC	24		x
	PR_AS63_r7_189_G	AGCTTTWTKKCCGCAAATTTCTAT	24		x
	PR_AS63_r8_189_G	AGCTTTWTKKCCGCAAATTTCTAC	24		x
	PR_AS63_r9_189_G	AGCTTTWTKTCCACAGATTTCTAT	24		x
	PR_AS63_r10_189_G	AGCTTTWTKTCCACAGATTTCTAC	24		x
	PR_AS63_r11_189_G	AGCTTTWTKTCCACAAATTTCTAT	24		x
	PR_AS63_r12_189_G	AGCTTTWTKTCCACAAATTTCTAC	24		x
	PR_AS63_r13_189_G	AGCTTTWTKKCCACAGATTTCTAT	24		x
	PR_AS63_r14_189_G	AGCTTTWTKKCCACAGATTTCTAC	24		x
	PR_AS63_r15_189_G	AGCTTTWTKKCCACAAATTTCTAT	24		x
	PR_AS63_r16_189_G	AGCTTTWTKKCCACAAATTTCTAC	24		x

71	211	PR_AS71_r_211_C	TACTAATACTGTACCTATAG	20	x	x
		PR_AS71_r2(25)_211_C	GGTCCTAYTAAYACTGTACCTATAG	25	x	x
		PR_AS71_r3(25)_211_C	GGTCCTAYTAAYACTGTACCTACAA	25	x	x
		PR_AS71_r4(25)_211_C	GGTCCTAYTAAYACTGTACCTATAA	25	x	
		PR_AS71_r6(25)_211_C	GGTCCTAYTAAYACTGTACCTACAG	25	x	x
	212	PR_AS71_r_212_G	CTACTAATACTGTACCTATA	20	x	x
		PR_AS71_r2(24)_212_G	GGTCCTAYTAAYACTGTACCTATA	24	x	x
		PR_AS71_r3(24)_212_G	GGTCCTAYTAAYACTGTACCTACA	24	x	x
	213	PR_AS71_r_213_A	CCTACTAATACTGTACCTAT	20	x	x
		PR_AS71_r2(30)_213_A	AGGTGTRGGTCCTAYTAATACTGTACCTAT	30	x	x
		PR_AS71_r3(30)_213_A	AGGTGTRGGTCCTAYTAATACTGTACCTAC	30	x	x
		PR_AS71_r4(30)_213_A	AGGTGTRGGTCCTAYTAACACTGTACCTAT	30	x	x
		PR_AS71_r5(30)_213_A	AGGTGTRGGTCCTAYTAACACTGTACCTAC	30	x	x
	73	217	PR_AS73_r_217_C	AGGTCCCTACTAATACTGTAC	20	x
		PR_AS73_r2(27)_217_C	CAGGTGTRGGTCCTAYTAATACTGTAC	27	x	x
		PR_AS73_r3(27)_217_C	CAGGTGTRGGTCCTAYTAACACTGTAC	27	x	x
		PR_AS73_r4(27)_217_C	CAGGTGTRGGTCCTAYTAATACTGTAG	27	x	x
		PR_AS73_r5(27)_217_C	CAGGTGTRGGTCCTAYTAACACTGTAG	27	x	x
218		PR_AS73_r_218_C	TAGGTCCTACTAATACTGTA	20	x	x
		PR_AS73_r2(26)_218_C	CAGGTGTRGGTCCTAYTAATACTGTA	26	x	x
219		PR_AS73_r_219_A	GTAGGTCCTACTAATACTGT	20	x	x
		PR_AS73_r2(30)_219_A	GTTGRCAGGTGTRGGTCCTACTAATACTGT	30	x	x
		PR_AS73_r3(30)_219_A	GTTGRCAGGTGTRGGTCCTACTAACACTGT	30	x	x
	PR_AS73_r4(30)_219_A	GTTGRCAGGTGTRGGTCCTATTAATACTGT	30	x	x	
	PR_AS73_r5(30)_219_A	GTTGRCAGGTGTRGGTCCTATTAACACTGT	30	x	x	
76	226	PR_AS76_r1(30)_226_A	CAATTATGTTGRCAGGTGTRGGTCCTACTA	30	x	x
		PR_AS76_r2(30)_226_A	CAATTATGTTGRCAGGTGTRGGTCCTATTA	30	x	x
	227	PR_AS76_r1(30)_227_A	CCAATTATGTTGRCAGGTGTRGGTCCTACT	30	x	x
		PR_AS76_r2(30)_227_A	CCAATTATGTTGRCAGGTGTRGGTCCTATT	30	x	x
	228	PR_AS76_r1(30)_228_T	TCCAATTATGTTGRCAGGTGTRGGTCCTAC	30	x	
		PR_AS76_r2(30)_228_T	TCCAATTATGTTGRCAGGTGTRGGTCCTAT	30	x	x
		PR_AS76_f3_228_A	GGACATAARGYTATAGGTACAGTATT	26		x
		PR_AS76_f4_228_A	GGACATAARGYTATAGGTACAGTAGT	26		x
		PR_AS76_f5_228_A	GGACATAARGYTATAGGTACAGTGTT	26		x
		PR_AS76_f6_228_A	GGACATAARGYTATAGGTACAGTGGT	26		x
		PR_AS76_f7_228_A	GGACATAARGYTGTAGGTACAGTATT	26		x
		PR_AS76_f8_228_A	GGACATAARGYTGTAGGTACAGTAGT	26		x
		PR_AS76_f9_228_A	GGACATAARGYTGTAGGTACAGTGTT	26		x
		PR_AS76_f10_228_A	GGACATAARGYTGTAGGTACAGTGGT	26		x
		PR_AS76_f11_228_A	GGACATAARGYTGTAGGTACAGTATT	26		x
		PR_AS76_f12_228_A	GGACATAARGYTGTAGGTACAGTAGT	26		x
		PR_AS76_f13_228_A	GGACATAARAYTATAGGTACAGTGTT	26		x
		PR_AS76_f14_228_A	GGACATAARAYTATAGGTACAGTGGT	26		x
PR_AS76_f15_228_A		GGACATAARAYTATAGGTACAGTATT	26		x	
PR_AS76_f16_228_A		GGACATAARAYTATAGGTACAGTAGT	26		x	
PR_AS76_f17_228_A		GGACATAARAYTGTAGGTACAGTGTT	26		x	
PR_AS76_f18_228_A		GGACATAARAYTGTAGGTACAGTGGT	26		x	
77	229	PR_AS77_r_229_C	GTTGACAGGTGTAGGTCCTA	20	x	x
		PR_AS77_r2_229_C	GTTGRCAGGTGTGGGTCCTA	20	x	x
	230	PR_AS77_r_230_A	TGTTGACAGGTGTAGGTCCT	20	x	x
		PR_AS77_r2_230_A	TGTTGRCAGGTGTGGGTCCT	20	x	x
	231	PR_AS77_r_231_T	ATGTTGACAGGTGTAGGTCC	20	x	x
		PR_AS77_r2_231_T	ATGTTGGCAGGTGTGGGTCCT	20	x	x
PR_AS77_r3_231_T		ATGTTGACAGGTGTGGGTCCT	20	x	x	
PR_AS77_r4_231_T		ATGTTGGCAGGTGTAGGTCC	20	x	x	
82	244	PR_AS82_f_244_G	TATTAGTAGGACCTACACCT	20	x	x
		PR_AS82_f(29)_244_G	TAGGTACAGTRTTARTAGGACCCYACACCT	29	x	x
	245	PR_AS82_f_245_T	ATTAGTAGGACCTACACCTG	20	x	x
		PR_AS82_f2_245_T	RTTARTAGGACCCACACCTG	20	x	x
		PR_AS82_f2(30)_245_T	TAGGTACAGTRTTARTAGGACCCYACACCTG	30	x	x
		PR_AS82_f3(30)_245_T	TAGGTACAGTRTTARTAGGACCCYACACCTA	30	x	x
		PR_AS82_f4(30)_245_T	TAGGTACAGTRTTARTAGGACCCYACACCTT	30	x	x
	PR_AS82_f5(30)_245_T	TAGGTACAGTRTTARTAGGACCCYACACCTC	30	x	x	
	246	PR_AS82_f_246_C	TTAGTAGGACCTACACCTGT	20	x	x
		PR_AS82_f2_246_C	TTARTAGGACCCACACCTGC	20	x	x
		PR_AS82_f3_246_C	TTARTAGGACCTACACCTGC	20	x	x
		PR_AS82_f4_246_C	TTARTAGGACCCACACCTGT	20	x	x

