IV. Diskussion

4.1 Genetische Kartierung in der Maus

Genetische Karten werden erstellt, um Eigenschaften wie z.B. sichtbare, morphologische Mutationen oder physiologische Merkmale auf chromosomale Regionen der 20 Mauschromosomen zu lokalisieren und ein oder mehrere Gene zu finden, die für das Merkmal verantwortlich sind. Dieser Ansatz wird 'reverse Genetik' oder ,positionelle Klonierung' des Gens genannt, denn er begibt sich, vom ,äusseren' Merkmal ausgehend, auf die Suche nach dem zugrunde liegenden Gen. Traditionell geht positionelle Klonierung von einem Krankheitsphänotypen aus zum Genotyp. Zuerst muss eine Kandidatenregion ermittelt werden, dann kann mit der physikalischen Kartierung der Region (Kontigbuilding) begonnen werden und letztendlich soll das der Krankheit zugrunde liegende Gen kloniert werden

Genetische Karten werden aus Vererbungsdaten von DNA-Markern erstellt. Als DNA-Marker dienen definierte DNA-Sequenzen, deren Vererbung in einer Mausfamilie mit verschiedenen Methoden leicht verfolgt werden kann. DNA-Sequenzen, die sich zwischen zwei ingezüchteten Mäusestämmen unterscheiden, nennt man polymorph und nur polymorphe Marker können für die Erstellung einer genetischen Karte verwendet werden. Ein ingezüchteter Mäusestamm ist für das vom Vater und das von der Mutter stammende Gen eines Genorts auf den Chromosomenpaaren 1-20 homozygot, d.h. an diesem Genort (Locus) befinden sich identische DNA-Sequenzen (Allele). Bei der Kreuzung von zwei Inzuchtstämmen entstehen Nachkommen, die F1-Generation, die für jeden Genort heterozygot sind, d.h. je ein Allel jedes Elterntieres enthält. Die Kreuzung von F1-Tieren mit einem Elternstamm wird als Rückkreuzung bezeichnet. Rückkreuzungstiere (B1-Tiere) enthalten einen Chromosomensatz mit Allelen des Parentalstamms und einen Chromosomensatz des F1-Tiers, der durch Rekombination während der Meiose Allele beider Parentalstämme trägt. Das B1-Tier ist für seine Genorte also entweder homo- oder heterozygot. Je weiter zwei Genorte räumlich voneinander entfernt auf dem Chromosom liegen desto häufiger liegt der eine homozygot und der andere heterozygot vor, d.h. während der Meiose des F1-Tieres hat Rekombination zwischen den Chromosomenpaaren stattgefunden und die parentalen Allele sind ausgetauscht worden. Diese Rekombinationen auf jedem Chromosom mittels DNA-Markern zu finden und die Häufigkeit von Rekombination zwischen DNA-Markern zu zählen heisst eine genetische Karte aus DNA-Loci zu erstellen.

Die Auflösung einer genetischen Karte hängt von der Zahl der Rückkreuzungstiere ab, die für die Analyse zur Verfügung stehen: Genetische Karten werden nach dem Drosophilagenetiker Morgan (Morgan und Cattell 1912) in der Einheit 'centiMorgan' erstellt. Findet man zwischen zwei Markerloci in 100 Rückkreuzungstieren eine Rekombination der parentalen Allele, so haben die Loci per definitionem den genetischen Abstand von 1 centiMorgan (cM). Welcher physikalischen Länge in Megabasen DNA

entspricht das? Die physikalische Länge des Mausgenoms beträgt ca. 3000 Megabasen. Die genetische Länge des Mausgenoms wird mit ca. 1400cM angegeben (Copeland et al. 1993; Dietrich et al. 1996; Rhodes et al. 1998). Zwei Markerloci, die einen Abstand von 1cM auf einer genetischen Karte haben, liegen daher durchschnittlich ca. 2 Megabasen (2000 Kilobasen) voneinander entfernt. Eine Erhöhung der Auflösung einer genetischen Karte erreicht man mit der Analyse einer grossen Zahl von Meiosen, z. B. von 1000 Rückkreuzungstieren.

Die Analyse polymorpher Markerallele in einer Mauskreuzung kann mit Hilfe verschiedener Markersysteme durchgeführt werden und wird als Genotypisierung bezeichnet.

4.2 Markersysteme

Das RFLP-Markersystem (Restriktionsfragmentlängenpolymorphismus; Lander und Botstein 1986) beruht darauf, dass sequenzspezifische Restriktionsendonukleasen genomische DNA zweier Inzuchtmäusestämme aufgrund variierender DNA-Sequenzen eines Locus in verschieden lange DNA-Fragmente spalten. Die Spaltprodukte jedes Stammes werden auf einem Agarosegel getrennt, auf eine DNA-bindende Membran übertragen (Southern Blot) und mit DNA-Fragmenten hybridisiert, die längenpolymorphe Fragmente der beiden Stämme nachweisen. Der Nachteil der Methode liegt darin, dass für einen Southern-Blot 5 µg DNA benötigt werden und für die Analyse jedes Restriktionsenzyms ein eigener Blot angefertigt werden muss.

SSLP-Marker (simple sequence length polymorphism; Love et al. 1990) - auch unter 'Mikrosatellitenmarker' bekannt - sind hochpolymorphe Loci im Genom der Maus, die sich in der Anzahl ihrer Dinukleotidwiederholungen unterscheiden (z.B. Allel1: $(CA)_n$ und Allel2: $(CA)_{n+1}$). Sie werden mit Primern für benachbarte Einzelkopiesequenzen von genomischer DNA amplifiziert. Unterscheiden sich die Amplifikate zweier Mäusestämme in ihrer Länge, so kann die Vererbung dieses Locus in einer Mäusefamilie verfolgt werden, indem man die PCR-Produktlängen auf einem Agarose- oder Polacrylamidgel analysiert. Für die PCR-Amplifikation eines Mikrosatellitenlocus werden nur 20ng genomischer DNA einer Maus benötigt. Die in dieser Arbeit verwendeten Mikrosatellitenmarker fallen in diese Kategorie von Markern.

Eine weitere auf PCR beruhende Methode, DNA-Polymorphismen nachzuweisen, ist die IRS-PCR (interspersed repetitive sequence PCR; Cox et al. 1991). Während jeder Mikrosatellitenlocus durch ein sequenzspezifisches Primerpaar definiert ist, das die polymorphe Dinukleotidsequenz amplifiziert, verhält es sich mit IRS-Markern genau umgekehrt. Mit einem einzigen Primer, der sequenzspezifisch an ein repetitives DNA-Element bindet, in unserem Fall das ca. 300bp grosse repetitive Element SINE- Element (short interspersed nuclear element), B1' der Maus, werden von genomischer DNA Einzelkopiesequenzen

90

amplifiziert. Das B1-Element ist mit 130000 – 180000 Kopien das häufigste repetitive Element im Genom der Maus (Krayev et al. 1980). Die Komplexität der eingesetzten DNA, wie z.B. Gesamtpräparation genomischer DNA, YAC- oder BAC-Klon DNA, bestimmt die Anzahl der amplifizierten IRS-PCR-Produkte. Liegen die repetitiven B1-Elemente in einem Abstand von weniger als 2000 Nukleotiden voneinander entfernt und in der passenden Orientierung zueinander, so wird ein PCR-Produkt gebildet. IRS-PCR von YAC- oder BAC-Klonen ergibt eine geringe Anzahl von PCR Produkten (ca. 5 von YACs und 0 – 3 von BACs).

Die Grundlage für Polymorphismen bildet die Verteilung der repetitiven Elemente im Genom verschiedener Mäusestämme. Der Polymorphismus kann entweder in der Länge der amplifizierten Einzelkopiesequenz liegen, wenn eines der benachbarten repetitiven Elemente verschoben ist oder in Vorhandensein oder Fehlen eines PCR-Produktes, wenn ein repetitives Element ganz fehlt. Quantitative Unteschiede in der Menge eines PCR-Produktes lassen sich auf eine Mutation im B1-Element zurückführen, die zum Beispiel die Spezifität der Primerbindung herabsetzt.

Für die Genotypisierung einer Mausrückkreuzung wird genomische DNA der Parental-, der Hybrid- sowie der Rückkreuzungstiere mit dem B1-Primer amplifiziert. Die komplexen Produktmischungen beliebig vieler Tiere werden auf eine DNA-bindende Membran aufgebracht (,(Hybridisierungs)filter') und nacheinander mit einzelnen klonierten, radioaktiv markierten IRS-PCR Produkten – den IRS-PCR Markern - hybridisiert. Hybridisierungsexperimente mit IRS-PCR Markern liefern eine positiv/negativ Antwort über die Existenz eines IRS-Produktes im Genom der Maus.

Der Vorteil der IRS-PCR Marker liegt darin, dass eine einzige PCR-Reaktion an 50ng genomischer DNA einer Maus ausreicht, um nachfolgend in Hybridisierungen mit radioaktiven IRS-PCR-Markern Genotypisierungsdaten für 1000 oder mehr Loci zu erhalten. Ausserdem liefert eine einzige Hybridisierung eines DNA-Filters Genotypisierungsdaten für hunderte von Rückkreuzungstieren. Für die Herstellung der IRS-PCR Marker ist keinerlei Sequenzinformation erforderlich. Roger Cox (Cox et al. 1991) zeigte, dass IRS-PCR-Produkte von genomischer DNA der Mäusespezies *Mus musculus* und *Mus spretus* eine Polymorphismusrate von >30% aufweisen und daher für die Segregationsanalyse der EUCIB Rückkreuzung als Quelle neuer Marker nützlich sind.

Die Auflösung einer genetischen Karte hängt von der Zahl der Rückkreuzungstiere ab, die für die Analyse zur Verfügung stehen. Die 1000 Rückkreuzungstiere des EUCIB-Projekts erlauben eine durchschnittliche Auflösung von 1/1000 oder 0,1 cM. Das entspricht einer physikalischen Länge von 200 Kilobasen und macht das Zusammensetzen von YAC-Klonen zu zusammenhängenden Sequenzen – das ,YAC-Contigbuilding' – möglich. Die hohe Auflösung der EUCIB-Kreuzung dient der sehr genauen Kartierung von Genen, während für die hier durchgeführte Proteinkartierung 64 Tiere eingesetzt wurden, was eine durchschnittliche Auflösung von 1,56 cM (das sind ca. 3 Megabasen) erlaubt.

4.3 Die Beobachtung: Proteinspots eines 2DE-Musters zweier Mäusestämmen zeigen Variationen

Wie echt sind die Variationen, die ein Proteinspot auf 2D-Gelen zeigt? Der 2D-Elektrophorese wird oft nachgesagt, dass das, was sie zum Vorschein bringe – Daten über die quantitative und qualitative Veränderung von separierten Proteinspots verschiedener Gewebe oder Organismen – nicht reproduzierbar und eventuell sogar Artefakte der Probenaufarbeitung ohne biologische Bedeutung seien. Die Position eines Proteinspots sei nicht konstant, weil die 2D-Muster Verzerrungen zeigen und die Menge eines Proteinspots variiere mit der Färbungsintensität der Silberfärbung und der aufgetragenen Probenmenge, das menschliche Auge sei nur in der Lage 70 Graustufen zu unterscheiden und daher für die Analyse von Quantitäten nicht ausreichend, ganz zu schweigen von der (Un)möglichkeit computergestützter glaubhafter Bildanalyse.

