

5 Diskussion

5.1 SNPs als Marker für Kopplungs- und Assoziationsstudien: Anforderungen und Schwierigkeiten

Das humane Genom besteht aus etwa 3 Milliarden Basenpaaren, und ein Einzelbasenpolymorphismus tritt durchschnittlich alle 1000 bp auf. Die Sequenzierung der menschlichen DNS-Sequenz durch das internationale "Genome Sequencing Consortium" erbrachte 1,42 Millionen SNPs, und durch Hochrechnungen wurde eine Zahl von insgesamt 1,6 – 3,2 Millionen SNPs ermittelt (Sachidanandam et al., 2001; Stoneking, 2001). Ein anderer Ansatz der Sequenzierung des menschlichen Genoms durch die Firma "Celera Genomics" ergab 2,1 Millionen SNPs (Venter et al., 2001).

SNPs stellen den größten Anteil an Sequenzunterschieden zwischen Individuen dar, und ihre Bedeutung für phänotypische Unterschiede zwischen Menschen ist von ihrer Position im Genom abhängig. SNPs in kodierenden Regionen können die Funktion oder Struktur eines Proteins beeinflussen, z.B. sind zwei Polymorphismen in der kodierenden Region des Apolipoprotein-E-Gens mit Arteriosklerose assoziiert, und ein Polymorphismus im Gerinnungsfaktor-V-Gen mit einem erhöhten Thromboserisiko (Davignon et al., 1988; Bertina RM et al., 1994). Es ist anzunehmen, dass auch SNPs in regulatorischen Genregionen phänotypische Auswirkungen haben. Die meisten SNPs liegen jedoch in nicht-kodierenden Regionen des Genoms. Ihre Häufigkeit, Stabilität und relativ regelmäßige Verteilung über das gesamte Genom machen sie zu wertvollen genetischen Markern. Von groß angelegten Assoziationsstudien mit vielen, über das Genom verteilten SNPs erhofft man sich, Bereiche zu identifizieren, in denen sich SNP-Allele zwischen Patientengruppen und Kontrollgruppen unterscheiden. In der Annahme, dass diese SNP-Allele zusammen mit dem Krankheits-prädisponierenden Allel vererbt werden, sollen auf diese Weise Krankheitsgene lokalisiert und identifiziert werden. Wissenschaftler debattieren derzeit intensiv, wie viele SNPs für solche genomweiten

Analysen benötigt werden. Bisherige Schätzungen reichen von 30000 bis 1 Million SNPs (Roberts et al., 2000; Kruglyak, 1999; Collins et al., 1999).

Neuere Untersuchungen belegen jedoch, dass das menschliche Genom in Abschnitte eingeteilt werden kann, innerhalb derer eine geringe Haplotyp-Diversität besteht (Patil et al., 2001). D.h., dass SNP-Allele innerhalb solcher Abschnitte korrelieren, so dass jeder Abschnitt durch wenige SNPs charakterisiert werden kann. Für genomweite Assoziationsstudien würden demnach nicht gleichmäßig über das Genom verteilte SNPs benötigt, sondern jeweils wenige SNPs pro Haplotyp-Abschnitt. Die Abschnitte müssten aber zunächst identifiziert werden, wofür ein sehr dichtes Raster an SNPs benötigt wird. Diese Information könnte dann verwendet werden, um eine viel kleinere Zahl an SNP-Markern zu definieren. Eine Alternative zu genomweiten Assoziationsstudien stellt die Analyse von SNPs innerhalb oder in der Nachbarschaft von Kandidatengenen dar. Solche Untersuchungen sind deutlich weniger aufwendig, und waren bereits in einigen Fällen erfolgreich. Z.B. ist es mehreren Forscherteams gelungen, ein erstes Krankheitsgen (NOD2) für die Morbus Crohn Krankheit auf Chromosom 16q12 zu identifizieren (Hugot et al., 2001; Hampe et al., 2001). Mit Hilfe der positionalen Klonierung, basierend auf Kopplungsanalysen und die Erstellung von Karten für Kopplungsungleichgewichte, konnten drei genetische Varianten charakterisiert werden. Es konnte gezeigt werden, dass die Insertionsmutation (3020insC), welche zu einer Leserasterverschiebung führt, und zwei Aminosäureaustausche (R702W und G908R) unabhängig voneinander stark mit der M. Crohn Krankheit assoziiert sind (Cuthbert et al., 2002; Lesage et al., 2002; Hampe et al., 2002). Desweiteren ist durch den Einsatz dichter genetischer Karten von Mikrosatelliten-Marker und SNPs gelungen, ein weiteres Gen (IBD5) auf Chromosomen 5 zu identifizieren, welches für die Entstehung von Morbus Crohn mitverantwortlich sein könnte (Rioux et al., 2001).

Weiterhin ist bislang unklar, wie groß das Patienten- und Kontrollkollektiv für Assoziationsstudien sein muss. Die statistische Signifikanz einer Korrelation zwischen einem Einzelbasen-Polymorphismus und einem bestimmten Phänotyp hängt davon ab, wie häufig der betreffende SNP in der Bevölkerung vorkommt und wie viele Personen untersucht wurden (McCarthy et al., 2000). Ausserdem hängt die erforderliche Zahl an

zu untersuchenden Personen von weiteren den Phänotyp beeinflussenden Faktoren wie z.B. Umweltfaktoren ab.