		PR_AS82_f2(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTGT	30	x	x
		PR_AS82_f3(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTAT	30	x	x
		PR_AS82_f4(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTTT	30	x	x
		PR_AS82_f5(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTCT	30	x	x
		PR_AS82_f6(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTGC	30	x	x
		PR_AS82_f7(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTAC	30	x	x
		PR_AS82_f8(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTTC	30	x	x
		PR_AS82_f9(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTCC	30	x	x
		PR_AS82_f10(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTGG	30	x	x
		PR_AS82_f11(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTAG	30	x	x
		PR_AS82_f12(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTTG	30	x	x
		PR_AS82_f13(30)_246_C	AGGTACAGTRTTARTAGGACCYACACCTCG	30	x	x
84	250	PR_AS84_f_250_A	TAGGACCTACACCTGTCAAC	20	x	x
		PR_AS84_f2_250_A	TAGGACCYACACCTGCCAAC	20	x	x
	251	PR_AS84_f_251_T	AGGACCTACACCTGTCAACA	20	x	x
		PR_AS84_f2_251_T	AGGACCYACACCTGCCAAC	20	x	x
		PR_AS84_f2(25)_251_T	TTARTAGGACCYACACCTGTCAACA	25	x	x
		PR_AS84_f3(25)_251_T	TTARTAGGACCYACACCTGTCAACG	25	x	x
		PR_AS84_f4(25)_251_T	TTARTAGGACCYACACCTGCCAAC	25	x	x
		PR_AS84_f5(25)_251_T	TTARTAGGACCYACACCTGCCAACG	25	x	x
	252	PR_AS84_f_252_A	GGACCTACACCTGTCAACAT	20	x	x
		PR_AS84_f2_252_A	GGACCYACACCTGCCAACAT	20	x	x
		PR_AS84_f2(26)_252_A	TTARTAGGACCYACACCTGTCAACAT	26	x	x
		PR_AS84_f3(26)_252_A	TTARTAGGACCYACACCTGTCAACGT	26	x	x
		PR_AS84_f4(26)_252_A	TTARTAGGACCYACACCTGCCAACAT	26	x	x
		PR_AS84_f5(26)_252_A	TTARTAGGACCYACACCTGCCAACGT	26	x	x
88	262	PR_AS88_f_262_A	CTGTCAACATAATTGGAAGA	20	x	x
		PR_AS88_f2(30)_262_A	GGACCYACACCTGYCAACATAATTGGAAGA	30	x	x
	263	PR_AS88_f_263_A	TGTCAACATAATTGGAAGAA	20	x	x
		PR_AS88_f2(30)_263_A	GACCYACACCTGYCAACATAATTGGAAGAA	30	x	x
		PR_AS88_f3(30)_263_A	GACCYACACCTGYCAACATAATTGGAAGAG	30	x	x
	264	PR_AS88_f_264_T	GTCAACATAATTGGAAGAAA	20	x	x
		PR_AS88_f3(30)_264_T	ACCYACACCTGYCAACATAATTGGAAGAAA	30	x	x
		PR_AS88_f4(30)_264_T	ACCYACACCTGYCAACATAATTGGAAGAAG	30	x	x
		PR_AS88_f5(30)_264_T	ACCYACACCTGYCAACATAATTGGAAGAGA	30	x	x
		PR_AS88_f6(30)_264_T	ACCYACACCTGYCAACATAATTGGAAGAGG	30	x	x
90	268	PR_AS90_f_268_T	ACATAATTGGAAGAAATCTG	20	x	x
		PR_AS90_f(30)_268_T	ACACCTGYCAACATAATTGGAAGAAATCTG	30	x	x
		PR_AS90_f2(30)_268_T	ACACCTGYCAACATAATTGGAAGAAATATG	30	x	x
		PR_AS90_f3(30)_268_T	ACACCTGYCAACATAATTGGAAGAAATTTG	30	x	x
		PR_AS90_f4(30)_268_T	ACACCTGYCAACATAATTGGAAGAAACCTG	30	x	x
		PR_AS90_f5(30)_268_T	ACACCTGYCAACATAATTGGAAGAAACATG	30	x	x
		PR_AS90_f6(30)_268_T	ACACCTGYCAACATAATTGGAAGAAACTTG	30	x	x
		PR_AS90_r_268_A	ACTGCAGCCAATCTGAGTCA	20	x	x
	269	PR_AS90_f_269_T	CATAATTGGAAGAAATCTGT	20	x	x
		PR_AS90_f2(30)_269_T	CACCTGYCAACATAATTGGAAGAAATATGT	30	x	x
		PR_AS90_f3(30)_269_T	CACCTGYCAACATAATTGGAAGAAATTTGT	30	x	x
		PR_AS90_f4(30)_269_T	CACCTGYCAACATAATTGGAAGAAACCTGT	30	x	x
		PR_AS90_f5(30)_269_T	CACCTGYCAACATAATTGGAAGAAACATGT	30	x	x
		PR_AS90_f6(30)_269_T	CACCTGYCAACATAATTGGAAGAAACTTGT	30	x	x
		PR_AS90_f7(30)_269_T	CACCTGYCAACATAATTGGAAGAAATCTGA	30	x	x
		PR_AS90_f8(30)_269_T	CACCTGYCAACATAATTGGAAGAAATATGA	30	x	x
		PR_AS90_f9(30)_269_T	CACCTGYCAACATAATTGGAAGAAATTTGA	30	x	x
		PR_AS90_f10(30)_269_T	CACCTGYCAACATAATTGGAAGAAACCTGA	30	x	x
		PR_AS90_f11(30)_269_T	CACCTGYCAACATAATTGGAAGAAACATGA	30	x	x
		PR_AS90_f12(30)_269_T	CACCTGYCAACATAATTGGAAGAAACTTGA	30	x	x
		PR_AS90_f(30)_269_T	CACCTGYCAACATAATTGGAAGAAATCTGT	30	x	x
		PR_AS90_r_269_A	AACCTGCAGCCAATCTGAGTC	20	x	x
	270	PR_AS90_f_270_G	ATAATTGGAAGAAATCTGTT	20	x	x
		PR_AS90_f(30)_270_G	ACCTGYCAACATAATTGGAAGAAATCTGTT	30	x	x
		PR_AS90_f2(30)_270_G	ACCTGYCAACATAATTGGAAGAAATATGTT	30	x	x
		PR_AS90_f3(30)_270_G	ACCTGYCAACATAATTGGAAGAAATTTGTT	30	x	x
		PR_AS90_f4(30)_270_G	ACCTGYCAACATAATTGGAAGAAACCTGTT	30	x	x
		PR_AS90_f5(30)_270_G	ACCTGYCAACATAATTGGAAGAAACATGTT	30	x	x
		PR_AS90_f6(30)_270_G	ACCTGYCAACATAATTGGAAGAAACTTGT	30	x	x
		PR_AS90_f7(30)_270_G	ACCTGYCAACATAATTGGAAGAAATCTGAT	30	x	x
		PR_AS90_f8(30)_270_G	ACCTGYCAACATAATTGGAAGAAATATGAT	30	x	x
		PR_AS90_f9(30)_270_G	ACCTGYCAACATAATTGGAAGAAATTTGAT	30	x	x

		PR_AS90_f10(30)_270_G	ACCTGYCAACATAAATTGGAAGAAACCTGAT	30	x	x
		PR_AS90_f11(30)_270_G	ACCTGYCAACATAAATTGGAAGAAACATGAT	30	x	x
		PR_AS90_f12(30)_270_G	ACCTGYCAACATAAATTGGAAGAAACTTGAT	30	x	x
		PR_AS90_r_270_C	AAACTGCAGCCAATCTGAGT	20	x	x
93	277	PR_AS93_r1_277_T	AAAATTTAAACTGCAGCCAA	20	x	x
		PR_AS93_r1_278_A	AAAATTTAAACTGCAGCCA	19	x	x
		PR_AS93_r1_279_A	AAAATTTAAACTGCAGCC	18	x	x
		PR_AS93_f1_277_A	GAAGAAATCTGTTGACTCAG	20	x	x
		PR_AS93_f(30)_277_A	AACATAAATTGGAAGAAAYCTGTTGACTCAG	30	x	x
		PR_AS93_f2(30)_277_A	AACATAAATTGGAAGAAAYHTGTTGACTCAG	30	x	x
		PR_AS93_f2(30)_277_A	AACATAAATTGGAAGAAAYHTGTTGACTCAG	30	x	x
		PR_AS93_f3(30)_277_A	AACATAAATTGGAAGAAAYHTGTTGACTCAA	30	x	x
		PR_AS93_f4(30)_277_A	AACATAAATTGGAAGAAAYHTGATGACTCAG	30	x	x
		PR_AS93_f5(30)_277_A	AACATAAATTGGAAGAAAYHTGATGACTCAA	30	x	x
		PR_AS93_r2_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r3_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r4_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r5_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r6_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r7_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r8_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
		PR_AS93_r9_277_T	AATGGGAAAATTTARAGTGCAGCCAA	26		x
	278	PR_AS93_f1_278_T	AAGAAATCTGTTGACTCAGA	20	x	x
		PR_AS93_f(30)_278_T	ACATAAATTGGAAGAAAYCTGTTGACTCAGA	30	x	x
		PR_AS93_f2(30)_278_T	ACATAAATTGGAAGAAAYHTGTTGACTCAGA	30	x	x
		PR_AS93_f3(30)_278_T	ACATAAATTGGAAGAAAYHTGTTGACTCAA	30	x	x
		PR_AS93_f4(30)_278_T	ACATAAATTGGAAGAAAYHTGATGACTCAGA	30	x	x
		PR_AS93_f5(30)_278_T	ACATAAATTGGAAGAAAYHTGATGACTCAA	30	x	x
		PR_AS93_f6(30)_278_T	ACATAAATTGGAAGAAAYHTGTTGACTCAGC	30	x	x
		PR_AS93_f7(30)_278_T	ACATAAATTGGAAGAAAYHTGTTGACTCAAC	30	x	
		PR_AS93_f8(30)_278_T	ACATAAATTGGAAGAAAYHTGATGACTCAGC	30	x	x
		PR_AS93_f9(30)_278_T	ACATAAATTGGAAGAAAYHTGATGACTCAAC	30	x	x
	279	PR_AS93_f1_279_T	AGAAATCTGTTGACTCAGAT	20	x	x
		PR_AS93_f(30)_279_T	CATAAATTGGAAGAAAYCTGTTGACTCAGAT	30	x	x
		PR_AS93_f2(30)_279_T	CATAAATTGGAAGAAAYHTGTTGACTCAGAT	30	x	x
		PR_AS93_f3(30)_279_T	CATAAATTGGAAGAAAYHTGTTGACTCAAAT	30	x	x
		PR_AS93_f4(30)_279_T	CATAAATTGGAAGAAAYHTGATGACTCAGAT	30	x	x
		PR_AS93_f5(30)_279_T	CATAAATTGGAAGAAAYHTGATGACTCAAAT	30	x	x
		PR_AS93_f6(30)_279_T	CATAAATTGGAAGAAAYHTGTTGACTCAGCT	30	x	x
		PR_AS93_f7(30)_279_T	CATAAATTGGAAGAAAYHTGTTGACTCAACT	30	x	x
		PR_AS93_f8(30)_279_T	CATAAATTGGAAGAAAYHTGATGACTCAGCT	30	x	x
		PR_AS93_f9(30)_279_T	CATAAATTGGAAGAAAYHTGATGACTCAACT	30	x	x