Die dringende Frage, die es zu beantworten gilt: sind die 1324 polymorphen Proteinspots des Mausgehirnextrakts bedeutungstragende Variationen? Sind sie reproduzierbar? Spiegeln sie zufälliges Geschehen in der Physiologie des Mausgehirns wider oder stellen sie echte genetisch bedingte Entitäten dar?

Generell kann man zur technischen und biologischen Reproduzierbarkeit der Unterschiede sagen: für die genetische Analyse wurden nur Spotunterschiede zwischen den Ausgangsstämmen verwendet, die in drei Elternpaaren unabhängig gefunden wurden und in drei weiteren Paaren bestätigt wurden. Positionsverschiebungen sind meist einfach zu erkennen, quantitative Unterschiede wurden nur registriert, wenn sie ganz augenfällig waren (Klose, persönliche Mitteilung).

4.3.1 Sind die Variationen der Proteinspots der Mäusestämme *Mus spretus* und *Mus musculus* genetisch bedingt?

Um dieser Frage nachzugehen, wurden die 1324 polymorphen Spots zwischen den Mäusespezies in deren direkten Nachkommen, der F1-Generation, sowie in der Rückkreuzungsgeneration (B1) untersucht. In Abhängigkeit vom Variationstyp konnte für Variationen, die einem Mendelschen Erbgang folgten und entweder dominant in B6 oder kodominant (intermediär) vererbt wurden, die genetische Bedingtheit nachgewiesen werden und die Variation auf der genetischen Karte der Maus lokalisiert werden. Polymorphe Spots, die dem gleichen Variationstyp angehörten und identische Segregation zeigten, wurden unter demselben Variantennamen zusammengefasst. Die 1324 polymorphen Proteinspots wurden daher von Klose zu 936 Varianten zusammengefasst.

4.3.2 Welche Variationstypen wurden gefunden und welchen Erbgängen folgen sie?

Für qualitative Variationen, d.h. Verschiebung eines Spots relativ zu seinen Nachbarspots (mobility variation, mV) lässt sich der Erbgang einfach und sicher bestimmen: kodominant vererbt, erscheinen beide parentalen Spots auf dem 2D-Gel des F1-Tieres, während im Rückkreuzungstier entweder nur der Spot des Rückkreuzungelterntieres auftritt (homozygote Tiere) oder beide Spots (heterozygote Tiere) wie im F1-Tier. Für 327 (89%) Mobilitätsvarianten wurden kodominante Erbgänge gefunden (Klose, persönliche Mitteilung), für 39 (11%) liess sich kein mendelscher Erbgang attestieren. Dominant/rezessive Mobilitätsvarianten traten nicht auf.

Für quantitative Variationen eines Proteinspots ist es wesentlich schwieriger einen Mendelschen Erbgang zu attestieren: ein dicker Spot in der einen Mäusespezies und ein dünner Spot in der 2. Spezies resultieren - kodominant vererbt - in einem Spot mittleren Volumens in der F1-Generation und in Rückkreuzungstieren muss zwischen der mittleren Intensität heterozygoter Tiere und der Intensität des parentalen Rückkreuzungstieres (homozygot) unterschieden werden. 40% der quantitativen Varianen vom Typ paV und aV wurden kodominant vererbt (230 Varianten), nur 30% zeigten einen dominant/rezessiven Erbgang (169 Varianten) und für weitere 30% konnte kein Erbgang bestimmt werden (171 Varianten) (Klose, persönliche Mitteilung).

		1324 polymorphe Einzelspots			
		936 Varianten			
668 Einzelspots	366 (100%) qualitative Varianten (mV)	 327 (89%) kodominant 0 dominant/rezessiv 39 (11%) nicht mendelnd 	274 (75%) Varianten kartiert	409 Varianten der genetis	664 polymorp
656 Einzelspots	570 (100%) quantitative Varianten (paV, aV)	230 (40%) kodominant 169 (30%) dominant/rezessiv 171 (30%) nicht mendelnd	135 (24%) Varianten kartiert	kartierten auf chen Karte	he Einzelspots

Tab.4.3.2 Variationstyp und Vererbungmodus der 1324 Polymorphismen (936 Varianten) von *Mus musculus* und *Mus spretus*

Die 1324 polymorphen Spots wurden zu 936 Varianten zusammengefasst. Für 441 Varianten (744 Einzelspots) konnte die genetische Bedingtheit der Variation basierend auf der Segregation in der F1- und Rückkreuzungsgeneration nachgewiesen werden. Von diesen 441 konnten 409 (664 Einzelspots) auf der

genetischen Karte der Maus lokalisiert werden. Der Nachweis eines Mendelschen Erbganges für eine Mobilitätsvariante liefert im selben Zuge die Bestätigung dafür, dass die beiden speziesspezifischen Spots tatsächlich ein polymorphes Paar bilden. Für die Varianten mV146, mV147, mV148, mV269, mV363, aV232 wurde dies auch durch Massenspektrometrische Identifizierung des Spots von *Mus musculus* sowie des polymorphen Partners von *Mus spretus* bestätigt (Kap.3.4, Tab.3.4.a).

Ca. 600 Spots der 1324 polymorphen Spots zeigten keinen Mendelschen Erbgang und konnten deshalb mit der hier verwendeten Methodik nicht kartiert werden. Der grösste Teil dieser Varianten gehören dem quantitativen Typus an. Auch von den 441 Varianten mit identifiziertem Mendelschem Erbgang liessen sich 32 nicht auf der genetischen Karte der Maus einordnen. Das sind 20 qualitative Varianten (Mobilitäts-Varianten: mV) und 12 guantitative Varianten (Presence/Absence und Amount Varianten: paV, aV). Vier mV-Varianten zeigten ausschliesslich homozygote Vererbung und können daher nicht kartiert werden, 12 Varianten hatten bei der Auswertung Schwierigkeiten bereitet (Klose, persönliche Mitteilung) und wiesen daher wahrscheinlich unrichtig interpretierte Daten auf. Fünf quantitative Varianten allerdings zeigten ein gut auswertbares Segregationsmuster und konnten trotzdem nicht kartiert werden: dies lässt sich damit erklären, dass der Phänotyp der betreffenden Proteinspots nicht auf einem einzelnen, nämlich dieses Protein kodierenden Genortes und seinen Allelen beruhte, sondern einen Phänotyp darstellt, der durch mehrere Gene verursacht wird. Die genaue Menge eines polymorphen Spots in den einzelnen Rückkreuzungstieren kann mit dem Auge nicht erfasst und klassifiziert werden und auch computergestützte Analysen könnten nur mit ausgewählten, geeigneten Spots zur Identifizierung mehrerer Intensitätsklassen führen (der Spot gross genug sein, um mit Coomassie anfärbbar zu sein). Ein solcher Phänotyp, der in der Rückkreuzung in mehreren quantitativen Klassen aufritt, wird als ,quantitative trait" Locus (QTL) bezeichnet. Für weitere nicht-kartierte Varianten standen nur 27 – 33 Genotypisierungsdaten zur Verfügung, was nicht ausreicht, um für bestimmte Regionen des Genoms statistisch signifikante Kopplung zu finden (siehe Kap.3.3.2: Abb.3.3.2). Dies sind Regionen auf den Chromosomen 1, 2, 3, 4, 6, 7, 8, 10, 12, 15, 16, 19 und X.

Die Vererbung von polymorphen Spots wurde bisher an Drosophila (Spicer 1988; Zeng und Singh 1993), Mäusen (vier polymorphe Mausgehirnproteine von 2DE-Mustern der Rekombinanten Inzuchtlinie BxD; Goldman und Pikus 1986), Menschen (sechs von 19 polymorphen Lymphozytenproteinen; Goldman et al. 1991) und verschiedenen Pflanzenarten wie Pinie (65 Polymorphismen kartiert; Gerber et al. 1993), Mais (Damerval et al. 1994), Erbse und Gerste (de Vienne et al. 1996) untersucht. Alle 170 Positionspolymorphismen, die von De Vienne in Pinie, Mais und Erbse untersucht wurden, zeigten eine Positionsverschiebung auf dem 2D-Gel und - wie er es nennt - monogenetische Vererbung. Damerval (Damerval et al. 1994) klassifiziert Proteinpolymorphismen zweier Maislinien in qualitative Positionsvarianten, die er als genetische Marker für die Erstellung einer genetischen Karte verwendet – diese entsprechen unseren mV-Varianten - und in quantitative Varianten, die er von vorne herein als multifaktoriellen Phänotyp betrachtet. Er quantifiziert die varianten Spots der B1-Generation und findet, dass sie verschiedenen quantitativen Klassen angehören. Als Ergebnis wurden für 42 von 72 polymorphen Proteinspots jeweils ein bis fünf unabhängige Loci gefunden, die an der Ausbildung des quantitativen Phänotyp des Proteinpolymorphismus beteiligt sind. Diese Untersuchung zeigt eindrucksvoll, dass zusätzlich zur Kartierung des Strukturgenlocus auch mehrere weitere über das Genom verteilte Genorte gefunden werden konnten, die an der Regulation der Menge des Proteins beteiligt sind.

4.3.3 Die molekulare Basis von genetischer Variation

Die einfachste Erklärung für einen Proteinpolymorphismus liegt in der Annahme einer Mutation des proteinkodierenden Gens, die in einem Aminosäureaustausch - einer Veränderung der Primärstruktur des Proteins - resultiert (siehe auch (Damerval 1994). Der Austausch einer ungeladenen gegen eine ladungstragende Aminosäure verändert die Nettoladung des Proteins und führt damit zu einer Verschiebung des Isoelektrischen Punktes und der Position des Proteins im IEF-Gel. Neben der veränderten Beweglichkeit kann die Teilnahme des Proteins an der Bildung von multimeren Proteinkomplexen verändert sein, die neue Aminosäure kann die Möglichkeit einer zusätzlichen posttranslationalen Modifizierung durch z.B. Glykosylierung bieten und damit das Molekulargewicht des Proteins verändern. Auch die aktuell vorhandene Menge des Proteins in der Zelle kann durch einen Aminosäureaustausch beeinflusst werden und zu veränderter Stabilität des Proteins führen. Dies wurde für Mutationen in der Phenylalaninhydroxylase nachgewiesen: 60% der 400 untersuchten Allele zeigten eine beschleunigte Degradation im Vergleich zum Wildtypprotein (Scriver und Waters 1999). Allein die veränderte Primärstruktur eines Proteins liefert somit viele Erklärungsmöglichkeiten für die Beobachtung eines qualitativen oder quantitativen Polymorphismus auf einem 2D-Gel. Tatsächlich hat die Identifizierung der polymorphen Spotpaare von Mus spretus und Mus musculus (Tab.3.4.a) für die Varianten mv269, mv148, mv363, mv146, mv147 und av232 ergeben, dass es sich um Allele desselben Genortes handelt. Auch Touzet hat durch Aminosäuresequenzierung festgestellt, dass polymorphe Proteine in fast allen Fällen Allele eines Proteins sind (Touzet et al. 1995).

Für das Auftreten einer Variation könnte jedoch auch ein modifizierendes/regulierendes Protein verantwortlich gemacht werden, das selbst polymorph ist: in Abhängigkeit von der Expression des Modifikators selbst wird ein anderes Protein modifiziert (Mausspezies1) oder nicht modifiziert (Mausspezies2). Im Falle eines intermediären Erbganges wie ihn Positionsvarianten zeigen, würden beide Spots im F1-Hybriden (Mausspezies1 x Mausspezies2) koexistieren, wenn man annimmt, dass der Modifikator in seiner Aktivität limitierend ist. In den Rückkreuzungstieren sind heterozygote Tiere mit zwei

95

Spots sowie homozygote Tiere mit einem Spot zu erwarten. Der Phänotyp einer quantitativen Variante (paV oder aV) lässt sich genauso entweder als ein intermediärer Erbgang beschreiben, den gleichen Phänotyp würde man aber auch erwarten, wenn man einen unabhängigen Genlocus – einen Modifieroder einen Regulatorgenort - postuliert, der die Proteinmenge wie oben beschrieben beeinflusst (s. Abb.4.3.3, links).