Die große Zahl an zu untersuchenden SNPs bei sehr vielen Personen stellt enorme Anforderungen an Techniken zur SNP-Typisierung. Das wichtigste Ziel von Biotechnologie-Firmen und Forschungsinstituten, welche SNPs als Marker einsetzen wollen, besteht derzeit darin, preisgünstige Methoden der SNP-Typisierung mit erhöhtem Durchsatz zu entwickeln.

Mit der DNS-Mikroraster-Technologie können auf einer kleinen Fläche viele Reaktionen von DNS-Einzelsträngen gleichzeitig durchgeführt werden. Sie ist deshalb besonders für Hochdurchsatzmessungen geeignet, und wurde in der vorliegenden Arbeit als Plattform für den Nachweis von Einzelbasenaustauschen gewählt.

Ein DNS-Chip hat im allgemeinen eine Fläche von 18 cm². Bei einem „spot“-Durchmesser von 200 µm (wie in dieser Arbeit) können somit theoretisch 10.000 Sonden auf einen DNS-Chip aufgebracht werden, also wesentlich mehr als bei der Verwendung von Mikrotiterplatten (typischerweise 96 oder 384 getrennte Vertiefungen, die jeweils eine Sonde enthalten), welche ebenfalls zur SNP-Typisierung verwendet werden. (Landegren et al., 1988; Nickersen et al., 1990; Samiotaki et al., 1994).

5.2 Prinzip der Allel-spezifischen Primer-Elongationsreaktion auf Oligonukleotid-Mikrorastern

DNS-Mikroraster sind kleine Hybridisierungsplattformen, mit denen die Ausbildung von spezifischen DNS-DNS oder DNS-RNS Doppelsträngen gemessen werden kann (Southern et al. 1999; Brown et al. 1999). Diese Technologie hat ihren Ursprung in der vergleichenden Messung von Transkripten bestimmter Gene. Dabei wird die Bindung Fluoreszenz-markierter cDNS-Moleküle von durchschnittlich 2 kb an cDNS-Sonden von 500-1500 bp, die auf dem Chip gebunden sind, gemessen. Einzelne Basenunterschiede zwischen der Sonde und des Targets haben hier aufgrund der Länge der DNS-Moleküle

keinen entscheidenden Einfluß auf die Hybridisierung ansonsten komplementärer DNS-Stränge. Die Stabilität von DNS-Doppelsträngen hängt v.a. von der Stringenz der Reaktionsbedingungen, der Gesamtlänge und des Basengehaltes komplementärer Abschnitte ab (Conner et al., 1983).

Je mehr Wasserstoffbrücken ausgebildet werden, d.h. je länger die Sequenz und je höher der Gehalt an Guanin oder Cytosin, desto stabiler der Doppelstrang. Eine Möglichkeit, den Einfluß einzelner Basenaustausche auf die Stabilität zu erhöhen, besteht darin, kürzere DNS-Sequenzen zu verwenden. Zum Nachweis einzelner Basenunterschiede in genomischer DNS durch Hybridisierung auf DNS-Chips werden deshalb kurze Oligonukleotide, z.B. 25mere, als gebundene Sonden eingesetzt (Hacia et al., 1996). Jedoch hat auch die Position eines Basenunterschiedes innerhalb einer Sequenz Einfluß auf die Stabilität von DNS-Doppelsträngen, z. B. haben zentrale Basenunterschiede einen stärker destabilisierenden Effekt als terminale, und auch die Basensequenz direkt neben dem nicht komplementären Nukleotids wirkt sich auf die Stabilität des Hybridmoleküls aus. Man muß mehrere Sequenzkombinationen, welche die zu untersuchende Basenposition flankieren, als Sonden einsetzen, und die Interpretation der Hybridisierungsergebnisse ist nur möglich durch Einbeziehung der Signale all dieser Sonden, was entsprechende mathematische Algorithmen erfordert (Hacia et al. 1998; Wang et al. 1998; Cho et al. 1999).

Ein anderes Prinzip zum Nachweis von Einzelbasenunterschieden nach einer Hybridisierung basiert auf der Fähigkeit von Enzymen, komplementäre Basenpaarungen eindeutig zu erkennen. In der vorliegenden Arbeit wurde dieses Prinzip mit einer Hybridisierung auf DNS-Chips kombiniert. Der Vorteil einer enzymatischen Erkennung von SNPs liegt zum einen in der verbesserten Diskriminierung von Basenunterschieden, zum anderen müssen nur zwei Allel-spezifische Oligonukleotide als Sonden eingesetzt werden (Syvänen et al., 1990; Pastinen et al., 1997; Ross et al., 1998; Dubilay et al., 1999; Pastinen et al., 2000).

Die in dieser Arbeit beschriebene, neue Methode zur Typisierung von SNPs stützt sich auf die Fähigkeit von DNS-Polymerasen, Einzelbasenaustausche in einer Elongations-Reaktion, von auf Chips immobilisierten Oligonukleotiden, zu unterscheiden. Für jeden

zu untersuchenden SNP wurden zwei Allel-spezifische Oligonukleotide auf Microrastern immobilisiert. Die Allel-spezifische Elongationsreaktion fand in einem Ansatz mit allen vier dNTPs und Cy3-dUTP in Anwesenheit einer DNS-Polymerase sowie eines komplementären DNS-Targets statt. Bei Komplementarität der 3'-terminalen Base resultiert dieser Ansatz in einem Extensions-Produkt, dessen Größe vom Abstand zwischen der variablen Base und dem 5'-Ende der Matrizensequenz abhängig ist (siehe Ergebnisse, Abb. 4.1). Um die Ausbeute des verlängerten Produktes zu steigern, wurden Zyklen von Denaturierungs-, Annealing- und Extensionsschritten durchgeführt.