Oligonukleotide für den Reverse Transkriptase-Chip

Position	Name	Sequenz 5'-3'	Länge	
41	121	RT_AS41_f1_121_A	ATAAAGCATTARTAGAAATYTGACAGAA	30
	122	RT_AS41_f1_122_T	TAAAAGCATTARTAGAAATYTGACAGAAA	30
		RT_AS41_f2_122_T	TAAAAGCATTARTAGAAATYTGACAGAAC	30
		RT_AS41_f3_122_T	TAAAAGCATTARTAGAAATYTGACAGAAT	30
	123	RT_AS41_f1_123_G	AAAAGCATTARTAGAAATYTGACAGAAAT	30
		RT_AS41_f2_123_G	AAAAGCATTARTAGAAATYTGACAGAACT	30
		RT_AS41_f3_123_G	AAAAGCATTARTAGAAATYTGACAGAATT	30
44	130	RT_AS44_r1_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCTT	30
		RT_AS44_r2_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCAT	30
		RT_AS44_r3_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCGT	30
		RT_AS44_r4_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCTG	30
		RT_AS44_r5_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCAG	30
		RT_AS44_r6_130_C	TTTCAGGCCCAATTTTTGAAATTTTYCCGG	30

131	RT_AS44_r1_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTTCCT	30	
	RT_AS44_r2_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTCCA	30	
	RT_AS44_r3_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTCCG	30	
	RT_AS44_r4_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTCCCT	30	
	RT_AS44_r5_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTCCA	30	
	RT_AS44_r6_131_T	TTTTTCAGGCCCAATTTTTGAAATTTTTCCG	30	
132	RT_AS44_r1_132_T	ATTTTCAGGCCCAATTTTTGAAATTTTTCC	30	
	RT_AS44_r2_132_T	ATTTTCAGGCCCAATTTTTGAAATTTTTCCC	30	
62	184	RT_AS62_f1_184_G	CCTGAAAATCCATAYAATACTCCAGTATTT	30
	185	RT_AS62_f1_185_C	CTGAAAATCCATAYAATACTCCAGTATTTG	30
	186	RT_AS62_f1_186_C	TGAAAATCCATAYAATACTCCAGTATTTGC	30
		RT_AS62_f2_186_C	TGAAAATCCATAYAATACTCCAGTATTTGT	30
65	193	RT_AS65_f1_193_A	CCATAYAATACTCCAGTATTTGYCATAAAG	30
		RT_AS65_f2_193_A	CCATAYAATACTCCAGTATTTGYTATAAAG	30
	194	RT_AS65_f1_194_A	CATAYAATACTCCAGTATTTGYCATAAAGA	30
		RT_AS65_f2_194_A	CATAYAATACTCCAGTATTTGYTATAAAGG	30
	195	RT_AS65_f1_195_A	ATAYAATACTCCAGTATTTGYCATAAAGAA	30
		RT_AS65_f2_195_A	ATAYAATACTCCAGTATTTGYTATAAAGAA	30
		RT_AS65_f3_195_A	ATAYAATACTCCAGTATTTGYCATAAAGGA	30
		RT_AS65_f4_195_A	ATAYAATACTCCAGTATTTGYTATAAAGGA	30
		RT_AS65_f5_195_A	ATAYAATACTCCAGTATTTGYCATAAAGAG	30
		RT_AS65_f6_195_A	ATAYAATACTCCAGTATTTGYTATAAAGAG	30
		RT_AS65_f7_195_A	ATAYAATACTCCAGTATTTGYCATAAAGGG	30
		RT_AS65_f8_195_A	ATAYAATACTCCAGTATTTGYTATAAAGGG	30
67	199	RT_AS67_f1_199_G	AATACTCCAGTATTTGYYATAAAGAAAAAA	30
		RT_AS67_f1_201_C	TACTCCAGTATTTGYYATAAAGRAAAAAGA	30
		RT_AS67_r1_199_C	TAATTTTCTCCATYTAITACTGT	23
		RT_AS67_r2_199_C	TAATTTTCTCCATYTAITACTGC	23
		RT_AS67_r3_199_C	TAATTTTCTCCATYTAITACTTT	23
		RT_AS67_r4_199_C	TAATTTTCTCCATYTAITACTTC	23
		RT_AS67_r5_199_C	TAATTTTCTCCATYTAITACTCT	23
		RT_AS67_r6_199_C	TAATTTTCTCCATYTAITACTCC	23
		RT_AS67_r7_199_C	TAATTTTCTCCATYTAICTACTGT	23
		RT_AS67_r8_199_C	TAATTTTCTCCATYTAICTACTGC	23
		RT_AS67_r9_199_C	TAATTTTCTCCATYTAICTACTTT	23
		RT_AS67_r10_199_C	TAATTTTCTCCATYTAICTACTTC	23
		RT_AS67_r11_199_C	TAATTTTCTCCATYTAICTACTCT	23
		RT_AS67_r12_199_C	TAATTTTCTCCATYTAICTACTCC	23
	200	RT_AS67_f1_200_A	ATACTCCAGTATTTGYYATAAAGRAAAAAG	30
		RT_AS67_r1_200_T	TAATTTTCTCCATYTAITACTG	22
		RT_AS67_r2_200_T	TAATTTTCTCCATYTAITACTT	22
	201	RT_AS67_r1_201_G	TAHTTTTCTCCATTTAGTACT	21
		RT_AS67_r2_201_G	TAHTTTTCTCCATTTAGCACT	21
		RT_AS67_r3_201_G	TAHTTTTCTCCATTTATTACT	21
		RT_AS67_r4_201_G	TAHTTTTCTCCATTTATCACT	21
		RT_AS67_r5_201_G	TAHTTTTCTCCATTTACTACT	21
		RT_AS67_r6_201_G	TAHTTTTCTCCATTTACCACT	21
		RT_AS67_r7_201_G	TAHTTTTCTCCATTTAATACT	21
		RT_AS67_r8_201_G	TAHTTTTCTCCATTTAACACT	21
		RT_AS67_r9_201_G	TAHTTTTCTCCATCTAGTACT	21
		RT_AS67_r10_201_G	TAHTTTTCTCCATCTAGCACT	21
		RT_AS67_r11_201_G	TAHTTTTCTCCATCTATTACT	21
		RT_AS67_r12_201_G	TAHTTTTCTCCATCTATCACT	21
		RT_AS67_r13_201_G	TAHTTTTCTCCATCTACTACT	21