4.3.4 Häufigkeit von Proteinpolymorphismen zwischen zwei Mäusespezies

Unter etwa 8700 analysierten cytosolischen Proteinen von *Mus musculus* und *Mus spretus* kamen etwa 1324 (ca. 15%). zum Vorschein, die einen Polymorphismus zeigten. Davon liessen sich 664 Spots genetisch kartieren, das sind etwa 50% (oder 7,6% aller Spots). Die anderen 50% varianter Spots folgen entweder einem komplexeren Erbgang und sind von mehreren Faktoren kontrolliert oder sie variieren mit dem physiologischen Zustand der Tiere. Proteinextrakte weiterer Organe einer Kreuzung von *Mus musculus* und *Mus spretus* (Herz, Leber, Niere, Muskel, ...) bergen das Potential, weitere organspezifisch exprimierte Proteinpolymorphismen ausfindig und damit kartierbar zu machen, zumal die Polymorphismusrate für organspezifische Proteine höher zu sein scheint als für nicht-organspezifische (Klose 1982; de Vienne et al. 1988). Hier schliesst sich die Frage an, wieviele Proteinpolymorphismen zwischen den beiden Mäusespezies aufgrund der derzeitigen Kenntnisse über die Häufigkeit von DNA-Polymorphismen zu erwarten sind. Vergleicht man das Auftreten von Nukleotidaustauschen in

kodierenden Sequenzen von Genen von Mus musculus und Mus spretus, so findet man Einzelnukleotidaustausche (single nucleotide polymorphisms, SNPs) mit einer Häufigkeit von 0.4% (35 SNPs in 8909bp von 9 Genen) (Himmelbauer, persönliche Mitteilung). Aufgrund des degenerierten genetischen Codes werden hiervon nur ca. zwei Drittel in Aminosäuresubstitutionen resultieren, also 0.27%. Analysen der Häufigkeit von SNPs in menschlichen Genen ergab, dass nicht zwei Drittel sondern nur etwa 50% der SNPs in einem Aminosäureaustausch resultieren. Nimmt man an, dass diese Selektion gegen SNP-Mutationen in Nagern in gleichem Ausmass zum Tragen kommt, so kann man damit rechnen, dass nur 0,2% der SNPs in Aminosäuresubstitutionen resultieren. Das entspricht einem SNP in 500bp kodierender Sequenz. Welche Rückschlüsse lassen sich damit auf die Häufigkeit von SNP-bedingten Aminosäureaustauschen der Proteine des 2D-Gels unseres Mausgehirnextraktes ziehen? Eine DNA-Sequenz von 500bp kodiert für ein 166 Aminosäuren langes Protein von ca. 20kD (gerechnet mit dem durchschnittlichen Gewicht einer Aminosäure von 120D). Die meisten Proteine des 2D-Gels, nämlich 7420 Spots oder 85%, haben eine Masse grösser als 20kD und sollten deshalb mindestens einen Aminosäureaustausch zeigen (Spots von 200kDa sogar 10). Die 1347 Proteinspots unter 20kD bis zu den kleinsten von 9kD sollten eine bis 0.5 Mutationen (durchschnittlich 0.75) tragen (1010 Spots). Insgesamt müsste man erwarten, dass 96% aller Spots wenigstens einen Aminosäureaustausch aufweisen. Andererseits berechnete Halushka aus der SNP-Analyse von 75 menschlichen Genen, dass 17% der Proteine in menschlichen Populationen nicht polymorph sein sollten (Halushka et al. 1999). Cargill ermittelt aus der Analyse von 114 Allelen von 106 Genen, dass ein SNP in durchschnittlich 734 kodierenden Basen vorkommt, obwohl in 26 Genen länger als 734 bp kein SNP entdeckt wurde, was 25% nichtpolymorphen Proteinen entspricht (Cargill et al. 1999). Für Mäuse kann derzeit noch keine Aussage gemacht werden, da nicht genügend Gensequenzen verschiedener Stämme in den Datenbanken existieren.

4.3.5 Detektierbarkeit von Mutationen mittels 2D-Elektrophorese

Wie kann man die Diskrepanz zwischen der Anzahl qualitativer Varianten des 2D-Musters (15%) und der theoretischen, auf SNP-Analyse basierenden Anzahl von 96% erklären? Zum einen könnte die 2D-Elektrophorese nicht empfindlich genug sein, um alle Mutationen in Proteinen sichtbar zu machen. Zum anderen haben die meisten Aminosäureaustausche keinen Effekt auf die Ladung, Masse, Konformation oder posttranslationale Modifizierbarkeit (Seitenketten) und verursachen daher keinen 2D-Gel-Phänotypen. Zur Verteidigung der Sensitivität der Elektrophorese lässt sich folgendes anführen: die Isoelektrische Fokussierung ist in der Lage, selbst den Austausch einer neutralen gegen eine andere neutrale (ungeladene) Aminosäure aufzudecken (Whitney et al. 1985). In der Hämoglobin alpha-Kette konnten sechs von sieben neutralen Aminosäuresubstitutionen mittels IEF nachgewiesen werden (Cossu

et al. 1986). Der Austausch von neutralen, ungeladenen Aminosäuren bedeutet eine geringe Veränderung der Hydrophobizität des Proteins, die sich auf den pK-Wert und die Ionisierbarkeit der Seitenketten von umliegenden Aminosäuren auswirkt und zu einer geringen Nettoladungsänderung führte. Das in der vorliegenden Untersuchung verwendete System zur Isoelektrischen Fokussierung erreicht eine sehr hohe Auflösung des pH-Gradienten (pH3 – pH11) über eine Trennstrecke von 40cm und ist daher besonders sensitiv für Ladungsunterschiede zum Beispiel konnten die einzelnen Phosphorylierungsstufen der hochmolekularen Isoform 1 des Tau-Proteins (758 AS, ca. 91KD, (Andreadis et al. 1992)) in 96 Isospots auf dem 2D-Gel sichtbar gemacht werden (Janke et al. 1996). Die SDS-Gelelektrophorese trennt Proteine aufgrund ihrer Masse. SDS bindet an Proteine in einem konstanten Verhältnis von 1.4g SDS pro Gramm Protein und führt daher für alle Proteine zu einem relativ konstanten Masse-Ladungsverhältnis (Gerstner et al. 2000). Obwohl man dachte, dass SDS-Gelelektrophorese nicht empfindlich genug ist, um Massenänderung von Proteinen aufzuzeigen, die auf dem Austausch einer einzelnen Aminosäure basieren, haben Studien das Gegenteil bewiesen (Noel et al. 1979). Diese Effekte wurden auf geringe Konformationsänderungen zurückgeführt, die durch den Austausch einer Aminosäure gegen eine andere bleiben. hervorgerufen wurden und selbst unter denaturierenden Bedingungen erhalten Konformationsänderungen können zur Bindung leicht abweichender Mengen SDS und damit zu leicht veränderter elektrophoretischer Beweglichkeit des Proteins führen. Selbst wenn man annimmt, dass mit der hier verwendeten hochauflösenden 2D-Elektrophorese nur die Hälfte aller Aminosäuresubstitutionen entdeckt werden, bliebe ein grosser Anteil von über 4000 Proteinspots übrig, von denen nur etwa1300 Spots einen Polymorphismus auf dem Gel zeigen (siehe Tab.4.3.2). Für einen grossen Anteil der Spots muss man annehmen, dass die Aminosäuresubstitution zu keinen weiteren strukturellen Veränderungen des Proteins geführt haben. Für diese Polymorphismen ist anzunehmen, dass sie auch funktionell neutral sind.

Quantitative Variabilität der Proteine

Die Kartierung von guantitativen Varianten erbrachte, dass von den guantitativen Varianten (paV, aV) nur 24% kartiert werden konnten, während von den Mobilitäts-Varianten 75% kartiert werden konnten (siehe Tab.4.3.2). Diese Tatsache spiegelt die grosse Variabilität der quantitativen Varianten wider. Die quantitative Variation eines Proteinspots konnte nur dann zur genetischen Kartierung des Spots herangezogen werden, wenn sie auffällig genug war, um auch den Mengenunterschied im heterozygoten homozygoten (Proteinmenge (mittlere Proteinmenge) und des Mus spretus Elterntieres) Rückkreuzungstier zu erkennen. Es ist deshalb anzunehmen, dass neben den nicht klassifizierbaren und daher nicht kartierbaren guantitativen Varianten von Mus musculus und Mus spretus eine grosse Anzahl, wenn nicht alle Proteinspots quantitative Variationen zeigen, dass diese Variation jedoch unter der Detektionsgrenze liegt: ein Grund dafür liegt in der limitierten Detektierbarkeit von Proteinspots des 2D-Gels und deren Mengenänderungen durch die Silberfärbung. Die Silberfärbung hat den Vorteil, dass sie die sensitivste aller Färbungen ist, und daher auch kleine Proteinspots anfärbt (ca. 1-10ng Protein, 20-200 femtomol eines 50kD Proteins). Lineare Detektion von Färbungsintensität und Proteinmenge existiert jedoch nur in einem kleinen Bereich von ca. 1 - 50ng, was eine verlässliche Quantifizierung nur über einen kleinen Mengenbereich möglich macht (Damerval 1994; Yan et al. 1999).

Für diejenigen der qualitativen Varianten, die genetisch kartierbar waren, kann man annehmen, dass sie auf SNP-Mutationen im Strukturgenort oder - wie für die abweichend kartierten Varianten paV410, paV419 und mv82 auf einen varianten Modifier-Genort – beruhen. Man weiss, dass Aminosäuremutationen zu veränderter Stabilität eines Proteins führen können, z. B. dadurch, dass die Degradation beschleunigt oder vermindert wird.

Welche Hinweise ergeben sich aber aus der grossen Anzahl nichtkartierbarer, quantitativer Varianten und der grossen Anzahl erwarteter aber nicht auf dem 2D-Gel nachweisbarer SNP-Mutationen?

Vergleicht man die Häufigkeit des Vorkommens von Nukleotidaustauschen in kodierenden und nichtkodierenden Sequenzen von Genen von *Mus musculus* und *Mus spretus*, so findet man SNPs mit einer Häufigkeit von 0.4% (35 SNPs in 8909bp von 9 Genen) in kodierenden Sequenzen und 1,4-2,2% in nichtkodierenden Sequenzen (Himmelbauer, persönliche Mitteilung). Nicht-kodierende Sequenzen spielen jedoch für die Regulation der Genexpression eine grosse Rolle. Man kann also annehmen, dass praktisch jedes kodierende Gen noch eine Anzahl regulatorischer SNP-Polymorphismen besitzt, die unabhängig vom Strukturgen in der Reuckkreuzung rekombinieren und für die quantitativen Variationen verantwortlich sind.