5.3 Ermittlung geeigneter Allel-spezifischer Oligonukleotide

Voraussetzung der in dieser Arbeit vorgestellten Reaktion ist eine sequenzspezifische Anlagerung eines DNS-Targets an die komplementäre Oligonukleotid-Sonde. Für eine spezifische Hybridisierung des Targets mit der Sonde müssen beide in einem ausreichend langen Sequenzabschnitt zueinander komplementär sein. Die Anlagerung des Targets hängt aber auch von der Zugänglichkeit der am 5' Ende fixierten Sonde ab. Dies schliesst sehr wahrscheinlich die am weitesten 5'-wärts gelegenen Basen von der Hybridisierung aus; sie können deshalb durch einen unspezifischen Abstandshalter ersetzt werden. In dieser Arbeit wurden verschiedene Kombinationen an Abstandshalter und spezifischer Sequenz getestet (siehe Ergebnisse, Abb. 4.8). Eine spezifische Sequenz von 20 Basenpaaren ohne Abstandshalter erbrachte nur sehr schwache Signale. Mit Oligonukleotiden, die aus einem poly-T-Abstandshalter von 15 Basen und einer spezifischen Sequenz von 20 bp (Gesamtlänge von 35 bp) bestanden, wurden hingegen deutliche Signale erhalten (Ergebnisse Teil 4.6). Eine noch höhere Spezifität konnte jedoch durch den Einsatz von Oligonukleotiden mit 50bp langen spezifischen Sequenzen erreicht werden. Allel-spezifische Oligonukleotide, die bei ähnlichen Primer-Elongationsmethoden auf Chips verwendet wurden, wie z.B. der Minisequenzierungs-Reaktion, sind aus 9-15 bp polyT-Abstandshalter und 15-27 bp spezifischer Sequenz

zusammengesetzt und somit zu den in dieser Arbeit verwendeten vergleichbar (Pastinen et al., 1997; Pastinen et al., 2000; Raito et al., 2001).

5.4 Entwicklung einer geeigneten Chip-Oberfläche

Die Immobilisierung der Oligonukleotid-Sonden auf Glasoberflächen muss stabil und außerdem gerichtet sein, d.h. mit der Allel-spezifischen Base am freien 3'-Ende. Zu Beginn der Arbeit lag ein von Guo et al. publiziertes Protokoll vor, welches auf einer kovalenten Bindung des N`N Phenylendiisothiocyanat (PDITC) an die 5`Amino-Gruppe der Oligonukleotide beruht (siehe auch Methoden Abb. 3.1 und 3.2). In zwei anderen Arbeiten war diese Methode ebenfalls erfolgreich für die Bindung von Oligonukleotid-Sonden auf N`N Phenylendiisothiocyanat beschichteten Glasoberflächen in identischer Weise (Pastinen et al., 2000) bzw. geringfügig modifiziert (Beier et al., 2000) angewendet worden. In der vorliegenden Arbeit wurden sowohl bereits silanierte Objektträger von Perkin Elmer verwendet als auch unbehandeltes Glas nach einem etwas veränderten Protokoll selbst silaniert. Die wesentlichen von Guo et al. abweichenden Schritte betreffen die Reinigung, Silanisierungszeit und Silankonzentration. Eine gründlichere Säuberung der unbeschichteten Glasobjektträger mit 1M NaOH (siehe Methoden, Abschnitt 3.1.1) im Gegensatz zu 10 min in Aceton erwies sich besser für die nachfolgende Haftung des Amino-propyltrimethoxysilans. Ausserdem führte eine Verlängerung der Silanisierungszeit (30 min bei RT auf dem Schüttler in der vorliegenden Arbeit, anstelle 2 min bei RT nach Guo et al., 1994) zu höheren Signalintensitäten bei der Primer-Elongationsreaktion. Die verschiedenen Silankonzentrationen (1%,2%,3%) stellten sich als ein kritischer Parameter für die Vermeidung falsch positiver Signale heraus (siehe Ergebnisse, Abschnitt 4.4).