		RT_AS67_r14_201_G	TAHTTTTCTCCATCTACCACT	21
		RT_AS67_r15_201_G	TAHTTTTCTCCATCTAATACT	21
		RT_AS67_r16_201_G	TAHTTTTCTCCATCTAACACT	21
69	205	RT_AS69_r1_205_T	ATCYAYTAITTTTCTCCATTTAG	23
		RT_AS69_r2_205_T	ATCYAYTAITTTTCTCCATTTAT	23
		RT_AS69_r3_205_T	ATCYAYTAITTTTCTCCATTTAC	23
		RT_AS69_r4_205_T	ATCYAYTAITTTTCTCCATTTAA	23
		RT_AS69_r5_205_T	ATCYAYTAITTTTCTCCATCTAG	23
		RT_AS69_r6_205_T	ATCYAYTAITTTTCTCCATCTAT	23
		RT_AS69_r7_205_T	ATCYAYTAITTTTCTCCATCTAC	23
		RT_AS69_r8_205_T	ATCYAYTAITTTTCTCCATCTAA	23
		RT_AS69_r9_205_T	ATCYGYTAITTTTCTCCATTTAG	23
		RT_AS69_r10_205_T	ATCYGYTAITTTTCTCCATTTAT	23
		RT_AS69_r11_205_T	ATCYGYTAITTTTCTCCATTTAC	23
		RT_AS69_r12_205_T	ATCYGYTAITTTTCTCCATTTAA	23
		RT_AS69_r13_205_T	ATCYGYTAITTTTCTCCATCTAG	23
		RT_AS69_r14_205_T	ATCYGYTAITTTTCTCCATCTAT	23
		RT_AS69_r15_205_T	ATCYGYTAITTTTCTCCATCTAC	23
		RT_AS69_r16_205_T	ATCYGYTAITTTTCTCCATCTAA	23
	206	RT_AS69_r1_206_G	TCTCTGARATCYACTAITTTTCTCCATTTA	30
		RT_AS69_r2_206_G	TCTCTGARATCYACTAITTTTCTCCATCTA	30
		RT_AS69_r3_206_G	TCTCTGARATCYATTAITTTTCTCCATTTA	30
		RT_AS69_r4_206_G	TCTCTGARATCYATTAITTTTCTCCATCTA	30
		RT_AS69_r5_206_G	TCTCTGARATCYGCTAITTTTCTCCATTTA	30
		RT_AS69_r6_206_G	TCTCTGARATCYGCTAITTTTCTCCATCTA	30
		RT_AS69_r7_206_G	TCTCTGARATCYGTTAITTTTCTCCATTTA	30
		RT_AS69_r8_206_G	TCTCTGARATCYGTTAITTTTCTCCATCTA	30
	207	RT_AS69_r1_207_A	TTCTCTGARATCYACTAITTTTCTCCATTT	30
		RT_AS69_r2_207_A	TTCTCTGARATCYACTAITTTTCTCCATCT	30
		RT_AS69_r3_207_A	TTCTCTGARATCYATTAITTTTCTCCATTT	30
		RT_AS69_r4_207_A	TTCTCTGARATCYATTAITTTTCTCCATCT	30
		RT_AS69_r5_207_A	TTCTCTGARATCYGCTAITTTTCTCCATTT	30
		RT_AS69_r6_207_A	TTCTCTGARATCYGCTAITTTTCTCCATCT	30
		RT_AS69_r7_207_A	TTCTCTGARATCYGTTAITTTTCTCCATTT	30
		RT_AS69_r8_207_A	TTCTCTGARATCYGTTAITTTTCTCCATCT	30
70	208	RT_AS70_r1_208_T	GTTCTCTGARATCYACTAATTTTCTCCATT	30
		RT_AS70_r2_208_T	GTTCTCTGARATCYACTAATTTTCTCCATC	30
		RT_AS70_r3_208_T	GTTCTCTGARATCYACTACTTTTCTCCATT	30
		RT_AS70_r4_208_T	GTTCTCTGARATCYACTACTTTTCTCCATC	30
		RT_AS70_r5_208_T	GTTCTCTGARATCYACTATTTTCTCCATT	30
		RT_AS70_r6_208_T	GTTCTCTGARATCYACTATTTTCTCCATC	30
		RT_AS70_r7_208_T	GTTCTCTGARATCYATTAATTTTCTCCATT	30
		RT_AS70_r8_208_T	GTTCTCTGARATCYATTAATTTTCTCCATC	30
		RT_AS70_r9_208_T	GTTCTCTGARATCYATTACTTTTCTCCATT	30
		RT_AS70_r10_208_T	GTTCTCTGARATCYATTACTTTTCTCCATC	30
		RT_AS70_r11_208_T	GTTCTCTGARATCYATTATTTTCTCCATT	30
		RT_AS70_r12_208_T	GTTCTCTGARATCYATTATTTTCTCCATC	30
		RT_AS70_r13_208_T	GTTCTCTGARATCYGCTAATTTTCTCCATT	30
		RT_AS70_r14_208_T	GTTCTCTGARATCYGCTAATTTTCTCCATC	30
		RT_AS70_r15_208_T	GTTCTCTGARATCYGCTACTTTTCTCCATT	30
		RT_AS70_r16_208_T	GTTCTCTGARATCYGCTACTTTTCTCCATC	30
		RT_AS70_r17_208_T	GTTCTCTGARATCYGCTATTTTCTCCATT	30
		RT_AS70_r18_208_T	GTTCTCTGARATCYGCTATTTTCTCCATC	30
		RT_AS70_r19_208_T	GTTCTCTGARATCYGTTAATTTTCTCCATT	30
		RT_AS70_r20_208_T	GTTCTCTGARATCYGTTAATTTTCTCCATC	30

		RT_AS70_r21_208_T	GTTCTCTGARATCYGTTACTTTTCTCCATT	30
		RT_AS70_r22_208_T	GTTCTCTGARATCYGTTACTTTTCTCCATC	30
		RT_AS70_r23_208_T	GTTCTCTGARATCYGTTATTTTTCTCCATT	30
		RT_AS70_r24_208_T	GTTCTCTGARATCYGTTATTTTTCTCCATC	30
209		RT_AS70_r1_209_T	AGTTCTCTGARATCYACTAATTTTTCTCCAT	30
		RT_AS70_r2_209_T	AGTTCTCTGARATCYACTACTTTTCTCCAT	30
		RT_AS70_r3_209_T	AGTTCTCTGARATCYACTATTTTTCTCCAT	30
		RT_AS70_r4_209_T	AGTTCTCTGARATCYATTAATTTTTCTCCAT	30
		RT_AS70_r5_209_T	AGTTCTCTGARATCYATTACTTTTCTCCAT	30
		RT_AS70_r6_209_T	AGTTCTCTGARATCYATTATTTTTCTCCAT	30
		RT_AS70_r7_209_T	AGTTCTCTGARATCYGCTAATTTTTCTCCAT	30
		RT_AS70_r8_209_T	AGTTCTCTGARATCYGCTACTTTTCTCCAT	30
		RT_AS70_r9_209_T	AGTTCTCTGARATCYGCTATTTTTCTCCAT	30
		RT_AS70_r10_209_T	AGTTCTCTGARATCYGTTAATTTTTCTCCAT	30
		RT_AS70_r11_209_T	AGTTCTCTGARATCYGTTACTTTTCTCCAT	30
		RT_AS70_r12_209_T	AGTTCTCTGARATCYGTTATTTTTCTCCAT	30
210		RT_AS70_r1_210_T	AAGTTCTCTGARATCYACTAATTTTTCTCCA	30
		RT_AS70_r2_210_T	AAGTTCTCTGARATCYACTACTTTTCTCCA	30
		RT_AS70_r3_210_T	AAGTTCTCTGARATCYACTATTTTTCTCCA	30
		RT_AS70_r4_210_T	AAGTTCTCTGARATCYATTAATTTTTCTCCA	30
		RT_AS70_r5_210_T	AAGTTCTCTGARATCYATTACTTTTCTCCA	30
		RT_AS70_r6_210_T	AAGTTCTCTGARATCYATTATTTTTCTCCA	30
		RT_AS70_r7_210_T	AAGTTCTCTGARATCYGCTAATTTTTCTCCA	30
		RT_AS70_r8_210_T	AAGTTCTCTGARATCYGCTACTTTTCTCCA	30
		RT_AS70_r9_210_T	AAGTTCTCTGARATCYGCTATTTTTCTCCA	30
		RT_AS70_r10_210_T	AAGTTCTCTGARATCYGTTAATTTTTCTCCA	30
		RT_AS70_r11_210_T	AAGTTCTCTGARATCYGTTACTTTTCTCCA	30
		RT_AS70_r12_210_T	AAGTTCTCTGARATCYGTTATTTTTCTCCA	30
74	220	RT_AS74_r1_220_A	TTATTAAGTTCTCTGARATCTACTA	25
		RT_AS74_r2_220_A	TTATTAAGTTCTCTGARATCTATTA	25
		RT_AS74_r3_220_A	TTATTAAGTTCTCTGARATCTGCTA	25
		RT_AS74_r4_220_A	TTATTAAGTTCTCTGARATCTGTTA	25
		RT_AS74_r5_220_A	TTATTAAGTTCTCTGARATCCACTA	25
		RT_AS74_r6_220_A	TTATTAAGTTCTCTGARATCCATTA	25
		RT_AS74_r7_220_A	TTATTAAGTTCTCTGARATCCGCTA	25
		RT_AS74_r8_220_A	TTATTAAGTTCTCTGARATCCGTTA	25
221		RT_AS74_r1_221_A	TYTTATTAAGTTCTCTGARATCTACT	26
		RT_AS74_r2_221_A	TYTTATTAAGTTCTCTGARATCTATT	26
		RT_AS74_r3_221_A	TYTTATTAAGTTCTCTGARATCTGCT	26
		RT_AS74_r4_221_A	TYTTATTAAGTTCTCTGARATCTGTT	26
		RT_AS74_r5_221_A	TYTTATTAAGTTCTCTGARATCCACT	26
		RT_AS74_r6_221_A	TYTTATTAAGTTCTCTGARATCCATT	26
		RT_AS74_r7_221_A	TYTTATTAAGTTCTCTGARATCCGCT	26
		RT_AS74_r8_221_A	TYTTATTAAGTTCTCTGARATCCGTT	26
222		RT_AS74_r1_222_T	TYTTATTAAGTTCTCTGARATCTAC	25
		RT_AS74_r2_222_T	TYTTATTAAGTTCTCTGARATCTAT	25
		RT_AS74_r3_222_T	TYTTATTAAGTTCTCTGARATCTGC	25
		RT_AS74_r4_222_T	TYTTATTAAGTTCTCTGARATCTGT	25
		RT_AS74_r5_222_T	TYTTATTAAGTTCTCTGARATCCAC	25
		RT_AS74_r6_222_T	TYTTATTAAGTTCTCTGARATCCAT	25
		RT_AS74_r7_222_T	TYTTATTAAGTTCTCTGARATCCGC	25
		RT_AS74_r8_222_T	TYTTATTAAGTTCTCTGARATCCGT	25
75	223	RT_AS75_r1_223_C	GAGTTYTYTTATTAAGTTCTCTGAAATCTA	30
		RT_AS75_r2_223_C	GAGTTYTYTTATTAAGTTCTCTGAAATCTG	30
		RT_AS75_r3_223_C	GAGTTYTYTTATTAAGTTCTCTGAAATCCA	30