4.3.6 Praktische Anwendungen: Sind Proteinpolymorphismen als genetische Marker einsetzbar?

Enzympolymorphismen wurden schon vor dem Anbruch der DNA-Ära für genetische Analysen und die Kartierung von Merkmalen in Drosophila (Kojima et al. 1970) und in Mäusen (Kaplan et al. 1973; Nichols und Ruddle 1973; Randerson 1973) verwendet. Nachdem DNA-Polymorphismen mittels PCR und Restriktionsenzymen sehr einfach und schnell analysierbar wurden, wurden phänotypische Merkmale wie z.B. die Vererbung einer bestimmten Krankheit in Mensch und Tier mit Hilfe kosegregierender DNA-Polymorphismen kartiert. Für die Kartierung eines monogenetisch vererbten Merkmals der Maus in eine genomische Region von ca. 25cM sind 50 Tiere einer Rückkreuzung erforderlich. Die aufgetretenen Rekombinationsereignisse können mit Hilfe von ca. 80 polymorphen Markern nachgewiesen werden. Setzt man Mikrosatellitenmarker ein, müssen 4000 PCR-Reaktionen durchgeführt und auf Agarosegelen analysiert werden. Der Einsatz der aus dieser Arbeit erwachsenen polymorphen Proteinspots als Marker für die genetische Kartierung eines Merkmals würde die Analyse von 50 Maus-Gehirn 2D-Gelen der 50 Rückkreuzungstiere erfordern. Die Probenaufarbeitung der Gewebe, Erstellung von 50 2D-Gelen, und Analyse von 80 ausgewählten, polymorphen Proteinspots würde einige Monate in Anspruch nehmen, während die Genotypisierung mit Mikrosatellitenmarkern im Bereich einiger Wochen läge. Proteinpolymorphismen bilden daher den zeitlichen und methodischen Arbeitsaufwand betreffend keine Konkurrenz zu DNA Markern. Proteinpolymorphismen sind zudem im Vergleich zu den Tandemrepeats der Mikrosatellitenmarker nicht hochpolymorph, sondern zeigen nur wenige Allele pro Genort, so dass je nach Mauskreuzung viele Proteinmarker aufgrund des fehlenden Polymorphismus nicht informativ und damit wertlos für die Kartierung eines Phänotyps sind.

Die Bedeutung der Kartierung von Proteinpolymorphismen liegt im Nachweis von kodierenden Genorten. Hier kommt gegenüber der Kartierung von cDNA- und EST-Sequenzen noch ein Vorteil hinzu: cDNA Kartierungen werden durchgeführt, indem der Genort mit spezifischen Primerpaaren von genomischer DNA der Tiere einer Kreuzung amplifiziert wird oder indem die cDNA gegen DNA von Tieren einer Kreuzung hybridisiert wird (Southern Blot). Beide Methoden beinhalten die Möglichkeit, dass der genomische Ort eines Pseudogens anstelle des eigentlichen Genortes amplifiziert wird oder Hybridisierungssignale zeigt und damit die cDNA falsch kartiert ist. De Vienne et al. (de Vienne et al. 1996) geben dies als eine wichtige Motivation für ihre Proteinkartierung an. Die 2DE liefert Kartierungsergebnisse echter, translatierter Sequenzen im Gegensatz zur anonymen Markierung des Genoms mit DNA-Markern (Mikrosatelliten und STS-Markern) oder funktionsunbekannten cDNAs. Gleichzeitig ist bekannt, aus welchem Gewebe und welcher Fraktion der kartierte Polymorphismus stammt.

4.4 Die Identifizierung kartierter, polymorpher Proteinspots mit MALDI Massenspektrometrie enthüllt abweichende Kartierungsergebnisse.

Kartierte Proteinspots wurden mit MALDI Massenspektrometrie von Dr.C.Gauss oder Dr.E.Krause identifiziert. Die Identifizierung ermöglichte es, zu prüfen, ob für das identifizierte Protein schon eine genetische Kartierungsposition in Maus oder Mensch bekannt war. Sieben Proteinspots zeigten eine neue, von der bisher bekannten Kartierungsposition Position. Die Abweichung kann durch verschiedene Ursachen bedingt sein: (a) Die Identifizierung des Proteinspots mit MALDI Massenspektrometrie war falsch, (b) Die referierte Kartierung entspricht der eines Pseudogens, während wir das echte Strukturgen kartiert haben, (c) unsere abweichende Kartierung entspricht der Kartierung eines Locus, der den Proteinphänotyp verursacht (Modifier- oder Regulatorischer Genort).

Zur Identifizierung der Spots mittels Massenspektrometrie lässt sich sagen: Spots, die nicht zweifelsfrei identifiziert werden konnten, wurden wiederholt analysiert.

Wenn wir uns auf die bereits vorhandenen, bekannten Kartierungsdaten eines Proteins beziehen, so stammen diese von der Kartierung von DNA-Sequenzen mittels Hybridisierungsexperimenten oder der Amplifizierung des Genorts mit sequenzspezifischen Primern. Da Gene oft Sequenzen enthalten, die mehrmals im Genom vorkommen (Repeats, homologe Regionen zu nahverwandten Genen oder Pseudogenen), können wir - um auszuschliessen, dass es sich bei der referierten Kartierungsposition um die irrtümliche Kartierung eines Pseudogenlocus handelt - die näheren Umstände der Kartierung überprüfen: Kartiert die cDNA in eine chromosomale Region, die an mehreren, benachbarten Loci konservierte Syntenie zu orthologen Loci des menschlichen Genoms zeigt? Falls man davon ausgeht, dass die Kartierung des referierten Genlocus korrekt ist, kommt als Erklärungsmöglichkeit für die abweichende Kartierung unserer Proteinspots nur die dritte Möglichkeit in Frage: Der Polymorphismus wird durch einen weiteren, entfernt kartierenden Genort verursacht und liegt nicht in der Sequenz des identifizierten Proteins selbst begründet.

4.4.1 Die abweichende Kartierung des polymorphen Proteinspots mV147

Die Variante mV147 hat die SpotID B5_463 des 2D-Referenzgels von Mausgehirn (cytosolische Fraktion) (Gauss et al. 1999). Spot B5_463 wurde als Acat1 identifiziert und kartiert in dieser Untersuchung proximal von D9Nds8/Crya2 während Acat1 von Xia (Xia et al. 1996) distal von Crya2 kartiert wurde (ein Rekombinationsereignis Crya2 - Acat1). Crya2 ist identisch mit dem Locus D9Mit96. Crya2/D9Mit96 und D9Nds8 wurden bisher jedoch nicht in einer Kreuzung kartiert, daher kann keine endgültige Reihenfolge der beiden Marker festgelegt werden.



4.4.2 Die abweichende Kartierung des polymorphen Proteinspots paV410

PaV410 kartiert auf Maus Chromosom 17 und wurde als als cytosolische Acyl-CoA Thioester Hydrolase (Ratte) identifiziert. Das homologe menschliche Protein hat 95% Homologie zum identifizierten Rattenprotein und kartiert auf Chromosom 1, D1S214-D1S244 (16,4 - 22,9 cM), die gesamte Region ist homolog zu Maus Chromosom 4. Da die gesamte Region von Mensch und Maus viele Orthologe aufweist, kann man die irrtümliche Kartierung eines Pseudogens der <u>menschlichen</u> Acyl-CoA Thioester Hydrolase ausschliessen (http://www3.ncbi.nlm.nih.gov/Homology/index.html). Unsere genetische Kartierung des Proteinspots wurde basierend auf Daten von 35 Rückkreuzungstieren gewonnen und ist sicher.

Als Erklärungsmöglichkeiten ergeben sich folgende: (a) wir haben mit paV410 ein Protein identifiziert und kartiert, das ein eng verwandtes, dupliziertes, mausspezifisches Homolog darstellt und weder in Ratte noch Mensch existiert. (b) wir haben nicht die Acyl-CoA Thioester Hydrolase selbst kartiert, sondern ein Protein, das die Expression der Acyl-CoA Thioester Hydrolase reguliert. In welche Stoffwechselwege ist die Acyl-CoA Thioester Hydrolase eingebunden, welche Proteine kommen als Kandidaten für ihre Regulation in Frage? Welche QTLs kartieren bei Marker D17Mit25, dem Centromer von Mauschromosom 17? Es finden sich keine QTLs in Centromernähe von Chromosom 17, die mit unserem Proteinphänotyp zusammenfallen.

Die Acyl-CoA Thioester wird im Gehirn konstitutiv exprimiert während das auf Aminosäureebene zu 94% homologe Enzym in der Leber nur nach Induktion gebildet wird. Yamada (Yamada et al. 1997) vermutet, dass die beiden Enzyme, die sich auch in ihrer Substratspezifität unterscheiden, vom selben Gen durch alternatives Splicing der mRNA entstehen - ein interessanter Mechanismus, um gewebespezifische Expression eines Gens zu regulieren. Die beobachtete Mengenvariation des Proteins in Gehirnen der beiden Mäusestämme B6 und MSP kann damit nicht erklärt werden.

Wie sicher ist unsere Kartierung: besteht Kopplung zu anderen Regionen des Genoms?