5.5 Falsch-positive Signale unter verschiedenen experimentellen Bedingungen

Primer-Elongationsreaktionen wurden generell mit und ohne Zusatz eines DNS-Targets durchgeführt, um Hintergrund-Signale bestimmen zu können. Sogar ohne ein DNS-Target wurden in einigen Fällen starke Signale erhalten. Aus diesen Ergebnissen wurde gefolgert, dass die Bildung von inter- oder intramolekularen DNS-Doppelsträngen der immobilisierten Oligonukleotide zu einer Elongationsreaktion führen können. Eine nähere Untersuchung der Sequenz immobilisierter Oligonukleotide zeigte, dass die Bildung von Oligonukleotid-Dimeren benachbarter Moleküle in einem „spot“ die wahrscheinlichste Erklärung ist. Bei 5 der 6 intensivsten falsch positiven Signale ergab die Sequenzanalyse der betreffenden Oligonukleotide die Möglichkeit einer Bildung von Oligonukleotid-Dimeren (siehe Ergebnisse, Abschnitt 4.4). Dieser Mechanismus, der zu falsch positiven Signalen führt, wurde bereits 1994 in einer Arbeit von Nikiforov et al. beschrieben. Wie in der vorliegenden Arbeit gezeigt, können die Intensitäten derartiger falsch positiver Signale durch Modifikationen der experimentellen Bedingungen beeinflusst werden. Um die Bedeutung der Oberflächenchemie für das Vorkommen falsch positiver Signale zu untersuchen, wurden 96 Allel-spezifische Oligonukleotide auf unterschiedliche Aminosilan-beschichtete Glasobjektträger immobilisiert. Die Silanisierung wurde mit 1, 2 und 3% Aminopropyltrimethoxysilan durchgeführt und ausserdem kommerziell erhältliche Aminosilan-Glasobjektträger getestet (siehe Materialien, Abschnitt 2.7 und Methoden, Abschnitt 3.1.1). Die Ergebnisse der Primer-Elongationsreaktion mit diesen vier verschiedenen Arten von Objektträgern und ohne das Hinzufügen eines DNS-Targets sind im Ergebnisteil Abschnitt 4.4 zu finden. Die Intensität und Anzahl der falsch positiven Signale nahm bei den selbst silanisierten Objektträgern mit zunehmender Konzentration an Aminopropyltrimethoxysilan zu. Die höchste Anzahl falsch positiver Signale wurde bei den kommerziell erhältlichen Objektträgern festgestellt. Kontrollexperimente, bei denen ein asymmetrisches PCR-Produkt als DNS-Target verwendet wurde (siehe Ergebnisteil, Abschnitt 4.9.2) zeigten, dass unter den gleichen experimentellen Bedingungen die beste Allel-Diskriminierung

unter Verwendung von 1% Aminopropyltrimethoxysilan-beschichteten Objektträgern erreicht wurde.

Es wurde ausserdem beobachtet, dass verschiedene Parameter des Elongations-Protokolls die Anzahl der falsch-positiven Signale beeinflussen. Auf 1% Aminopropyltrimethoxysilan beschichteten Objektträgern wurden 96 Allel-spezifische Oligonukleotide immobilisiert. Die Zahl der falsch-positiven Signale stieg

- mit zunehmender Zahl an Zyklen (25 Zyklen anstatt 15 Zyklen),
- mit abnehmender Annealing-Temperatur (48°C anstatt 56°C)
- mit verlängerter Annealing-Zeit (60sec. anstatt 30sec)
- und mit verlängerter Elongationszeit (30sec. anstatt 15 sec.)

Diese Ergebnisse sind in Abschnitt 4.4 dargestellt.

Unter Standardbedingungen war die Intensität falsch positiver Signale am geringsten. Unter diesen Bedingungen wurde außerdem die beste Unterscheidung von Allelen erreicht, wie der Einsatz anderer PCR-Produkte als Kontrolle zeigte. Generell können Allel-spezifische Oligonukleotide, welche sogar unter optimalen Bedingungen zu falsch positiven Signale führen, durch ihre komplementäre DNS-Sequenz ersetzt werden.

5.6 Ermittlung geeigneter Targets

Die größte Schwierigkeit bei der Typisierung von SNPs in großem Maßstab stellt die Target-Herstellung dar. Ideal wäre in diesem Zusammenhang die direkte Verwendung von genomischer DNS aus Blut von Patienten oder Probanden. Allerdings können durch derzeitige Methoden SNPs erst nach einer PCR-Amplifikation der entsprechenden genomischen Loci nachgewiesen werden, da die Sensitivität der Protokolle zu gering bzw. das menschliche Genom zu komplex ist. D.h. pro SNP muss eine PCR-Amplifikation mit diesen Locus-flankierenden spezifischen Primern durchgeführt werden,

und für die Typisierung tausender SNPs müssen ebensoviele PCR-Amplifikationen erfolgen.

Ein in dieser Arbeit verfolgter Ansatz war deshalb, die Komplexität des menschlichen Genoms zu reduzieren, indem nur Bereiche zwischen repetitiven Alu-Sequenzen in die Untersuchung einbezogen werden sollten (s. Abschnitt 4.14).

Alu-Sequenzen treten bevorzugt in genreichen (Korenberg und Engels, 1978) und in rekombinationsaktiven Bereichen des humanen Genoms (Kuhn und Therman, 1986) auf. Um genreiche Regionen im humanen Genom zu erhalten, müssen die Bereiche zwischen den Alu-Sequenzen amplifiziert werden. Hierzu werden in einer PCR-Reaktion Primer (C12) verwendet, die spezifisch an Alu-Sequenzen binden. Liegen zwei Alu-Wiederholungen nahe genug beieinander und in der passenden Orientierung vor, so kann die dazwischen liegende humane DNS-Sequenz, durch eine PCR amplifiziert werden.

SNPs in solchen Inter-Alu-Abschnitten im Genom von 6 Probanden wurden durch Sequenzierung in einer Kollaboration mit der Firma MetaGen identifiziert, entsprechende Allel-spezifische Oligonukleotide wurden entworfen und auf Chips immobilisiert.

Trotz Reduktion der Komplexität war die Menge der auf diese Weise hergestellten Targets immer noch zu gering, um Fluoreszenzsignale durch Primer-Elongationsreaktionen auf Chips zu erhalten (Ergebnisse Abschnitt 4.14).