		RT_AS75_r4_223_C	GAGTTYTYTTATTAAGTTCTCTGAAATCCG	30
		RT_AS75_r5_223_C	GAGTTYTYTTATTAAGTTCTCTGAGATCTA	30
		RT_AS75_r6_223_C	GAGTTYTYTTATTAAGTTCTCTGAGATCTG	30
		RT_AS75_r7_223_C	GAGTTYTYTTATTAAGTTCTCTGAGATCCA	30
		RT_AS75_r8_223_C	GAGTTYTYTTATTAAGTTCTCTGAGATCCG	30
224		RT_AS75_r1_224_A	TGAGTTYTYTTATTAAGTTCTCTGAAATCT	30
		RT_AS75_r2_224_A	TGAGTTYTYTTATTAAGTTCTCTGAAATCC	30
		RT_AS75_r3_224_A	TGAGTTYTYTTATTAAGTTCTCTGAGATCT	30
		RT_AS75_r4_224_A	TGAGTTYTYTTATTAAGTTCTCTGAGATCC	30
225		RT_AS75_r1_225_T	TTGAGTTYTYTTATTAAGTTCTCTGAAATC	30
		RT_AS75_r2_225_T	TTGAGTTYTYTTATTAAGTTCTCTGAGATC	30
77	229	RT_AS77_r1_229_A	TCTTGAGTTYTYTTATTAAGTTCTCTGA	28
	230	RT_AS77_r1_230_A	TCTTGAGTTYTYTTATTAAGTTCTCTG	27
	231	RT_AS77_r1_231_G	TCTTGAGTTYTYTTATTAAGTTCTCT	26
98	292	RT_AS98_f1_292_G	TGGGAAGTTCAATTAGGAATACCACATCCT	30
	293	RT_AS98_f1_293_C	GGGAAGTTCAATTAGGAATACCACATCCTG	30
		RT_AS98_f2_293_C	GGGAAGTTCAATTAGGAATACCACATCCTT	30
		RT_AS98_f3_293_C	GGGAAGTTCAATTAGGAATACCACATCCCG	30
		RT_AS98_f4_293_C	GGGAAGTTCAATTAGGAATACCACATCCCT	30
294		RT_AS98_f1_294_A	GGAAGTTCAATTAGGAATACCACATCCTGC	30
		RT_AS98_f2_294_A	GGAAGTTCAATTAGGAATACCACATCCTTC	30
		RT_AS98_f3_294_A	GGAAGTTCAATTAGGAATACCACATCCCGC	30
		RT_AS98_f4_294_A	GGAAGTTCAATTAGGAATACCACATCCCTC	30
		RT_AS98_f5_294_A	GGAAGTTCAATTAGGAATACCACATCCTGG	30
		RT_AS98_f6_294_A	GGAAGTTCAATTAGGAATACCACATCCTTG	30
		RT_AS98_f7_294_A	GGAAGTTCAATTAGGAATACCACATCCCGG	30
		RT_AS98_f8_294_A	GGAAGTTCAATTAGGAATACCACATCCCTG	30
100	298	RT_AS100_f1_298_T	GTTCAATTAGGAATACCACATCCYGCAGGG	30
		RT_AS100_f2_298_T	GTTCAATTAGGAATACCACATCCYGGAGGG	30
		RT_AS100_f3_298_T	GTTCAATTAGGAATACCACATCCYTCAGGG	30
		RT_AS100_f4_298_T	GTTCAATTAGGAATACCACATCCYTGAGGG	30
	299	RT_AS100_f1_299_T	TTCAATTAGGAATACCACATCCYGCAGGGT	30
		RT_AS100_f2_299_T	TTCAATTAGGAATACCACATCCYGGAGGGT	30
		RT_AS100_f3_299_T	TTCAATTAGGAATACCACATCCYTCAGGGT	30
		RT_AS100_f4_299_T	TTCAATTAGGAATACCACATCCYTGAGGGT	30
		RT_AS100_f5_299_T	TTCAATTAGGAATACCACATCCYGCAGGGA	30
		RT_AS100_f6_299_T	TTCAATTAGGAATACCACATCCYGGAGGGA	30
		RT_AS100_f7_299_T	TTCAATTAGGAATACCACATCCYTCAGGGA	30
		RT_AS100_f8_299_T	TTCAATTAGGAATACCACATCCYTGAGGGA	30
	300	RT_AS100_f1_300_A	TCAATTAGGAATACCACATCCYGCAGGGTT	30
		RT_AS100_f2_300_A	TCAATTAGGAATACCACATCCYGGAGGGTT	30
		RT_AS100_f3_300_A	TCAATTAGGAATACCACATCCYTCAGGGTT	30
		RT_AS100_f4_300_A	TCAATTAGGAATACCACATCCYTGAGGGTT	30
		RT_AS100_f5_300_A	TCAATTAGGAATACCACATCCYGCAGGGAT	30
		RT_AS100_f6_300_A	TCAATTAGGAATACCACATCCYGGAGGGAT	30
		RT_AS100_f7_300_A	TCAATTAGGAATACCACATCCYTCAGGGAT	30
		RT_AS100_f8_300_A	TCAATTAGGAATACCACATCCYTGAGGGAT	30
101	301	RT_AS101_f1_301_A	CAATTAGGAATACCACATCCYKCAGGGTTA	30
		RT_AS101_f2_301_A	CAATTAGGAATACCACATCCYKCAGGGATA	30
		RT_AS101_f3_301_A	CAATTAGGAATACCACATCCYKCAGGGTTA	30
		RT_AS101_f4_301_A	CAATTAGGAATACCACATCCYKCAGGGATA	30
	302	RT_AS101_f1_302_A	AATTAGGAATACCACATCCYKCAGGGTTAA	30
		RT_AS101_f2_302_A	AATTAGGAATACCACATCCYKCAGGGATAA	30

		RT_AS101_f3_302_A	AATTAGGAATACCACATCCYKGAGGGTTAA	30
		RT_AS101_f4_302_A	AATTAGGAATACCACATCCYKGAGGGATAA	30
		RT_AS101_f5_302_A	AATTAGGAATACCACATCCYKCAGGGTTAG	30
		RT_AS101_f6_302_A	AATTAGGAATACCACATCCYKCAGGGATAG	30
		RT_AS101_f7_302_A	AATTAGGAATACCACATCCYKGAGGGTTAG	30
		RT_AS101_f8_302_A	AATTAGGAATACCACATCCYKGAGGGATAG	30
		RT_AS101_f9_302_A	AATTAGGAATACCACATCCYKCAGGGTTAC	30
		RT_AS101_f10_302_A	AATTAGGAATACCACATCCYKCAGGGATAC	30
		RT_AS101_f11_302_A	AATTAGGAATACCACATCCYKGAGGGTTAC	30
		RT_AS101_f12_302_A	AATTAGGAATACCACATCCYKGAGGGATAC	30
303		RT_AS101_f1_303_A	ATTAGGAATACCACATCCYKCAGGGTTAAA	30
		RT_AS101_f2_303_A	ATTAGGAATACCACATCCYKCAGGGATAAA	30
		RT_AS101_f3_303_A	ATTAGGAATACCACATCCYKGAGGGTTAAA	30
		RT_AS101_f4_303_A	ATTAGGAATACCACATCCYKGAGGGATAAA	30
		RT_AS101_f5_303_A	ATTAGGAATACCACATCCYKCAGGGTTAGA	30
		RT_AS101_f6_303_A	ATTAGGAATACCACATCCYKCAGGGATAGA	30
		RT_AS101_f7_303_A	ATTAGGAATACCACATCCYKGAGGGTTAGA	30
		RT_AS101_f8_303_A	ATTAGGAATACCACATCCYKGAGGGATAGA	30
		RT_AS101_f9_303_A	ATTAGGAATACCACATCCYKCAGGGTTACA	30
		RT_AS101_f10_303_A	ATTAGGAATACCACATCCYKCAGGGATACA	30
		RT_AS101_f11_303_A	ATTAGGAATACCACATCCYKGAGGGTTACA	30
		RT_AS101_f12_303_A	ATTAGGAATACCACATCCYKGAGGGATACA	30
		RT_AS101_f13_303_A	ATTAGGAATACCACATCCYKCAGGGTTAAC	30
		RT_AS101_f14_303_A	ATTAGGAATACCACATCCYKCAGGGATAAC	30
		RT_AS101_f15_303_A	ATTAGGAATACCACATCCYKGAGGGTTAAC	30
		RT_AS101_f16_303_A	ATTAGGAATACCACATCCYKGAGGGATAAC	30
		RT_AS101_f17_303_A	ATTAGGAATACCACATCCYKCAGGGTTAGC	30
		RT_AS101_f18_303_A	ATTAGGAATACCACATCCYKCAGGGATAGC	30
		RT_AS101_f19_303_A	ATTAGGAATACCACATCCYKGAGGGTTAGC	30
		RT_AS101_f20_303_A	ATTAGGAATACCACATCCYKGAGGGATAGC	30
		RT_AS101_f21_303_A	ATTAGGAATACCACATCCYKCAGGGTTACC	30
		RT_AS101_f22_303_A	ATTAGGAATACCACATCCYKCAGGGATACC	30
		RT_AS101_f23_303_A	ATTAGGAATACCACATCCYKGAGGGTTACC	30
		RT_AS101_f24_303_A	ATTAGGAATACCACATCCYKGAGGGATACC	30
		RT_AS101_f25_303_A	ATTAGGAATACCACATCCYKCAGGGTTAAG	30
		RT_AS101_f26_303_A	ATTAGGAATACCACATCCYKCAGGGATAAG	30
		RT_AS101_f27_303_A	ATTAGGAATACCACATCCYKGAGGGTTAAG	30
		RT_AS101_f28_303_A	ATTAGGAATACCACATCCYKGAGGGATAAG	30
		RT_AS101_f29_303_A	ATTAGGAATACCACATCCYKCAGGGTTAGG	30
		RT_AS101_f30_303_A	ATTAGGAATACCACATCCYKCAGGGATAGG	30
		RT_AS101_f31_303_A	ATTAGGAATACCACATCCYKGAGGGTTAGG	30
		RT_AS101_f32_303_A	ATTAGGAATACCACATCCYKGAGGGATAGG	30
		RT_AS101_f33_303_A	ATTAGGAATACCACATCCYKCAGGGTTACG	30
		RT_AS101_f34_303_A	ATTAGGAATACCACATCCYKCAGGGATACG	30
		RT_AS101_f35_303_A	ATTAGGAATACCACATCCYKGAGGGTTACG	30
		RT_AS101_f36_303_A	ATTAGGAATACCACATCCYKGAGGGATACG	30
103	307	RT_AS103_r1_307_T	CCCACATCCAGTACTGTYAYTGATTTTT	28
		RT_AS103_r2_307_T	CCCACATCCAGTACTGTYAYTGATTTTC	28
		RT_AS103_r3_307_T	CCCACATCCAGTACTGTYAYTGATTTTG	28
		RT_AS103_r4_307_T	CCCACATCCAGTACTGTYAYTGATTTGT	28
		RT_AS103_r5_307_T	CCCACATCCAGTACTGTYAYTGATTTGC	28
		RT_AS103_r6_307_T	CCCACATCCAGTACTGTYAYTGATTTGG	28
		RT_AS103_r7_307_T	CCCACATCCAGTACTGTYAYTGATTTAT	28
		RT_AS103_r8_307_T	CCCACATCCAGTACTGTYAYTGATTTAC	28
		RT_AS103_r9_307_T	CCCACATCCAGTACTGTYAYTGATTTAG	28
		RT_AS103_r10_307_T	CCCACATCCAGTACTGTYGYTGATTTTT	28