Linked Markers in Sequence at min LOD 4.00, max Distance 50.0					
Marker-1	Marker-2	Theta	LOD	cM	
pav410s	b5_42m17	0.03	8.56	2.94	
pav410s	cr15g19a	0.04	6.27	3.85	
pav410s	cr9h7x	0.03	8.56	2.94	
pav410s	pav410a	0	10.54	0	
pav410s	pav413a	0	9.33	0	
pav410s	pav433a	0.03	7.41	3.33	
pav410s	mv126a	0.03	8.56	2.94	
pav410s	mv191a	0.03	8.56	2.94	
pav410s	mv220a	0.03	8.56	2.94	
pav410s	mv347a	0.03	8.28	3.03	
pav410s	mv82a	0	10.54	0	
pav410s	mv95a	0.03	7.99	3.13	
pav410s	D17Mit25	0	9.93	0	
nav/110e		0.12	4.64	13.88	
11> near n	D17Mit9				
11> near p Linked Mar	D17Mit9 av410s:3 rkers in Seque	ence at mir	n LOD 3.00	, max Distanc	e 50.0
11> near p Linked Mar	D17Mit9 av410s:3 'kers in Seque <u>Marker-2</u>	ence at mir Theta	n LOD 3.00	, max Distand	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s	D17Mit9 pav410s:3 rkers in Seque <u>Marker-2</u> b5_42m17	ence at mir <u>Theta</u> 0.03	n LOD 3.00 LOD 8.56	, max Distand <u>cM</u> 2.94	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s	D17Mit9 av410s:3 rkers in Seque <u>Marker-2</u> b5_42m17 cr15g19a	ence at mir <u>Theta</u> 0.03 0.04	LOD 3.00 LOD 8.56 6.27	, max Distanc <u>cM</u> 2.94 3.85	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque <u>Marker-2</u> b5_42m17 cr15g19a cr9h7x	ence at mir <u>Theta</u> 0.03 0.04 0.03	LOD 3.00 LOD 8.56 6.27 8.56	, max Distanc <u>cM</u> 2.94 3.85 2.94	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque <u>Marker-2</u> b5_42m17 cr15g19a cr9h7x pav410a	ence at mir <u>Theta</u> 0.03 0.04 0.03 0	LOD 3.00 LOD 8.56 6.27 8.56 10.54	, max Distanc <u>cM</u> 2.94 3.85 2.94 0	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque <u>Marker-2</u> b5_42m17 cr15g19a cr9h7x pav410a pav413a	ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0 0	LOD 3.00 8.56 6.27 8.56 10.54 9.33	, max Distanc <u>cM</u> 2.94 3.85 2.94 0 0	ce 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a	ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0 0.03	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41	, max Distanc <u>cM</u> 2.94 3.85 2.94 0 0 3.33	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a	ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0 0.03 0.18	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14	, max Distanc 2.94 3.85 2.94 0 0 3.33 22.6	e 50.0
11> near p Linked Mar <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a	Theta 0.03 0.04 0.03 0 0 0.03 0.03 0.18 0.03	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56	, max Distanc 2.94 3.85 2.94 0 0 3.33 22.6 2.94	ce 50.0
11> near p Linked Man <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a	Theta 0.03 0.04 0.03 0 0 0.03 0.18 0.03 0.03 0.03	LOD 3.00, 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56	, max Distanc 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94	ce 50.0
11> near p Linked Man <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a mv220a	Ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0.03 0.18 0.03 0.03 0.03 0.03 0.03	LOD 3.00, 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.56	, max Distanc 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94	ce 50.0
11> near p Linked Man Marker-1 pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a mv220a mv347a	Ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0.03 0.18 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.56 8.28	, max Distanc 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94 3.03	e 50.0
11> near p Linked Man Marker-1 pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a mv220a mv347a mv82a	Ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0.03 0.18 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.04 0.03 0.03 0.04 0.03 0.03 0.04 0.03 0.03 0.04 0.03	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.56 8.28 10.54	, max Distance 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94 3.03 0	e 50.0
11> near p Linked Man <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a mv220a mv347a mv82a mv95a	Ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0.03 0.18 0.03	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.56 8.28 10.54 7.99	, max Distance 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94 3.03 0 3.13	e 50.0
11> near p Linked Man Marker-1 pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav413a pav433a pav500a mv126a mv191a mv220a mv347a mv82a mv95a D17Mit25	Ence at mir <u>Theta</u> 0.03 0.04 0.03 0 0.03 0.18 0.03 0.03 0.03 0.03 0.03 0.03 0.03 0 0.03 0 0.03 0 0 0 0 0 0 0 0 0 0 0 0 0	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.56 8.28 10.54 7.99 9.93	, max Distance 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94 3.03 0 3.13 0	e 50.0
11> near p Linked Man <u>Marker-1</u> pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s pav410s	D17Mit9 av410s:3 rkers in Seque b5_42m17 cr15g19a cr9h7x pav410a pav410a pav433a pav500a mv126a mv191a mv220a mv347a mv82a mv95a D17Mit25 D17Mit9	Theta 0.03 0.04 0.03 0 0 0.03 0.18 0.03 0.03 0.03 0.03 0 0.03 0 0.03 0 0.03 0 0.03 0 0.03 0 0.03 0 0.03 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0	LOD 3.00 8.56 6.27 8.56 10.54 9.33 7.41 3.14 8.56 8.56 8.56 8.28 10.54 7.99 9.93 4.64	, max Distance 2.94 3.85 2.94 0 0 3.33 22.6 2.94 2.94 2.94 2.94 3.03 0 3.13 0 13.88	e 50.0

Abb.4.4.2 Kopplung der Variante paV410. Die Kopplung der Variante paV410 an den centromernahen Bereich von Chromosom 17 wird demonstriert. Der Befehl "near" des Mapmaker Programms sucht für paV410 aus dem gesamten Datensatz der Proteine, Ankermarker, IRS- und Mikrosatellitenmarker diejenigen heraus, die mit einem LOD-Wert von 3 bzw. 4 Kopplung zeigen.

Sowohl die Analyse mit LOD-Wert 4 wie auch die mit LOD-Wert 3 als Grenzwert findet ausschliesslich Kopplung der Variante PaV410 zu Proteinen und Markern von Chromosom 17 und macht damit einen Fehler in unserer Kartierung sehr unwahrscheinlich.

4.4.3 Die abweichende Kartierung des polymorphen Proteinspots mV134

mV134 zeigt Kopplung an einen Marker auf Chromosom 8, sie wurde als Carbonische Anhydrase identifiziert. Diese ist in der Maus auf Chromosom 3 lokalisiert.

16> near mv134a:3 Linked Markers in Sequence at min LOD 3.00 , max Distance 50.0						
Marker-1	Marker-2		Theta	LOD	сM	
mv134a	pav227a	chr3	0	6.62	0	
mv134a	pav229a	nicht kart.	0.09	5.57	10.03	
mv134a	mv138a	chr3	0.24	3.77	33.07	
mv134a	mv169a	nicht kart.	0.1	6.91	10.57	
mv134a	mv172a	nicht kart.	0.23	3.45	30.95	
mv134a	mv174a	chr3	0.08	4.24	9.12	
mv134a	mv285a	chr3	0.22	4.06	29.74	
mv134	D8Pas3	chr8		LOD 3.49	32.0	

Abb.4.4.3 Kopplung der Variante mV134. MV134 zeigt Kopplung zum proximalen Ende von Chromosom 3 sowie zu Marker D8Pas3 von Chromosom 8

MV134 ist als Carbonische Anhydrase identifiziert und zeigt Kopplung zum proximalen (centromer-nahen) Ende von Chromosom 3 sowie zu Marker D8Pas3 von Chromosom 8. Ihr Segregationsverhalten weicht stark ab von dem der Variante mV138, die ebenfalls als Carbonische Anhydrase identifiziert wurde und in die korrekte Region von Chromosom 3 kartiert (siehe Abb. Chromosomenkarten im Anhang). Die Kopplung zu Proteinen am centromernahen Ende von Chromosom 3 ist mit hohen Rekombinationsfrequenzen verbunden (mV134 - mV138: 33,07 cM, mV134 - mV285: 29,74 cM), die eine gemeinsame Kartierung beider Varianten in die proximale Region von Chromosom 3 ausschliessen

Die Existenz zweier unabhängiger Varianten mV134 und mV138, beide als Carbonische Anhydrase identifiziert jedoch mit stark unterschiedlichem Segregationsverhalten, legt die regulatorische oder modifizierende Einwirkung mindestens eines weiteren Proteins auf das Protein Carbonische Anhydrase nahe. Dies zeigt sich in der Varianten mV134, während der Polymorphismus der Variante mV138 einen Strukturgenpolymorphismus darstellt. Der folgende Polymorphismus wurde für die Carbonische Anhydrase beschrieben: es sind mehrere verschiedene Isoenzyme der Carbonischen Anhydrase bekannt, drei davon sind cytoplasmatisch. Interessanterweise haben Venta et al. (Venta et al. 1985) gefunden, dass in Mausinzuchtstämmen zwei Allele der Carbonischen Anhydrase II vorkommen: CAIIa und CAIIb, sie unterscheiden sich in der Aminosäure in Position 38: Glutamin (GIn), einer Aminosäure mit einer polaren,

ungeladenen Seitenkette oder Histidin (His), einer schwach basischen, geladenen Aminosäure. Diese Mutation könnte dem wie erwartet kartierenden Polymorphismus mV138 zugrunde liegen.

4.4.4 Die abweichende Kartierung des polymorphen Proteinspots paV419

PaV419 wurde auf Chromosom 7 kartiert und als Fructose-Bisphosphat-Aldolase C identifiziert. Diese wurde zuvor jedoch auf Maus Chromosom 11 bei 45cM kartiert (Segre et al. 1995).

Wie sicher ist unsere Kartierung: besteht Kopplung zu anderen Regionen des Genoms?

13> near pav419a:4 Linked Markers in Sequence at min LOD 4.00 , max Distance 50.0					
Marker-1	Marker-2		Theta	LOD	сМ
pav419a	b5_39o12	chr7	0.1	7.74	11.41
pav419a	cm43c14a	chr7	0.1	7.74	11.41
pav419a	pav160a	chr7	0.16	4.65	19.7
pav419a	pav417a	chr7	0.06	6.66	6.46
pav419a	pav78a	chr7	0.14	6.02	16.82
pav419a	mv276a	chr7	0.02	12.63	2.08
pav419a	mv279a	chr7	0.02	12.34	2.13
pav419a	mv43a	chr7	0.14	6.02	16.82
pav419a	D7Mit40	chr7	0.02	11.46	2.27
14> near pa Linked Mar <u>Marker-1</u>	av419a:3 kers in Sequ <u>Marker-2</u>	uence a	at min LOD 3.00 , Theta	max Distar	nce 50.0 cM
pav419a	b5_39o12	chr7	0.1	7.74	11.41
pav419a	cm43c14a	chr7	0.1	7.74	11.41
pav419a	pav160a	chr7	0.16	4.65	19.7
pav419a	pav26a	chr7	0.22	3.42	29.8
pav419a	pav417a	chr7	0.06	6.66	6.46
pav419a	pav78a	chr7	0.14	6.02	16.82
pav419a	mv276a	chr7	0.02	12.63	2.08
pav419a	mv279a	chr7	0.02	12.34	2.13
pav419a	mv343a	chr7	0.23	3.23	30.66
pav419a	mv43a	chr7	0.14	6.02	16.82
pav419a	D7Mit40	chr7	0.02	11.46	2.27

Abb.4.4.4 Kopplung der Variante paV419. Die Variante paV419 zeigt ausschliesslich Kopplung zu Markern von Chromosom 7

PaV419 zeigt ausschliesslich Kopplung zu Markern von Chromosom 7. Zu Markern von Chromosom 11 existiert keine Kopplung. PaV419 kartiert ca. 8cM proximal von D7Mit40 und distal von D7Mit171. Das

entspricht der Region zwischen 48 und 53 cM auf Chromosom 7 der Maus-Konsensuskarte (1999). Das Protein Fructose-Bisphosphat-Aldolase C wird gehirnspezifisch exprimiert. In Muskel und Leber werden die Isoenzyme Aldolase A und B exprimiert. Aldolase A ist wurde in dieser Arbeit auf Chromosom 6 kartiert (siehe Abb.4.4.4) und Aldolase B kartiert auf Chromosom 4 (Pilz et al. 1993). Zwischen Aldolase A und C besteht 81% Homologie, während zwischen Aldolase A und B, sowie B und C 70% Homologie besteht (Kukita et al. 1988). Der hier kartierte Locus auf Chromosom 7 kann also auch keinen Verwechslung mit der zu 81% homologen Aldolase A darstellen. Die Kartierung der AldolaseC auf Chromosom 11 erfolgte über direkte Sequenzierung eines Kontigs der Region und ist sicher (Segre et al. 1995). Eine weitere Bestätigung der Kartierung des Strukturgenlocus auf Chromosom 11 bedeutet die später vom Labor Prof.H.Meyer, Bochum durchgeführte Identifizierung der Variante mV159: sie wurde in dieser Arbeit ebenfalls auf Chromosom 11 in die entsprechende Region kartiert (siehe Anhang Chromosomenkarten).