5.6.1 Typisierung von Einzelbasenaustauschen mit synthetischen Oligonukleotid-Targets

Zu Beginn der Arbeit wurde eine Allel-spezifische PCR in Lösung durchgeführt, um die Fähigkeit der Taq-Polymerase zur Erkennung von 3'-ständigen Basenfehlpaarungen zu analysieren. Tatsächlich kann es nur bei Vorliegen einer korrekten Basenpaarung zur Amplifikation kommen, was durch gelelektrophoretische Auftrennung der PCR-Produkte nachgewiesen wurde (siehe Ergebnisse, Abschnitt 4.1). Im Anschluss wurden

synthetische 50mer Oligonukleotide als Target eingesetzt, um die Genauigkeit der von der Taq-Polymerase katalysierten Elongationsreaktion auf Oligonukleotid-Mikrorastern zu bestimmen. Als DNS-Target wurden vier verschiedene Oligonukleotide eingesetzt, die eine Länge von 50 Basen besaßen und deren Sequenz einem mtDNS Abschnitt entsprach. Diese vier Oligonukleotide unterschieden sich jeweils in der zentralen Base an Position 25. Vier Allel-spezifische Oligonukleotide der Länge von 26 bp wurden als Duplikate auf Objektträger immobilisiert. Ihre Sequenz war komplementär zu 25 bp aller vier Oligonukleotid-Sonden, sie unterschieden sich nur in ihrer Base am 3'-Ende, die entweder ein A, T, G oder C enthielt. Jedes 50mer Oligonukleotid-Target wurde in einer getrennten Primer-Elongationsreaktion auf diesem Mikroraster eingesetzt und die Signalintensitäten an der Position jeder Sonde (immobilisierte Oligonukleotide) gemessen. Die Signal-Verhältnisse zwischen korrekten Basenpaarungen zu Basenfehlpaarungen wurden ermittelt. Die relativen Signal-Intensitäten lagen dabei im Bereich von 1:10 bis 1:50, mit der höchsten Diskriminierung bei A(Sonde):G(Target), T:G und die niedrigste bei C:T Basenfehlpaarungen (siehe Ergebnisse, Abschnitt 4.7). Diese Ergebnisse sind analogen, in Lösung durchgeführten Primer-Elongationsexperimenten vergleichbar (Huang et al., 1992). Dabei wurde die Bindungsaffinität und Extensionszeit der Taq-Polymerase an komplementären und nicht komplementären DNS-Termini untersucht und die höchste Spezifität des Enzyms für A:G,- G:A,- G:G,- C:C,- A:A-Basenfehlpaarungen, die schlechteste Diskriminierung für bei C:T-Fehlpaarungen gefunden.

5.6.2 Detektion von Einzelbasenpolymorphismen mittels einzelsträngigen asymmetrischen PCR-Produkten als Target

Da bei Verwendung von Doppelstrang DNS-Sequenzen nur sehr schwache Signale resultierten, wurden Einzelstrang DNS-Targets auf drei verschiedenen Wegen hergestellt und getestet.

Eine Methode zur Herstellung von Einzelstrang-Nukleinsäuren ist die *in vitro* Transkription. Bei der *in vitro* Transkription wurde ein Strang der DNS in RNS umgeschrieben (transkribiert). Dieses Einzelstrang RNS-Produkt wurde als Target in der Allel-spezifischen Primer-Elongationsreaktion eingesetzt. Mittels dieses RNS-Targets konnten 5 "stromabwärts" liegende SNPs korrekt typisiert werden (siehe Ergebnisteil, Abschnitt 4.11). Eine andere Methode bestand in der Durchführung einer asymmetrischen PCR, bei der nur ein Oligonukleotid "oberstrom" der zu amplifizierenden Sequenz eingesetzt wird und somit Produkte unterschiedlicher Länge linear amplifiziert werden. Zur Herstellung einzelsträngiger SNP-enthaltender DNS-Targets wurden entsprechende mitochondriale DNS-Bereiche zunächst mittels einer konventionellen PCR amplifiziert und diese doppelsträngigen Produkte dann als Matrize in einer asymmetrischen PCR eingesetzt. Zur Typisierung mehrerer benachbarter SNPs wurde sowohl ein einziger Oligonukleotid-Primer oberhalb dieser Loci oder aber mehrere Oligonukleotid-Primer jeweils oberhalb der einzelnen Loci in einer asymmetrischen PCR eingesetzt, da die einfache Elongation in einem asymmetrischen PCR-Ansatz immer von dem eingesetzten Primer ausgeht und in unterschiedlicher Entfernung davon abbricht.

Mit Hilfe eines einzigen Primers konnten 9 stromabwärts liegende SNPs typisiert werden. Als DNS-Matrize zur Herstellung des asymmetrischen Produktes diente ein 426 bp PCR-Produkt. Die Primer-Elongationsreaktion mit diesen Einzelstrang-Produkten wurde auf Mikrorastern durchgeführt, die pro SNP zwei identische Oligonukleotid-Primer trugen, welche sich nur in der terminalen 3'-Base unterschieden. Alle 9 SNPs wurden typisiert, indem die Verhältnisse der Signale korrekter Basenpaarungen zu Basenfehlpaarungen gebildet wurden, die im Bereich von 1,67 bis 6,28 lagen und einen Mittelwert von 3,60 ergaben.