		RT_AS103_r11_307_T	CCCACATCCAGTACTGTYGYTGATTTTC	28
		RT_AS103_r12_307_T	CCCACATCCAGTACTGTYGYTGATTTTG	28
		RT_AS103_r13_307_T	CCCACATCCAGTACTGTYGYTGATTTGT	28
		RT_AS103_r14_307_T	CCCACATCCAGTACTGTYGYTGATTTGC	28
		RT_AS103_r15_307_T	CCCACATCCAGTACTGTYGYTGATTTGG	28
		RT_AS103_r16_307_T	CCCACATCCAGTACTGTYGYTGATTTAT	28
		RT_AS103_r17_307_T	CCCACATCCAGTACTGTYGYTGATTTAC	28
		RT_AS103_r18_307_T	CCCACATCCAGTACTGTYGYTGATTTAG	28
308		RT_AS103_r1_308_T	CCCACATCCAGTAYTGTYACTGATTTT	27
		RT_AS103_r2_308_T	CCCACATCCAGTAYTGTYACTGATTTG	27
		RT_AS103_r3_308_T	CCCACATCCAGTAYTGTYACTGATTTA	27
		RT_AS103_r4_308_T	CCCACATCCAGTAYTGTYATTGATTTT	27
		RT_AS103_r5_308_T	CCCACATCCAGTAYTGTYATTGATTTG	27
		RT_AS103_r6_308_T	CCCACATCCAGTAYTGTYATTGATTTA	27
		RT_AS103_r7_308_T	CCCACATCCAGTAYTGTYGCTGATTTT	27
		RT_AS103_r8_308_T	CCCACATCCAGTAYTGTYGCTGATTTG	27
		RT_AS103_r9_308_T	CCCACATCCAGTAYTGTYGCTGATTTA	27
		RT_AS103_r10_308_T	CCCACATCCAGTAYTGTYGTTGATTTT	27
		RT_AS103_r11_308_T	CCCACATCCAGTAYTGTYGTTGATTTG	27
		RT_AS103_r12_308_T	CCCACATCCAGTAYTGTYGTTGATTTA	27
309		RT_AS103_r1_309_T	CCCACATCYAGTAYTGTTACTGATTT	26
		RT_AS103_r2_309_T	CCCACATCYAGTAYTGTTATTGATTT	26
		RT_AS103_r3_309_T	CCCACATCYAGTAYTGTTGCTGATTT	26
		RT_AS103_r4_309_T	CCCACATCYAGTAYTGTTGTTGATTT	26
		RT_AS103_r5_309_T	CCCACATCYAGTAYTGTCACTGATTT	26
		RT_AS103_r6_309_T	CCCACATCYAGTAYTGTCATTGATTT	26
		RT_AS103_r7_309_T	CCCACATCYAGTAYTGTCGCTGATTT	26
		RT_AS103_r8_309_T	CCCACATCYAGTAYTGTCGTTGATTT	26
106	316	RT_AS106_r1_316_C	ATGCATCACCCACATCYAGTAYTGTTA	27
		RT_AS106_r2_316_C	ATGCATCACCCACATCYAGTAYTGTTG	27
		RT_AS106_r3_316_C	ATGCATCACCCACATCYAGTAYTGTC	27
		RT_AS106_r4_316_C	ATGCATCACCCACATCYAGTAYTGTCG	27
	317	RT_AS106_r1_317_A	ATGCATCACCCACATCYAGTACTGTT	26
		RT_AS106_r2_317_A	ATGCATCACCCACATCYAGTACTGTC	26
		RT_AS106_r3_317_A	ATGCATCACCCACATCYAGTATTGTT	26
		RT_AS106_r4_317_A	ATGCATCACCCACATCYAGTATTGTC	26
	318	RT_AS106_r1_318_T	AAWATGCATCACCCACATCCAGTACTGT	28
		RT_AS106_r2_318_T	AAWATGCATCACCCACATCCAGTATTGT	28
		RT_AS106_r3_318_T	AAWATGCATCACCCACATCTAGTACTGT	28
		RT_AS106_r4_318_T	AAWATGCATCACCCACATCTAGTATTGT	28
108	322	RT_AS108_r1_322_C	AAWATGCATCACCCACATCCAGTA	24
		RT_AS108_r2_322_C	AAWATGCATCACCCACATCTAGTA	24
	323	RT_AS108_r1_323_A	AAWATGCATCACCCACATCCAGT	23
		RT_AS108_r2_323_A	AAWATGCATCACCCACATCTAGT	23
	324	RT_AS108_r1_324_T	AAWATGCATCACCCACATCCAG	22
		RT_AS108_r2_324_T	AAWATGCATCACCCACATCTAG	22
115	343	RT_AS115_f1_343_T	ACARTACTRGATGTGGGTGATGCA	24
	344	RT_AS115_f1_344_A	ACARTACTRGATGTGGGTGATGCAT	25
	345	RT_AS115_f1_345_T	ACARTACTRGATGTGGGTGATGCATA	26
		RT_AS115_f2_345_T	ACARTACTRGATGTGGGTGATGCATT	26
116	346	RT_AS116_r1_346_A	CTGAAKTCTTYATCTAAGGGAAGTAAAA	28
		RT_AS116_r2_346_A	CTGAAKTCTTYATCTAAGGGAAGTGAAT	28
		RT_AS116_r3_346_A	CTGAAKTCTTYATCTAAGGGAAGTGAAGA	28