Wir können annehmen, dass der von uns kartierte Genort auf Maus Chromosom 7 den Polymorphismus der Aldolase C verursacht. Etwas ähnliches wurde für einen anderen Genort auf Maus Chromosom 7 bei 49cM schon beschrieben (Nesbitt et al. 1979), und zwar für den Locus "hypoxanthine phosphoribosyltransferase mobility alteration" (Hma). Dieser Locus weist in den Mäusestämmen A/J, AKR/J, AU/SsJ, BALB/cJ, CBA/J, C3H/HeJ, DBA/2J, LP/J, RF/J, SEA/Gn, ST/BJ, und 129/J das Hma-a Allel auf und in den Mäusestämmen C57BL/6J, C57BL/KsJ, C58/J, LT/Sv, MA/MyJ, SJL/J, und SWR /J das Allel Hma-b. Das Allel Hma-a ist dominant über Hma-b. Tiere, die homozygot für Hma-b sind, zeigen bei der Elektrophorese von Erythrozyten- oder anderen Gewebsextrakten zwei Hauptbanden mit HPRT-Aktivität. Heterozygote oder für Hma-a homozygote Tiere haben in Erythrozytenextrakten 3 Banden, jedoch nur zwei Banden in allen anderen Gewebsextrakten. Dieser Fall demonstriert, dass der Polymorphismus eines Genortes von einem anderen, selbst polymorphen abhängen kann, und das in gewebespezifischer Weise. Da die Kartierung des kartierten Locus paV419 mit der von Hma zumindest überlappt, könnte man – falls Hma und paV419 sich als identische Loci erweisen würden - die Hypothese aufstellen, dass hier ein Locus gefunden wurde, der für einen pleiotropen Effekt auf mindestens zwei Gene, nämlich die Fructose-Bisphosphat-Aldolase C und die Hypoxanthine Phosphoribosyltransferase (HPRT) ausübt. Die HPRT ist ein Enzym, dessen Defiziens zur Ausbildung des Lesch-Nyhan-Syndroms führt. Dies ist eine Stoffwechselerkrankung, die dadurch verursacht wird, dass Purine nicht ausreichend entsorgt werden. Das führt zu Schädigung von dopaminergen Nervenendigungen und reduzierten Dopaminkonzentrationen (Torres-Jimenez et al. 1998; Smith und Friedmann 2000). In diesem Zusammenhang wurde allerdings bisher nicht von einer gemeinsamen Regulation von HPRT und von Fructose-Bisphosphat-Aldolase C Aktivität berichtet.

Abb.4.4.4 Kartierung der Fructose-Bisphosphat-Aldolase A, B und C

	Chrom	osom
Fructose-Bisphosphat-Aldolase A	Chr 6	paV398, diese Arbeit
Fructose-Bisphosphat-Aldolase B	Chr 4	Pilz et al.
Fructose-Bisphosphat-Aldolase C	Chr 11 Segre et al.;	
	Chr11	mV159, diese Arbeit
	Chr 7	paV419, diese Arbeit

4.4.5 Die abweichende Kartierung des polymorphen Proteinspots aV19

AV19 zeigt schwache Kopplung an Marker und kartierte Proteine von Chromosom 18 und ist als Triosephosphat Isomerase identifiziert worden. Diese wurde auf Maus Chromosom 6 kartiert (Bulfield et al. 1987).

Wie sicher ist unsere Kartierung: besteht Kopplung zu anderen Regionen des Genoms?

17> near av19a:4 Linked Markers in Sequence at min LOD 4.00, max Distance 50.0					
Marker-1	Marker-2		Theta	LOD	сМ
av19a	no linked		markers		
18> near av19a:3 Linked Markers in Sequence at min LOD 3.00, max Distance 50.0					ance 50.0
Marker-1	Marker-2		Theta	LOD	cM
av19a	mv10a	chr18	0.27	3.18	37.88
av19a	mv123a	chr18	0.27	3.18	37.88
av19a	mv124a	chr18	0.27	3.18	37.88
av19a	mv23a	chr18	0.27	3.18	37.88
av19a	mv264a	chr18	0.27	3.18	37.88
av19a	mv7a	chr18	0.27	3.18	37.88
av19a		- h = 1 0	0.05	0.00	05 54
aviou	D18Nds1	CULIS	0.25	3.23	35.51

Abb.4.4.5 Die Kopplung der Variante aV19. Schwache genetische Kopplung kann fuer die Variante aV19 ausschliesslich zu Markern und Proteinen des telomernahen Endes von Chromosom 18 nachgewiesen werden.

Alle Marker und Proteine zu denen aV19 eine schwache Kopplung mit hoher Rekombinationsfrequenz (35 - 37cM) zeigt, stammen vom distalen (telomernahen) Bereich von Chromosom 18. Zu Markern/Proteinen von Chromosom 3 existiert keine Kopplung.

Für aV19 kann keine genetische Kartierungsposition angegeben werden. Das legt die Vermutung nahe, dass es sich hier um einen Spotphänotypen handelt, der nicht nur von einem einzigen weiteren

Regulatorgenort beeinflusst wird wie die oben besprochene Variante paV419 (Kap.4.4.4), sondern von mehreren verschiedenen Genen bedingt wird. Einen solchen Locus nennt man einen quantitative trait locus (QTL). Eine Quantifizierung der quantitativen Variante aV19 in den Rückkreuzungstieren könnte zur Identifizierung von verschiedenen Intensitätsklassen des Spots und damit zur Kartierung von mehreren Loci führen, die alle Einfluss auf die gebildete Proteinmenge des Spots aV19 haben. Dies wurde schon von Damerval an Proteinpolymorphismen zweier Mais-Inzuchtlinien gezeigt (Damerval et al. 1994). 72 ausgewählte Proteinpolymorphismen wurden im Rahmen der Genauigkeit der Silberfärbung quantifiziert. 42 Loci zeigten einen bis fünf regulatorische Faktoren, die auf verschiedenen Chromosomen unabhängig vom Strukturgenlocus kartierten. Zum Beispiel kartiert der Mobilitätslocus (ProteinShiftLocus PSL) PSL22 auf Chromosom 9, während 4 weitere regulatorische Loci, die die Proteinmenge verantworten, auf Chromosom 5, 6, 8 und zwei auf 9.

4.4.6 Die abweichende Kartierung des polymorphen Proteinspots mV82

MV82 wurde auf Maus Chromosom 17 nahe des Markers D17Mit25 kartiert und wurde als ER60 identifiziert. Das Mausprotein "probable protein disulfide isomerase ER-60 precursor" (Protein Accession P27773) ist identisch mit dem Protein "glucose regulated protein Grp58" (Protein Accession NP 031978). Das menschliche Homolog GRP58 kartiert auf Chromosom 15q15, dieser Bereich ist jedoch homolog zu Maus Chromosom 2 bei 69cM, wo auch das Maus Grp58-Gen kartiert wurde (Briquet-Laugier et al. 1998). Die Kartierung wurde mittels Southernblot-Analyse einer Rückkreuzung von 67 Tieren durchgeführt. Die Sonde für den Nachweis von Restriktionsfragmentlängenpolymorphismen (RFLPs) in Hybridisierungsexperimenten wurde mittels spezifischen Primern von RNA aus Mausleber gewonnen. Dabei kann nicht ausgeschlossen werden, dass ein Pseudogenort identifiziert und kartiert wurde. Dagegen spricht jedoch, dass das Maus-Grp58 in eine Region konservierter Syntenie mit dem menschlichen Grp58 kartiert. Als Erklärungsmöglichkeit bleibt hier die Annahme, dass wir auf Chromosom 17 einen Genort gefunden haben, der den Polymorphismus mV82 verursacht.

4.4.7 Die abweichende Kartierung des polymorphen Proteinspots mV248

MV248 kartiert auf Chromosom 15 der Maus und wurde als Aconitat-Hydratase (Aco2) identifiziert. Von Brady (Brady et al. 1997b) wurde Aco2 auf Chromosom 10 bei 18cM kartiert. Das menschliche Homolog kartiert auf Chromosom 22q13.2-q13.31. Eine grosse Anzahl benachbarter Gene haben orthologe Partner auf Chromosom 15 der Maus. Eine Kartierung auf Chromosom 15 der Maus wäre daher aufgrund der konservierten Syntenie der chromosomalen Regionen von Mensch und Maus wahrscheinlich

(http://www3.ncbi.nlm.nih.gov/Homology/index.html). Brady et al. haben ihre Kartierung basierend auf PCR-Amplifizierung der 3'-untranslatierten Region durchgeführt (Brady et al. 1997a). Diese Region eines Gens enthält, da sie nicht translatiert wird, vermehrt Polymorphismen. Für die Amplifizierung wurde jeweils ein Primer für das Poly-A-Ende und ein genspezifischer Primer aus der 3'-untranslatierten Region eines ESTs eingesetzt. Die Analyse der PCR-Produkte von Tieren einer Mausrückkreuzung wurde mittels SSCP (single strand conformational polymorphism) Analyse durchgeführt. Zwei Gründe werden für die irrtümliche Amplifizierung eines Locus angegeben: falls die Sequenz der Primer mehrmals im Genom vorhanden ist, kann entweder ein nicht funktioneller Pseudogenlocus oder verwandeter Genlocus amplifiziert worden sein. Unsere Kartierung eines Proteinpolymorphismus liefert daher neuen Beweis dafür, dass Aco2 nicht auf Chromosom 10, sondern auf Chromosom 15 kartiert.

4.5 Zusammenfassend: Anwendungen der Proteinkartierungen

Neukartierungen	 14 Neukartierungen: siehe Tab.3.4.b) 2 Proteine sind weder in Maus, Mensch und Ratte kartiert (mV301, mV229) 2 Kartierungen waren abweichend von der Kartierungsposition des menschlichen Homologs (paV410, mV82) 10 Neukartierungen in der Maus wurden bestätigt durch die Kartierung des menschlichen Homologs
Kartierungen von Proteinphänotypen	 5 Kartierungspositionen in der Maus sind abweichend der bisher bekannten Kartierungsposition (mV147, paV419, aV19, mV134, mV248)
Nichtkartierbare Polymorphismen	von 1324 Polymorphismen zwischen <i>Mus spretus – Mus musculus</i> zeigen 744 Segregation, ca.600 sind nicht segregierend die 744 segregierenden wurden zu 441 Spotgruppen zusammengefasst: 409 Spotgruppen (664 Einzelspots) konnten kartiert werden

Tab.4.5.a) Anwendungen	der Proteinkartierungen:	Genkartierung und	Kartierung von I	Proteinphänotypen
	/ // • //		0		

4.5.1 Neukartierte Proteine als Kandidaten für Krankheitsloci

Die Familienanalyse der Vererbung von Merkmalen wie z.B. Bluthochdruck oder von Krankheiten wie der Alzheimer'schen Krankheit erbringt als Ergebnis eine genetische Kartierung des Merkmals auf eine oder mehrere chromosomale Positionen: Da längst nicht alle menschlichen Gene bekannt und kartiert sind auch die Fertigstellung der Genomsequenz des Menschen bedeutet nicht, dass die in der Sequenz verborgenen Gene alle bekannt wären - lohnt es sich weiterhin, neue Gene oder Proteine zu kartieren. Fällt die Kartierung eines Gens oder Proteins mit der Kartierung eines Krankheitslocus zusammen, so wird das Gen als Kandidat für die Beteiligung an der Krankheit betrachtet und auf Mutationen etc. untersucht. Neue bisher nicht kartierte Proteine stellen die beiden Proteinvarianten mV301 und mV229 dar. Sie wurden als Phosphoenolpyruvat Carboxykinase und Aldoketolase identifiziert und sind bisher weder in Maus noch Mensch kartiert. Für 10 weitere Varianten stellte sich nach der Identifizierung heraus, dass sie für die Maus neu kartiert worden waren. Ihre Kartierungsposition konnte anhand der Kartierungsposition des menschlichen Homologs in eine Region konservierter Syntenie bestätigt werden (siehe Tab.3.4.c). Die Identifizierung weiterer in dieser Arbeit kartierter Varianten wurde nachfolgend vom Labor Prof. Dr. J. Klose in Zusammenarbeit mit dem Labor Prof. Dr. Helmut Meyer in Bochum durchgeführt. Die Ergebnisse der Identifizierung mit Massenspektrometrie Tab.4.5 dargestellt und bringt zusätzlich zu den in Tab.3.4.b) aufgeführten 56 weitere neukartierte Proteine zum Vorschein: unter den etwa 300 bisher nicht identifizierten, kartierten Varianten sind daher noch viele weitere neukartierte Proteine zu erwarten.