Der Abstand zwischen den Primern, die in der asymmetrischen PCR-Reaktion eingesetzt werden, und der Position der SNPs in der DNS-Matrize bestimmt die Länge der synthetisierten Sequenz und so auch den Einbau von Fluoreszenz-markiertem dUTP. Der optimale Abstand des PCR-Primers zum SNP wurde als 25 bp ermittelt. In einem nächsten Experiment wurden ausgehend von der gleichen 426 bp Matrize asymmetrische

PCR-Produkte als Target hergestellt, wobei die PCR-Primer so ausgewählt wurden, dass sie maximal 50bp vom SNP entfernt lagen. 6 einzelne Oligonukleotid-Primer wurden dafür in einer multiplex PCR-Reaktion eingesetzt. Der Abstand dieser Oligonukleotid-Primer zu den SNPs lag im Bereich von 35 bp bis 49 bp. Das Ergebnis ist in Abschnitt 4.9.2 finden. Die Signal-Verhältnisse von Basenpaarungen zu Basenfehlpaarungen waren im Vergleich zu dem vorherigen Experiment mit nur einem Oligonukleotid-Primer nicht wesentlich besser; sie ergaben Werte von 1,6 bis 3,6 mit einem Durchschnittswert von 2,76.

Um 48 SNPs in einer einzigen Primer-Elongationsreaktion zu detektieren, wurden 3 doppelsträngige DNS-Matrizen mit einer Größe von jeweils ca. 4 kb in einer asymmetrischen PCR verwendet. 3 asymmetrische multiplex PCR-Reaktionen wurden mit jeweils 13,15 und 20 Primern durchgeführt und anschließend für den Einsatz in der Primer-Elongationsreaktion vereinigt. Wie in Abschnitt 4.10 zu sehen, konnten alle bis auf zwei SNPs typisiert werden. Sie ergaben relative Signal-Intensitäten korrekte Basenpaarung zu Basenfehlpaarungen von 1,6 bis 30,7. Für 2 SNPs wurden Signal-Verhältnisse von 0,9 bzw. 1,2 gefunden. Eine Heteroplasmie als Ursache für diese falsch-negativen Ergebnisse könnte durch Sequenzierung der untersuchten mt-DNS ausgeschlossen werden.

5.6.3 Einsatz von Exonuklease-behandelten PTO-modifizierten PCR-Produkten als Target

Eine weitere Methode zur Herstellung von Einzelstrang-DNS-Targets bestand aus einer PCR-Amplifikation mit zwei Oligonukleotiden, von denen eines an den drei terminalen 5'-Basen PTO-modifiziert war. Der Vorteil dieses Ansatzes ist die Möglichkeit, multiplexe PCR-Reaktionen durch den Nachweis von distinkten Doppelstrang-DNS-Proben in einem Agarosegel zu kontrollieren. Die Produkte asymmetrischer PCR-Reaktionen (eine Mischung aus unterschiedlich langen Produkten, die aus

Kettenabbrüchen der linearen Elongation resultieren) liessen sich gelelektrophoretisch nicht als diskrete Banden darstellen, so dass die Beurteilung solcher Multiplex-Reaktionen schwierig ist. PTO-Gruppen stellen einen Schutz gegen Exonukleaseaktivität dar, und die nachfolgende Behandlung mit einer 5` -> 3` Exonuklease führt zu einem selektiven Verdau von unmodifizierten Strängen und den Verbleib eines Einzelstrangs. Das gleiche Prinzip wurde bei der „Genetic Bit Analysis“ Methode für die Typisierung von SNPs verwendet, um Einzelstrang-DNS-Produkte mit hoher Effektivität herzustellen (Nikiforov, et al., 1994). 9 SNPs, die zuvor mit den asymmetrischen PCR-Targets untersucht worden waren, wurden unter Verwendung von Einzelstrang-DNS-Targets, die mittels Exonuklease-Verdau des gleichen 426 bp PTO-modifizierten PCR-Produkts hergestellt waren, typisiert (Ergebnisse Abschnitt 4.12). Signal-Verhältnisse von korrekten Basenpaarungen zu Basenfehlpaarungen lagen im Bereich von 1,89 bis 8,21 mit einem Mittelwert von 4,25. Dieses Ergebnis zeigte eine Verbesserung der Diskriminierung im Vergleich zu asymmetrischen-PCR-Produkten.

5.7 Protokoll zur Typisierung möglicherweise aller mitochondrialer SNPs durch nur 5 Primer-Elongationsreaktionen

Als nächstes wurde untersucht, ob längere PTO-modifizierte PCR-Produkte, welche mehrere SNPs enthalten, verwendet werden können. Solche Targets würden die Anzahl der Primer zur Amplifikation SNP-enthaltender DNS-Regionen reduzieren. Alle auf Mikrorastern basierenden Methoden, soweit bislang publiziert, benötigen 2 Primer pro SNP-Locus zur Target-Herstellung. Wie bereits erwähnt, liegt darin ein Hauptproblem für die Untersuchung von SNPs in großem Maßstab. Ein erster Versuch zur Verwendung sehr langer Targets wurde mit einem 4 kb Fragment durchgeführt (Produkt B), welches zahlreiche mitochondriale SNPs enthielt. Pro SNP wurden zwei Oligonukleotide verwendet, um insgesamt 16 in diesem Produkt enthaltene SNPs