		RT_AS116_r4_346_A	CTGAAKTCTTYATCTAAGGGAAGT	28
		RT_AS116_r5_346_A	CTGAAKTCTTYATCTAAGGGAATTGAAA	28
		RT_AS116_r6_346_A	CTGAAKTCTTYATCTAAGGGAATTGAAT	28
		RT_AS116_r7_346_A	CTGAAKTCTTYATCTAAGGGAATTGAGA	28
		RT_AS116_r8_346_A	CTGAAKTCTTYATCTAAGGGAATTGAGT	28
347		RT_AS116_r1_347_A	CTGAAKTCTTYATCTAAGGGAAGT	27
		RT_AS116_r2_347_A	CTGAAKTCTTYATCTAAGGGAAGT	27
		RT_AS116_r3_347_A	CTGAAKTCTTYATCTAAGGGAATTGAA	27
		RT_AS116_r4_347_A	CTGAAKTCTTYATCTAAGGGAATTGAG	27
348		RT_AS116_r1_348_A	CTGAAKTCTTYATCTAAGGGAAGT	26
		RT_AS116_r2_348_A	CTGAAKTCTTYATCTAAGGGAATTGAA	26
118	352	RT_AS118_r1_352_C	TATACTTYCTGAAKTCTTTATCTAAGGGAA	30
		RT_AS118_r2_352_C	TATACTTYCTGAAKTCTTCATCTAAGGGAA	30
	353	RT_AS118_r_353_A	GTATACTTYCTGAAKTCTTTATCTAAGGGAA	30
	354	RT_AS118_r1_354_A	AGTATACTTYCTGAAKTCTTTATCTAAGGG	30
		RT_AS118_r2_354_A	AGTATACTTYCTGAAKTCTTCATCTAAGGG	30
151	451	RT_AS151_r1_451_G	GGAATATTGCTGGTGATCCTTTCCATCCCT	30
		RT_AS151_r2_451_G	GGAATATTGCTGGTGATCCTTTCCATCCCA	30
	452	RT_AS151_r1_452_T	TGGAATATTGCTGGTGATCCTTTCCATCCC	30
	453	RT_AS151_r1_453_C	TTGGAATATTGCTGGTGATCCTTTCCATCC	30
179	353	RT_AS179_f1_535_G	GAGCCTTTTAGAAAAACAAAATCCAGACATA	30
		RT_AS179_f2_535_G	GAGCCTTTTAGAAAAACAAAATCCAGAAATA	30
	356	RT_AS179_f1_536_T	AGCCTTTTAGAAAAACAAAATCCAGACATAG	30
		RT_AS179_f2_536_T	AGCCTTTTAGAAAAACAAAATCCAGAAATAG	30
	357	RT_AS179_f1_537_C	GCCTTTTAGAAAAACAAAATCCAGACATAGT	30
		RT_AS179_f2_537_C	GCCTTTTAGAAAAACAAAATCCAGAAATAGT	30
181	451	RT_AS181_f1_541_T	TTTAGAAAAACAAAATCCAGAMATAGTCATC	30
		RT_AS181_f2_541_T	TTTAGAAAAACAAAATCCAGAMATAGTTATC	30
	452	RT_AS181_f1_542_A	TTAGAAAAACAAAATCCAGAMATAGTTATCT	30
		RT_AS181_f2_542_A	TTAGAAAAACAAAATCCAGAMATAGTCATCT	30
	453	RT_AS181_f1_543_T	TAGAAAAACAAAATCCAGAMATAGTTATCTA	30
		RT_AS181_r1_543_A	TACHTACAAATCATCIATGTATTG	24
		RT_AS181_r2_543_A	TACHTACAAATCATCIACGTATTG	24
184	550	RT_AS184_f1_550_A	CAAAATCCAGAMATAGTYATCTATCAATAC	30
		RT_AS184_f2_550_A	CAAAATCCAGAMATAGTYATCTGTCAATAC	30
	551	RT_AS184_f1_551_T	AAAATCCAGAMATAGTYATCTATCAATACA	30
		RT_AS184_f2_551_T	AAAATCCAGAMATAGTYATCTATCAATACG	30
		RT_AS184_f3_551_T	AAAATCCAGAMATAGTYATCTGTCAATACA	30
		RT_AS184_f4_551_T	AAAATCCAGAMATAGTYATCTGTCAATACG	30
	552	RT_AS184_f1_552_G	AAATCCAGAMATAGTYATCTATCAATACAT	30
		RT_AS184_f2_552_G	AAATCCAGAMATAGTYATCTATCAATACGT	30
		RT_AS184_f3_552_G	AAATCCAGAMATAGTYATCTGTCAATACAT	30
		RT_AS184_f4_552_G	AAATCCAGAMATAGTYATCTGTCAATACGT	30
188	562	RT_AS188_f1_562_T	ATAGTTATCTRTCAATACRTGGATGATTTG	30
		RT_AS188_f2_562_T	ATAGTTATCTRTCAATACRTTGATGATTTG	30
		RT_AS188_f3_562_T	ATAGTTATCTRTCAATACRTCGATGATTTG	30
		RT_AS188_f4_562_T	ATAGTTATCTRTCAATACRTAGATGATTTG	30
	563	RT_AS188_f1_563_A	TAGTTATCTRTCAATACRTGGATGATTTGT	30
		RT_AS188_f2_563_A	TAGTTATCTRTCAATACRTTGATGATTTGT	30
		RT_AS188_f3_563_A	TAGTTATCTRTCAATACRTCGATGATTTGT	30
		RT_AS188_f4_563_A	TAGTTATCTRTCAATACRTAGATGATTTGT	30

		RT_AS188_f5_563_A	TAGTTATCTRTCAATACRTGGATGATTTGC	30
		RT_AS188_f6_563_A	TAGTTATCTRTCAATACRTTGATGATTTGC	30
		RT_AS188_f7_563_A	TAGTTATCTRTCAATACRTCGATGATTTGC	30
		RT_AS188_f8_563_A	TAGTTATCTRTCAATACRTAGATGATTTGC	30
564		RT_AS188_f1_564_T	AGTTATCTRTCAATACATIGATGATTTGTA	30
		RT_AS188_f2_564_T	AGTTATCTRTCAATACATIGATGATTTGTT	30
		RT_AS188_f3_564_T	AGTTATCTRTCAATACATIGATGATTTGCA	30
		RT_AS188_f4_564_T	AGTTATCTRTCAATACATIGATGATTTGCT	30
		RT_AS188_f5_564_T	AGTTATCTRTCAATACGTIGATGATTTGTA	30
		RT_AS188_f6_564_T	AGTTATCTRTCAATACGTIGATGATTTGTT	30
		RT_AS188_f7_564_T	AGTTATCTRTCAATACGTIGATGATTTGCA	30
		RT_AS188_f8_564_T	AGTTATCTRTCAATACGTIGATGATTTGCT	30
190	568	RT_AS190_r1_568_C	TCTATGYTGYYCTATTTCTAAGTCAGATC	29
		RT_AS190_r2_568_C	TCTATGYTGYYCTATTTCTAAGTCAGATG	29
		RT_AS190_r3_568_C	TCTATGYTGYYCTATTTCTAAGTCAGATT	29
		RT_AS190_r4_568_C	TCTATGYTGYYCTATTTCTAAGTCAGAAC	29
		RT_AS190_r5_568_C	TCTATGYTGYYCTATTTCTAAGTCAGAAG	29
		RT_AS190_r6_568_C	TCTATGYTGYYCTATTTCTAAGTCAGAAT	29
569		RT_AS190_r1_569_C	TCTATGYTGYCCTATTTCTAAGTCAGAT	28
		RT_AS190_r2_569_C	TCTATGYTGYCCTATTTCTAAGTCAGAA	28
		RT_AS190_r3_569_C	TCTATGYTGYTCTATTTCTAAGTCAGAT	28
		RT_AS190_r4_569_C	TCTATGYTGYTCTATTTCTAAGTCAGAA	28
570		RT_AS190_r1_570_T	TCTATGYTGYCCTATTTCTAAGTCAGA	27
		RT_AS190_r2_570_T	TCTATGYTGYTCTATTTCTAAGTCAGA	27
210	628	RT_AS210_f1_628_T	CAAAAATAGAGGAACTGAGACAACATCTG	29
		RT_AS210_f2_628_T	CAAAAATAGAGGAACTGAGAGAACATCTG	29
629		RT_AS210_f1_629_T	CAAAAATAGAGGAACTGAGASAACATCTGT	30
		RT_AS210_f2_629_T	CAAAAATAGAGGAACTGAGASAACATCTGG	30
630		RT_AS210_f1_630_G	AAAAATAGAGGAACTGAGASAACATCTGTT	30
		RT_AS210_f2_630_G	AAAAATAGAGGAACTGAGASAACATCTGTG	30
		RT_AS210_f3_630_G	AAAAATAGAGGAACTGAGASAACATCTGGT	30
		RT_AS210_f4_630_G	AAAAATAGAGGAACTGAGASAACATCTGGG	30
215	643	RT_AS215_f1_643_A	AACATCTGKTIAGGTGGGGATTT	23
		RT_AS215_f2_643_A	AACATCTGKTIAGGTGGGGGTTT	23
		RT_AS215_f3_643_A	AACATCTGKTI AAGTGGGGATTT	23
		RT_AS215_f4_643_A	AACATCTGKTI AAGTGGGGGTTT	23
		RT_AS215_f5_643_A	AACATCTGKGIAGGTGGGGATTT	23
		RT_AS215_f6_643_A	AACATCTGKGIAGGTGGGGGTTT	23
		RT_AS215_f7_643_A	AACATCTGKGI AAGTGGGGATTT	23
		RT_AS215_f8_643_A	AACATCTGKGI AAGTGGGGGTTT	23
644		RT_AS215_f1_644_C	AACATCTGKTIAGGTGGGGATTTA	24
		RT_AS215_f2_644_C	AACATCTGKTIAGGTGGGGGTTTA	24
		RT_AS215_f3_644_C	AACATCTGKTI AAGTGGGGATTTA	24
		RT_AS215_f4_644_C	AACATCTGKTI AAGTGGGGGTTTA	24
		RT_AS215_f5_644_C	AACATCTGKGIAGGTGGGGATTTA	24
		RT_AS215_f6_644_C	AACATCTGKGIAGGTGGGGGTTTA	24
		RT_AS215_f7_644_C	AACATCTGKGI AAGTGGGGATTTA	24
		RT_AS215_f8_644_C	AACATCTGKGI AAGTGGGGGTTTA	24
		RT_AS215_f9_644_C	AACATCTGKTIAGGTGGGGATTTT	24
		RT_AS215_f10_644_C	AACATCTGKTIAGGTGGGGGTTTT	24
		RT_AS215_f11_644_C	AACATCTGKTI AAGTGGGGATTTT	24
		RT_AS215_f12_644_C	AACATCTGKTI AAGTGGGGGTTTT	24
		RT_AS215_f13_644_C	AACATCTGKGIAGGTGGGGATTTT	24
		RT_AS215_f14_644_C	AACATCTGKGIAGGTGGGGGTTTT	24