Für die hier besprochenen, neu kartierten Proteine gehen wir davon aus, dass es sich um monogenetisch vererbte Polymorphismen handelt, die auf einem Polymorphismus im Strukturgen des identifizierten Proteins beruht und daher über Ladungs- oder Konformationsänderung eine Positionsverschiebung des Spots auf dem 2D-Gel herbeiführt. Theoretisch würden wir daher erwarten, dass diese Polymorphismen dem qualitativen Typ (mV) angehören. Da das nicht der Fall ist (siehe Tab.3.4.c): pav110, paV171, av 274Varianten; sowie Tab.2 im Anhang) muss man fragen, was den Mengenpolymorphismus einer quantitativen Variante verursacht: für aV-Varianten kann die Ursache des Polymorphismus in der Veränderung der Menge der exprimierten RNA liegen. Für paV-Varianten kann dieselbe Ursache angenommen werden, jedoch mit dem Effekt, dass der fehlende (Absence)Spot in der einen Mäusespezies nicht wirklich fehlt, sondern unter dem Detektionslimit liegt. Einer der beiden Spot einer paV-Varianten kann auch unter einem anderen Spot des 2D-Gels verdeckt liegen und die paV-Variante daher in Wirklichkeit eine mV-Variante darstellen. Ebenso können bedingt durch allelische Unterschiede der beiden Mäusespezies zusätzliche Isoformen eines Proteins (modifizierte Formen, zusätzlich auftretende Degradationsprodukte) in einer Spezies auftreten. Dies sind wahrscheinliche Interpretationen der paV-Varianten, denn man kann kaum annehmen, dass in Mus musculus 175 Spots und in Mus spretus 166 Spots nicht exprimiert werden und für die Gehirnfunktion keine Rolle spielen. Im Gegensatz zur Annahme von Damerval (Damerval et al. 1994) können Positionsvarianten (mV) also nicht als die

ausschliesslich monogenetisch vererbten, strukturgenbedingte Polymorphismen angesehen werden und quantitative Varianten (aV, paV) nicht ausschliesslich als polygenetische, quantitative Merkmale (siehe auch Kap.4.5.2 und Kap.4.5.3)

Tab.4.5.b) Liste der Proteine, die in dieser Arbeit in der Maus neu kartiert wurden. Die Massenspektrometrische Identifizierung wurde später vom Labor Prof. Dr. Helmut Meyer in Bochum durchgeführt. ^a Wenn mehrere Spots derselben Variante identifiziert wurden, wurde das hinter dem Variantennamen vermerkt.

	identifizierte Proteine	Variantenname ^a	Chromosom
1	Acyl-coA-binding protein	mV131BS x 2	Chr 1
2	cDNA clone image 334048 (5')	mV242BS	Chr 1
3	ATP synthase gamma chain,	mV233BS	Chr 2
	mitochondrial		
4	cDNA clone image 400521 (5')	mV185BS	Chr 2
5	E1-enzyme	mV349BS	Chr 2
6	Glutathione synthetase type A1	mV72BS	Chr 2
7	Alcohol dehydrogenase, class 3	mV146BS	Chr 3
8	Glutathione-S-transferase GT 8.7	mV139BS	Chr 3
		mV215BS	
		mV9BS	
		mV210BS	
9	Hemoglobin delta chain	mV174BS	Chr 3
10	40kD Peptidyl-prolyl cis-trans	mV148BS	Chr 3
	isomerase(Cyclophilin 40)		
11	Quinone oxidoreductase	mV262BS	Chr 3
12	cDNA clone image: 2922806 (5')	mV30BS	Chr 4
		mV31BS x 4	
		mV120BS	
13	Delta-1-pyroline-5-carboxylate	paV492BS	Chr 4
	dehydrogenase		
14	cDNA clone image: 1450561 (5')	mV241BS x 2	Chr 4
15	N-acetylneuraminic acid phosphate	mV259BS	Chr 4
	synthase	paV412BS	
16	Zinc finger protein, type C3HC4	mV365BS x 2	Chr 4
17	COP9 complex subunit	mV45BS x 3	Chr 5
18	Coronin-3	paV494BS	Chr 5
19	Dihydrolipoamide acetyltransferase	mV217BS	Chr 5
20	Endoplasmic reticulum protein 29	mV128BS	Chr 5
21	Leucine aminopetidase	mV298BS	Chr 5
22	26S Proteasome regulatory subunit P27	mV35BS	Chr 5
23	T-complex protein, zeta subunit	mV299BS	Chr 5
24	Acyl-coA thioesterase	paV440BS	Chr 6
25	Fructose Bisphosphate aldolase A	paV398BS	Chr 6
	(muscle)		

	identifizierte Proteine	Variantenname ^a	Chromosom
26	GTP-spec. succinyl coA synthe-tase beta subunit	mV355BS x 2	Chr 6
27	P100 co-activator	mV286BS	Chr 6
28	Succinvl-coA ligase alpha-chain.	mV200BS	Chr 6
	mitochondrial		
29	E-stop protein	mV276BS x 2	Chr 7
30	PKCq interacting protein	mV47BS	Chr 7
31	Sumo-1activating enzyme	mV357BS	Chr 7
32	Proline synthetase ass.	paV292BS	Chr 8
33	Glucose regulated protein 170K	mV110BS x 3	Chr 9
34	Tropomyosin alpha-chain, brain-3	mV28BS	Chr 9
		mV38BS	
35	KNP-I alpha protein	mV366BS	Chr 10
		mV367BS	
36	Elongation factor TU, mitochondrial	paV468BS	Chr 11
37	Heat shock protein 110 kDa	mV109BS x 7	Chr 11
38	Hypothetical 32.1kD protein	mV300BS	Chr 11
39	Phosphoglycerate mutase, muscle form	aV380BS	Chr 11
40	Pyridoxine 5'-phosphate oxidase	mV209BS	Chr 11
41	45kD Secretory protein	mV237BS	Chr 11
42	Serine racemase	paV189BS x 2	Chr 11
43	Serine/threonine-protein kinase	mV351BS	Chr 11
44	Acyl-coA thioesterase long chain, mitochondrial	paV118BS	Chr 12
45	Methylmalonate-semialdehyd	paV463BS	Chr 12
	dehvdrogenase	paV486BS	0
46	Tubulin beta-5 chain	aV174BS	Chr 12
47	Aldo-keto reductase	mV229BS	Chr 13
48	Calpastatin	mV113BS x 3	Chr 13
49	Phosphofructokinase-1 C isozyme	mV290BS	Chr 13
50	Ubiquilin	mV91BS x 2	Chr 13
51	Ubiquitin	mV155BS	Chr 13
52	Glia maturation factor beta	mV8BS	Chr 14
53	Phosphenolpyruvate carboxy-kinase, mitochondrial	mV301BS	Chr 14
54	Adenylosuccinate lyase	mV266BS	Chr 15
55	3-Mercaptopyruvate sulfotransferase	mV363BS	Chr 15
56	MHC tum-transplantation antigen	paV467BS	Chr 15
	P35B	-	
57	Succinyl-coA:3-ketacid-coenzyme A tranferase	mV149BS x 2	Chr 15

58	Phosphomannomutase	mV125BS	Chr 16
59	Cytosolic acyl coenzyme A thioester	paV410BS	Chr 17
	hydrolase		
60	3,2-Trans-enoyl-coA isomerase	mV220BS	Chr 17
61	unknown protein	mV95BS	Chr 17
62	cDNA clone image: 2182265 (5')	mV347BS	Chr 17
	similar to		
	Brain acteylcholinesterase putative		
	membrane anchor		
63	dnaK-type Chaperonine, mitochondrial	mV84BS	Chr 18
64	Hydroxysteroid 17 beta dehydrogenase	paV426BS	Chr 18
	4		
65	α -Internexin neuronal intermediate	mV81BS	Chr 19
	filament protein		
66	1 Cys-peroxidoxin protein	aV51BS	Chr 19
67	Glutathione-S-transferase homolog	aV269BS	Chr 19
68	Guanine deaminase	mV71BS	Chr 19
69	KIAA 1167 protein	mV105BS x 2	Chr x
70	Pigment epithelium-derived factor	mV247BS	Chr x
	70 Proteine	103 Proteinspots	70 chromosomale
	/011000110	105 110001150015	Positionen

4.5.2 Proteinphenotypen - wer verursacht den Phänotyp?

Massenspektrometrische Identifizierung von Proteinspots hat ergeben, dass durchschnittlich 2/3 bis die Hälfte der Spots eines 2D-Gels Isoformen desselben Proteins sind (Gauss et al. 1999; Tab.1 im Anhang). Diese Proteinspots werden als Spotfamilie bezeichnet (Klose 1999b). Als homogene Spotfamilie sollen Spots bezeichnet werden, die alle den gleichen Typ von Variation zeigen: die Variation kann entweder in einem Polymorphismus des Strukturgens bedingt sein oder von einem anderen Genort (Regulator/Modifier) verursacht sein. Dies trifft für alle in Kap.4.5.2 besprochenen neukartierten Proteine zu sowie für die abweichend kartierten Varianten paV419, aV19, mV134 und mV248 (Kap.4.4.1 – 4.4.7). Auf dieser Ebene befinden wir uns immer noch auf der Stufe der monogenen Vererbung von Merkmalen und haben einen einzigen für den Polymorphismus verantwortlichen Locus kartiert. In Tab.2 des Anhangs unter "Einzelspot-Varianten" sind weitere 14 Varianten aufgelistet, deren Kartierungsposition von der erwarteten abweicht. Diese Identifizierungen wurden vom Labor Prof. H. Meyer, Bochum durchgeführt.

4.5.3 Identifizierung von variantem und nichtvariantem Spot des 2D-Gels

	Variantenname und ID des identifizierten varianten Spot des 2D-Gels, Mr und pI	Mr und pI des Proteins aus Swissprot	ID des identifizierten <i>nichtvarianten</i> Spots des 2D- Gels, Mr und pI
Gamma-Enolase: 8 Spots identifiziert	mV52 B3_522 (40,41 KD, pI 5,08) Chr 6 mV54 B2_423 (38,90 KD, pI 4,96) Chr 6 mV55 B2_415 (39,07 KD, pI 4,89) Chr 6 mV73 B3_359 (48,62 KD, pI 5,08) Chr 6 mV57 C3_006 (35,86 KD, pI 5,07) Chr 6 mV364 B2_309 (45,28 KD, pI 4,86) Chr 6 mV68 B2_525 (38,06 KD, pI 4,94) Chr 15	47KD, pI 4,84	C3_003 (35,86 KD, pI 5,03)
Triosephosphat: Isomerase 4 Spots identifiziert	aV19 D3_152 (16,24 KD, pI 5,54) Chr 18	27KD, pI 7,29	C5_250 (27,37 KD, pI 6,56) C5_254 (27,37 KD, pI 6,37) C5_264 (27,24 KD, pI 6,32)
Fruktose bisphosphat Aldolase C: 4 Spots identifiziert	mV159 D6_120 (15,63 KD, pI 6,92) Chr 11 paV419 B5_464 (43,20 KD, pI 6,49) Chr 7	39KD, pI 7,10	B5_467 (43,28 KD, pI 6,47) B5_490 (42,42 KD, pI 6,48)

Tab.4.5.3 Identifizierung von variantem und nichtvariantem Spot des 2D-Gels

Die Identifizierungen von mehreren Spots einer heterogenen Spotfamilie zeigte, dass Isospots desselben Proteins mehreren Variationstypen angehören können und auf verschiedene genetische Positionen kartieren können. An den heterogenen Spotfamilien lässt sich deshalb der Einfluss weiterer Gene auf das zugrunde liegende Protein ablesen. Wie der Fall der Proteine in Tab.4.5.b) zeigt, erbrachte die Identifizierung einiger Varianten als Gamma-Enolase, dass die Varianten mV52, mV54, mV55 und mV73 einen Polymorphismus des Strukturgens reflektieren und wie erwartet auf Chromosom 6 bei Marker D6Mit24 (57cM)kartieren, während die Varianten mV57 und mV364 auf Chromosom 6 an zwei weiteren Positionen kartiert wurden. Der Polymorphismus mV68 kartiert auf Chromosom 15. Auch ein nichtpolymorpher Spot, C3_003 wurde als Gamma-Enolase identifiziert. Da dieser Spot ein geringeres Molekulargewicht als die polymorphen Spots aufweist, kann man annehmen, dass durch fortschreitenden Abbau des Proteins Fragmente entstehen, die den Polymorphismus nicht mehr tragen und daher nicht polymorph erscheinen.