gemeinsam zu typisieren. In einer Primer-Elongationsreaktion konnten alle 16 SNPs typisiert werden, mit Signal-Verhältnissen von korrekten Basenpaarungen zu Basenfehlpaarungen im Bereich von 1,44 bis 5,90 und einem Mittelwert von 2,95 (Ergebnisse Abschnitt 4.12). Es stellte sich heraus, dass die Position einzelner SNPs innerhalb des 4 kb Produktes die Spezifität der Reaktion nicht beeinflusst. Die Anwendung von langen Einzelstrang-DNS-Targets wird auch durch wiederholte Denaturierungsschritte des entwickelten Protokolls ermöglicht (siehe Methodenteil, Abschnitt 3.11.3), welche zur Aufhebung von doppelsträngigen Strukturen führen. Dadurch wird die Zugänglichkeit von entsprechenden Abschnitten der DNS-Zielsequenz zu den auf dem Chip immobilisierten Oligonukleotiden verbessert. Da das mitochondriale Genom nur aus 16,569 kb besteht, wurden 4 zusätzliche Oligonukleotidpaare entworfen, um PCR-Produkte (A, C1, C2, D) herzustellen, welche zusammen mit Produkt B das mitochondriale Genom abdecken (Ergebnisteil Abschnitt 4.12). Allel-spezifische Oligonukleotidpaare, die 46 SNPs entsprechen, wurden auf einem Chip immobilisiert und diese 4 Exonuklease-verdauten PCR-Produkte gemeinsam in einer Primer-Elongationsreaktion als Target eingesetzt. Wie in Abschnitt 4.12 gezeigt, konnten 44 von 46 SNPs typisiert werden, mit einem Signal-Verhältnis von korrekten Basenpaarungen zu Basenfehlpaarungen von 1,4 bis 6,4, und einem Mittelwert von 2,99. Nur zwei SNPs ließen sich auf diese nicht typisieren; in einem Fall war kein Fluoreszenzsignal nachweisbar (SNP 29), und im anderen Fall (SNP 26) waren die Allel-spezifischen Signale nicht verschieden. Eine Primer-Elongationsreaktion mit einem längeren DNS-Target von 5,7 kb, die das Produkt C1 und C2 enthielt, versagte wiederholt. Dies deutet darauf hin, dass die Größe dieser Einzelstrang-DNS-Targets eine Länge zwischen 4,4 kb und 5,7 kb nicht überschreiten sollte.

5.8 Abschliessende Beurteilung der entwickelten Methode

Durchsatz

Die Zahl der in einem Experiment bestimmbaren SNPs ist v.a. begrenzt durch die Herstellung der SNP-enthaltenden Targets. Diese Limitierung betrifft alle Methoden zur SNP-Typisierung unabhängig von der Reaktionsplattform. Zwar wurde hier, wie auch in anderen Arbeiten, eine Multiplex-Strategie angewendet, keine dieser Protokolle stellt aber einen Durchbruch für die Erhöhung des Durchsatzes dar (Pastinen et al., 1996; Tully et al., 1996; Westin et al., 2000; Myakishev et al., 2001; Medintz et al., 2001; Fan et al., 2000; Tillib et al., 2001).

In dieser Arbeit konnte eine deutliche Vereinfachung der Target-Generierung für die Typisierung sehr wahrscheinlich aller in mitochondrialer DNS vorhandener SNPs entwickelt werden. Somit ist eine Detektion all dieser SNPs in einem Experiment möglich. Dies gilt auch für die Typisierung einer grossen Anzahl SNPs in umschriebenen Abschnitten genomischer DNS, welche bei dieser Technik in Fragmente von bis zu 4,4 kb aufgeteilt werden müssten.

Da kürzlich SNP-Haplotypen geringer Diversität im Genom gefunden wurden (siehe Diskussion, Abschnitt 5.1), ist es durchaus denkbar, dass solche Abschnitte durch wenige, repräsentative SNPs innerhalb eines PCR-Fragmentes dieser Größe charakterisiert werden können. Auch die Untersuchung von SNPs in der Nachbarschaft von Kandidatengenomen kann durch den Einsatz solch langer Targets anstelle von vielen einzelnen PCR-Fragmenten deutlich vereinfacht werden.

Automatisierbarkeit

Im Gegensatz zu anderen Methoden, insbesondere solcher, die auf der gelelektrophoretischen Trennung von DNS-Fragmenten beruhen, ist die Herstellung von

DNS-Chips sehr gut durch die Anwendung von Robotern („Arrayern“) automatisierbar (siehe Material, Abschnitt 2.8). Die Primer-Elongationsreaktion muss derzeit noch durch manuelle Pipettierschritte angesetzt werden, die anschliessenden Reaktionsschritte (Denaturierung, Annealing und Elongation) laufen jedoch unabhängig und elektronisch gesteuert in einem speziellen PCR-Gerät für Glasobjektträger ab (siehe Material, Abschnitt 2.8).

Die Beurteilung der abgelaufenen Elongations-Reaktionen erfolgt über das Messen von Fluoreszenzsignalen mittels konfokalem „Laser-Scanner“, und ist somit semiautomatisch durchführbar. Das Umrechnen von Bildpunktintensitäten der „spots“ wird von einem Computerprogramm übernommen, was wenige Eingaben des Benutzers erfordert.