		RT_AS215_f15_644_C	AACATCTGKGI AAGTGGGGATTTT	24
		RT_AS215_f16_644_C	AACATCTGKGI AAGTGGGGGTTTT	24
		RT_AS215_f17_644_C	AACATCTGKTI AAGTGGGGATTTG	24
		RT_AS215_f18_644_C	AACATCTGKTI AAGTGGGGGTTTG	24
		RT_AS215_f19_644_C	AACATCTGKTI AAGTGGGGATTTG	24
		RT_AS215_f20_644_C	AACATCTGKTI AAGTGGGGGTTTG	24
		RT_AS215_f21_644_C	AACATCTGKGI AAGTGGGGATTTG	24
		RT_AS215_f22_644_C	AACATCTGKGI AAGTGGGGGTTTG	24
		RT_AS215_f23_644_C	AACATCTGKGI AAGTGGGGATTTG	24
		RT_AS215_f24_644_C	AACATCTGKGI AAGTGGGGGTTTG	24
645		RT_AS215_f1_645_C	AACATCTGT KIARGTGGGGATTTAC	25
		RT_AS215_f2_645_C	AACATCTGT KIARGTGGGGATTTAA	25
		RT_AS215_f3_645_C	AACATCTGT KIARGTGGGGATTTAT	25
		RT_AS215_f4_645_C	AACATCTGT KIARGTGGGGATTTAG	25
		RT_AS215_f5_645_C	AACATCTGT KIARGTGGGGATTTTC	25
		RT_AS215_f6_645_C	AACATCTGT KIARGTGGGGATTTTA	25
		RT_AS215_f7_645_C	AACATCTGT KIARGTGGGGATTTTTT	25
		RT_AS215_f8_645_C	AACATCTGT KIARGTGGGGATTTTG	25
		RT_AS215_f9_645_C	AACATCTGT KIARGTGGGGATTTGC	25
		RT_AS215_f10_645_C	AACATCTGT KIARGTGGGGATTTGA	25
		RT_AS215_f11_645_C	AACATCTGT KIARGTGGGGATTTGT	25
		RT_AS215_f12_645_C	AACATCTGT KIARGTGGGGATTTGG	25
		RT_AS215_f13_645_C	AACATCTGT KIARGTGGGGGTTTAC	25
		RT_AS215_f14_645_C	AACATCTGT KIARGTGGGGGTTTAA	25
		RT_AS215_f15_645_C	AACATCTGT KIARGTGGGGGTTTAT	25
		RT_AS215_f16_645_C	AACATCTGT KIARGTGGGGGTTTAG	25
		RT_AS215_f17_645_C	AACATCTGT KIARGTGGGGGTTTTC	25
		RT_AS215_f18_645_C	AACATCTGT KIARGTGGGGGTTTTA	25
		RT_AS215_f19_645_C	AACATCTGT KIARGTGGGGGTTTTT	25
		RT_AS215_f20_645_C	AACATCTGT KIARGTGGGGGTTTTG	25
		RT_AS215_f21_645_C	AACATCTGT KIARGTGGGGGTTTGC	25
		RT_AS215_f22_645_C	AACATCTGT KIARGTGGGGGTTTGA	25
		RT_AS215_f23_645_C	AACATCTGT KIARGTGGGGGTTTGT	25
		RT_AS215_f24_645_C	AACATCTGT KIARGTGGGGGTTTGG	25
219	655	RT_AS219_f_655_A	CTGTTGAGGTGGGGATTTACCACACCAGAC	30
	656	RT_AS219_f_656_A	TGTTGAGGTGGGGATTTACCACACCAGACA	30
	657	RT_AS219_f_657_A	GTTGAGGTGGGGATTTACCACACCAGACAA	30
225	673	RT_AS225_r_673_G	GGAGTTCATAACCCATCCAAAGGAATGGAG	30
	674	RT_AS225_r_674_G	TGGAGTTCATAACCCATCCAAAGGAATGGA	30
	675	RT_AS225_r_675_A	ATGGAGTTCATAACCCATCCAAAGGAATGG	30
227	379	RT_AS227_r_379_A	CAGGATGGAGTTCATAACCCATCCAAAGGA	30
	380	RT_AS227_r_380_A	TCAGGATGGAGTTCATAACCCATCCAAAGG	30
	381	RT_AS227_r_381_G	ATCAGGATGGAGTTCATAACCCATCCAAAG	30
230	688	RT_AS230_r1_688_T	TATCARGATGGAGTTCATAACCCA	24
	689	RT_AS230_r1_689_A	GTCCADITATCAGGATGGAGTTCATAACCC	30
		RT_AS230_r2_689_A	GTCCADITATCAAGATGGAGTTCATAACCC	30
	690	RT_AS230_r1_690_C	TGTCCADITATCAGGATGGAGTTCATAACC	30
		RT_AS230_r2_690_C	TGTCCADITATCAAGATGGAGTTCATAACC	30
236	706	RT_AS236_r1_706_T	TATAGGCTGTACTGTCCATTTATCAG	26
		RT_AS236_f1_706_C	YTKGATGGGTTATGAACTCCAT	22
		RT_AS236_f2_706_C	YTKGCTGGGTTATGAACTCCAT	22
		RT_AS236_f3_706_C	YTKTATGGGTTATGAACTCCAT	22

		RT_AS236_f4_706_C	YTKTCTGGGTTATGAACTCCAT	22
707		RT_AS236_r1_707_A	AGCYTTATAGGCTGTACTGTCCATTTATCA	30
		RT_AS236_f1_707_C	TKGATGGGTTATGAACTCCATC	22
		RT_AS236_f2_707_C	TKGCTGGGTTATGAACTCCATC	22
		RT_AS236_f3_707_C	TKTATGGGTTATGAACTCCATC	22
708		RT_AS236_r1_708_C	CAGCYTTATAGGCTGTACTGTCCATTTATC	30
		RT_AS236_f4_707_C	TKTCTGGGTTATGAACTCCATC	22
		RT_AS236_f1_708_T	TKKATGGGTTATGAACTCCATCC	23
		RT_AS236_f2_708_T	TKKATGGGTTATGAACTCCATCT	23
		RT_AS236_f3_708_T	TKKCTGGGTTATGAACTCCATCC	23
		RT_AS236_f4_708_T	TKKCTGGGTTATGAACTCCATCT	23
238	712	RT_AS238_r1_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCATT	30
		RT_AS238_r2_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCATC	30
		RT_AS238_r3_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCATG	30
		RT_AS238_r4_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAAT	30
		RT_AS238_r5_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAAC	30
		RT_AS238_r6_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAAG	30
		RT_AS238_r7_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAGT	30
		RT_AS238_r8_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAGC	30
		RT_AS238_r9_712_T	CTGGCAGCWYTATAGGCTGTACTGTCCAGG	30
713		RT_AS238_r1_713_T	TCTGGCAGCWYTATAGGCTGTACTGTCCAT	30
		RT_AS238_r2_713_T	TCTGGCAGCWYTATAGGCTGTACTGTCCAA	30
		RT_AS238_r3_713_T	TCTGGCAGCWYTATAGGCTGTACTGTCCAG	30
714		RT_AS238_r1_714_T	TTCTGGCAGCWYTATAGGCTGTACTGTCCA	30

Für die Validierung verwendete Patientenproben

Die Patientenproben wurden vom Hygieneinstitut der Universität Heidelberg zur Verfügung gestellt (Abteilung Virologie). Sie sind mit zweistelligen Buchstabencodes bezeichnet, die zugehörigen Sequenzen lagen für die Validierung vor. Die Subtypenbestimmung wurde unter <http://dbpartners.stanford.edu/RegaSubtyping/> nach Oliveira *et al.* (2005) vorgenommen.

Patientencode	Proben Nr. PR-Chip	Proben Nr. RT-Chip	Subtyp / CRF
IN	1	1	B
IO	2		B
IQ	3	2	B
IR	4		A
IS	5	3	B
IT	6		nicht bestimmt
IU	7	4	B
JB	8		B
JD	9		A
JF	10		CRF02_AG
JG	11		B
JJ2	12	5	B
JM	13	6	CRF02_AG
JO	14	7	B
JQ	15		B
JR	16		B
JT2	17		B
JV	18		B
JW	19	8	B
JX	20		CRF02_AG
KA	21		B
KB	22		B
KH	23		B
KI2	24	9	B
KN	25		CRF06_cpx
KO2	26		B
KQ	27		B
KS	28		B
KT	29	10	B
KU	30	11	B
KW	31	12	B
KX	32	13	B
KY	33		B
LJ	34	14	B
LK	35	15	B
LN	36	16	B
LO	37		C
LP	38		B
LQ	39	17	B
LS	40		A
LT	41	18	CRF01_AE
LU	42	19	B

LV	43	20	B
LX	44	21	B
CN	45	22	C
DP	46		B
DQ	47	23	B
EN	48		B
EO	49	24	B
EQ	50		B
ER	51		B
ET	52		C
EV	53		CRF02_AG
EX	54	25	nicht bestimmt
FA		26	B
FE	55		B
FF	56	27	B
FG	57		nicht bestimmt
FI	58		CRF02_AG
FO	59		B
FQ	60		B
FT	61	28	B
FU	62	29	B
FV	63	30	B
FW	64	31	B
FZ	65		CRF02_AG
GB	66		B
GK2	67		B
GL2	68		B
GM	69		B
GP	70		nicht bestimmt
GR	71	32	B
GS	72	33	B
GT	73	34	B
GV	74	35	B
GW	75	36	CRF01_AE
GX	76	37	B
GY	77	38	B
HA	78	39	B
HB	79	40	B
HC	80		B
IC	81		B
ID	82	41	F1
IF	83	42	B
IG	84	43	B
II	85		nicht bestimmt
IK		44	nicht bestimmt
IL		45	B
CJ	86	46	B
QB	87		B
QC	88	47	B
QD	89	48	B
QE	90		B
QF	91		nicht bestimmt
NL4-3(WT)	92		B
NL4-3(WT)	93		B
NL4-3(WT)	94		B