Kosten

Zu den Herstellungskosten der DNS-Chips tragen im wesentlichen die erforderlichen Geräte wie „Arrayer“ und „Scanner“ bei. Ausserdem stellen die kommerziell erhältlichen Allel-spezifischen Oligonukleotide einen erheblichen Kostenfaktor dar. Allerdings sind diese Oligonukleotide auch für andere Methoden erforderlich (Abschnitt 1.3.2). Da einmal angesetzte Oligonukleotid-Lösungen für die Herstellung mehrerer hundert DNS-Chips verwendet werden können (nur sehr geringe Oligonukleotid-Mengen werden aufgebracht), werden die Herstellungskosten bei grossen Stückzahlen desselben Chips deutlich gesenkt. Laufende Kosten jeder Primer-Elongationsreaktion stellen v.a. das Fluoreszenz-markierte Oligonukleotid sowie auch die DNS-Polymerase dar.

Spezifität

Unter Spezifität der SNP-Typisierung versteht man die Güte der Diskriminierung zwischen verschiedenen Allelzuständen. Diese drückt sich bei der hier beschriebenen

Methode als Verhältnis der Signalintensitäten (d.h. Grad der Elongationsreaktion) zwischen korrekten Basenpaarungen und Basenfehlpaarungen aus.

In der vorliegenden Arbeit wurden Signalverhältnisse zwischen 1,44 und 50 gemessen. Andere Arbeiten, jedoch ebenfalls Chip-basierte Methoden zeigen Verhältnisse zwischen 7,9 und 18 auf (Pastinen et al., 2000). Die Verwendung von Enzymen zur Diskriminierung von Einzelbasenunterschieden scheint allgemein eine höhere Spezifität zu besitzen als die einfache Allel-spezifische Hybridisierung (Pastinen et al., 1997). Alle Allelzustände der hier untersuchten SNPs wurden unabhängig durch Sequenzierung untersucht. Bis auf einen SNP (Abschnitt 4.10) war die erreichte Diskriminierung ausreichend und die detektierten Allelzustände entsprachen dem Ergebnis der Sequenzierung.

Sensitivität

Die Sensitivität dieser Methode wird bestimmt durch das Verhältnis zwischen der Signalintensität von Chip-positionen mit Allel-spezifisch elongierten Oligonukleotiden und der Intensität benachbarter Chipoberflächen, letztere geht auf die Autofluoreszenz der Oberfläche sowie unspezifisch gebundenes Cy3-dUTP zurück. In der vorliegenden Arbeit wurden mit PCR-amplifizierten Targets bis auf einen von 47 Loci (siehe Ergebnisteil, Abschnitt 4.10) Intensitäten deutlich über dem lokalen Hintergrund gemessen. Andere Arbeiten berichten über Ausfallraten von bis zu 13%, insbesondere bei Anwendung einfacher Hybridisierungen zur SNP-Typisierung (Wang et al., 1998). Bei Verwendung komplexerer Zielsequenzen wie Inter-Alu PCR-Produkten konnten mit der hier entwickelten Methode jedoch keine ausreichend hohen Signale erhalten werden. Die Problematik der zu geringen Sensitivität beim Einsatz eines hochkomplexen Targetgemischs, in welchem einzelne Sequenzabschnitte in zu geringer Menge vorliegen, stellt die größte Herausforderung an zukünftige Entwicklungen zur SNP-Typisierung dar.

Bei der Detektion von Fluoreszenzsignalen könnte eine höhere Sensitivität durch eine Amplifikation des Signals erreicht werden. Z.B. könnte man bestimmte Polymermoleküle, sogenannte Dendrimere, verwenden, die hunderte von Farbstoffmolekülen tragen (Benters et al., 2001; Benters et al., 2002). Durch Einbau eines solchen Dendrimers anstelle nur eines einzigen Farbstoffmoleküls ist es möglich, ein deutlich stärkere Fluoreszenzsignale zu erhalten. Auch durch Weiterentwicklung der konfokalen Lasertechnologie sollte es möglich sein, die Sensitivität der hier vorgestellten Methode deutlich zu erhöhen.

Schließlich gibt es alternative Detektionsmethoden mit höherer Sensitivität, wie die Massenspektrometrie (Ross et al., 1997; Braun et al., 1997; Sauer et al., 2000; Buetow et al., 2001), Gold- und Silbernanopartikel anstelle von Fluorophoren (Taton et al., 2000), Fluoreszenzanregung durch ein "zwei-Photonen-System" (Hanninen et al., 2000), die "Fiberoptic-DNA-Sensor-Technik", welche auf der Bindung von Fluoreszenzmarkierter DNS an faseroptischen Lichtbündeln basiert (Healey, et al., 1997; Epstein et al., 2002), die Quantum-Dot-Fluoreszenzpartikel Technik (Han, et al., 2001) und das "Carbon nanotube probe"-System, bei dem markierte Oligonukleotide auf spezifische komplementäre Zielsequenzen hybridisiert und durch den Einsatz eines "atomic force microscop (AFM)" mit Hilfe von einwandigen Kohlenstoff-Nanotypes typisiert werden (Woolley et al., 2000).

Von diesen alternativen Detektionsmethoden wurden bislang nur die ersten drei mit der Chiptechnologie kombiniert. Gelänge damit eine Messung von SNPs mit genomischer DNS als Target, käme damit erstmalig eine genomweite SNP-Typisierung in Sicht.

Eine hocheffiziente Methode, die allein gesamt-genomische DNS als Target erfordert, und die Genotypisierung von SNPs mit hohem Durchsatz erlaubt, wird einen bedeutenden Einfluß auf unser Verständnis von Polymorphismen und deren Einfluß auf biologische Funktionen haben.

Mit der hier vorgestellten Methode ist ein entscheidender und zukunftsweisender Beitrag dazu geleistet worden.