Für die Triosephosphat Isomerase liegt der Fall genau umgekehrt: die drei nicht-polymorphen Spots des 2D-Gels zeigen die erwartete Grösse des Proteins von 27KD während ein kleiner Spot des Molekulargewichtes von 16KD eine Mengenvariation in den beiden Mäusestämmen zeigt. Da der Polymorphismus av19 zudem nicht an der Position (Chromosom 6) kartiert, die für die Triosephosphat Isomerase zu erwarten ist, sondern Kopplung zu Chromosom 18 zeigt wird die Annahme bestärkt, dass hier bei der Degradation des Proteins ein weiterer Faktor - z.B. zwei unterschiedliche Allele der

degradierenden Enzyme in den beiden Mäusestämmen - für den Polymorphismus verantwortlich gemacht werden kann.

Von den vier identifizierten Spots der Fructose-Bisphosphat-Aldolase zeigt ein Spot einen mV-Polymorphismus und kartiert auf Chromosom 11, eine Position, die von Segre gefunden wurde (Segre et al. 1995) und ein Spot einen presence/absence Polymorphismus. Dieser kartiert nicht wie erwartet auf Chromosom 11, sondern auf Chromosom 7 und zeigt ein höheres Molekulargewicht als der polymorphe Spot. Wie kann man sich die Bildung eines zusätzlichen, präsenten Proteinspots in einer der Mäusespezies erklären? Da der Spot paV419 eine geringe Abweichung in Molekulargewicht und pl von den nicht-varianten Gamma-Enolasespots zeigt, muss er eine posttranslationale Modifikation tragen. Beide Phänomene, die abweichende Kartierung der paV-Variante sowie die Abweichung in Molekulargewicht und pl machen die Einwirkung eines weiteren modifizierenden Genortes in einer der Mäusespezies sehr wahrscheinlich. Die Tabelle 2 im Anhang listet viele weitere Identifizierungen von Spots aus Spotfamilien auf und zeigt, dass bisher 25 heterogene Spotfamilien gefunden wurden.

Es stellt sich nun die Frage, welcher Anteil der Spots eines 2D-Gels lediglich den Strukturgenlocus reflektieren und welche Spots Ergebnis von Regulation und Modifikation sind.

Alle Proteine, von denen bisher ein oder mehrere Spots identifiziert wurden, die denselben Polymorphismus zeigen und daher eine homogene Spotfamilie bilden und deren Kartierung mit bisher bekannten Kartierungsdaten übereinstimmen, können als monogenetisch bedingter Strukturgenpolymorphismus betrachtet werden. Dies trifft für 70 der 409 kartierten Varianten zu (entsprechend 103 Einzelspots, siehe Tab.4.5.b). Polymorphismen von Spots, die abweichend kartiert wurden, müssen als Modifikator- oder Regulatorgen-bedingte Polymorphismen angesehen werden - das sind 20 Varianten (das sind fünf der in Kap.4.4.1 – 4.4.7 besprochenen Varianten, sowie die Varianten unter "Einzelspot-Varianten in Tab.2 des Anhangs). Spots einer heterogenen Spotfamilie, die wie die oben beschriebene Gamma-Enolase Spots enthält, die an verschiedene genomische Positionen kartieren, müssen ebenfalls als modifikator-/regulatorbedingte Spots betrachtet werden, ausgenommen der Polymorphismus, der an die Position des Strukturgens kartiert. Aus Tab.2 im Anhang lässt sich daher ableiten, dass von den 105 Spots der 25 Spotfamilien etwa 28 Spots einen Polymorphismus des Strukturgenlocus reflektieren, während der Polymorphismus von etwa 35 Spots eine Modifikation durch einen anderen Locus darstellt (Abb.4.5.1). Über die nicht-varianten Spots einer Spotfamilie kann keine Aussage gemacht werden, da sie nicht genetisch kartiert werden können. Welche Modifikation jeder einzelne analysierte Proteinspot aufweist, muss mittels Massenspektrometrie ermittelt werden. Was jedoch aus der genetischen Analyse der Varianten eines 2D-Gels hervorgeht, ist, dass auch ein einzelnes Protein nicht durch die Seguenz seines Gens festgelegt ist, sondern vielfältig reguliert und modifiziert wird und damit einen polygenen Phänotypen darstellt. Auf Einzelspotebene sind 70% der mit

Massenspektrometrie identifizierten polymorphen Spots strukturgenbedingt und 30% der Spotpolymorphismen sind durch einen Modifikator- oder Regulatorgenort bedingt.

Abb.4.5.3. Massenspektrometrisch identifizierten Polymorphismen des 2D-Musters von Mausgehirn (cytosolische Fraktion), die auf Strukturgenpolymorphismen beruhen und monogen bedingt sind und Polymorphismen, die auf der Einwirkung eines Modifier- oder Regulatorgenorts beruhen und vermutlich polygen bedingt sind.

Zahl der Proteinspots des 2D-Musters	ca. 8700 Spots			
Strukturgenpolymorphismen				
Zahl der Strukturgen-Polymorphismen (Tab.4.5.b):	103 Spots			
Strukturgenlocus in heterogenen Spotfamilien (Tab.2):	28 Spots			
	<u>131 Spots</u>			
Modifikator-/Regulatorbedingter Polymorphismus				
abweichend kartierte Varianten (Tab.4.5.a):	20 Spots			
Modifierloci in heterogenen Spotfamilien (Tab.2):	35 Spots			
	55 Spots			

4.5.4 Nichtkartierbare Proteinpolymorphismen

Von 8700 cytosolischen Gehirnproteinen der Maus zeigen zwischen *Mus spretus* und *Mus musculus* ca. 15 % einen Polymorphismus. Nur diese konnten auf mono- oder polygene Vererbung untersucht werden. Von den ca.1324 polymorphen Spots zeigten 744 Spots, die zu 441 Spotfamilien zusammen gefasst wurden (Klose 1999b)), Segregation in der F1- und F2-Generation und 409 (664 Spots) konnten genetisch kartiert werden. Für die 32 nicht kartierbaren Varianten sowie für die etwa 550 zwar polymorphen aber nicht konsistent segregierenden Spots kann folgende Erklärung gefunden werden: der grösste Teil dieser Varianten gehört dem quantitativen Typ an und dürften daher einem polygen bedingten Proteinphänotyp entsprechen, auf den andere Genorte Einfluss nehmen. Ein solcher quantitativer Locus lässt sich ohne exakte Quantifizierung des Spots in der Rückkreuzungsgeneration nicht erfassen. Ein weiterer Grund dafür, dass ca. die Hälfte der 1324 polymorphen Spots von *Mus musculus* und *Mus spretus* nicht kartiert werden konnten, liegt darin, dass sich dominant vererbte Spots von *Mus spretus* nicht erfassen lassen (s. Abb.4.5).



Abb.4.5.4 Die Segregation eines dominant vererbten Allels einer paV-Variante in der Rückkreuzungs-



Abschliessend kann man sagen: bei ca. 600/1324 (~45%) der polymorphen Spots beruht der Polymorphismus auf monogenetischer Ursache – entweder einem Polymorphismus des Strukturgens selbsts oder dem Polymorphismus eines Modifiergenorts - während die etwa 55% der nichtkartierten Proteinpolymorphismen durch einen mehrere Genorte (multifaktoriell) verursacht werden.

4.6 Die Biologische Aussage der Kartierung von Interspezies-Polymorphismen

Der Prozess der Artbildung

Die 2D-Elektrophorese wurde zur Untersuchung von nahe verwandten Drosophila-Arten benutzt, um Polymorphismen zu finden, die an der Ausbildung getrennter Arten beteiligt sind (Spicer 1988; Zeng und Singh 1993). Für Versuche, die Zahl der an der Artbildung beteiligten Gene zu schätzen, wurden nah verwandte Arten gewählt. Diese Experimente waren hauptsächlich in den 70er und 80er Jahren durchgeführt worden und zwar mittels Gelelektrophoresetechniken (Singh 1989). Den Anteil der artentrennenden Polymorphismen betreffend reichten die Ergebnisse von null % für teilweise isolierte Arten oder Unterarten bis 10% für Drosophila melanogaster und Drosophila simulans (Choudhary et al. 1992). Der Vergleich der Gehirnproteine der verwandten Arten Mus musculus und Mus spretus, die sich vor 3 Millionen Jahren zu getrennten Arten zu entwickeln begannen und daher nicht als nah verwandt bezeichnet werden können (siehe Abb.1.7.1), zeigte eine Polymorphismusrate von 15% (Klose 1999b). Ca. 40% der polymorphen Spots (das sind ca. 6% von 8700 Gehirn-Spots) konnten in dieser Arbeit kartiert werden. Welche Aussage lässt sich nun darüber machen, welche der Polymorphismen an der Artbildung beteiligt sein könnten? Ein grundlegendes Problem hierbei ist, relevante von zufälliger Variation zu unterscheiden. Die Spezies Mus musculus wurde ingezüchtet und ist als homozygote Inzuchtlinie C57BL/6 seit langem bekannt, auch die Spezies Mus spretus die für die EUCIB Rückkreuzung verwendet wurde, stammt aus einer kleinen Kolonie und weist daher einen Genpool auf, dessen Allele durch eine völlig zufällige Auswahl der Zuchttiere begründet wurde. Die Gesamtzahl der Polymorphismen zwischen Mus spretus und Mus musculus kann für die Speziesbildung nur als die weitest gefasste Menge an Kandidaten betrachtet werden, da sie innerartliche sowie zwischenartliche Variationen enthält. Sie ist eine Überschätzung. Die Schwierigkeit des Speziesvergleichs besteht also darin, die beobachteten Polymorphismen sowie zusätzlich vorhandene oder fehlende Proteinspots in ihrer Bedeutung für die Artbildung zu beurteilen. Zeng und Koautoren (Zeng und Singh 1993) kombinierten daher die molekulargenetische Analyse durch 2D-Elektrophorese geschickt mit der klassisch genetischen Segregationsanalyse des Phänotyps, Sterilität von Spezies-Hybriden der beiden Drosophilaarten D. simulans und D. sechellia. Sie untersuchten Spermienproteine von Parentalstämmen, Hybriden und fruchtbaren sowie unfruchtbaren Rückkreuzungstieren. Aus der Vielzahl der beobachteten Varianten (231 quantitative Variaten von ca. 1000 Spots einer 2D-Elektrphorese von Testis) wurden über Kosegregation der Varianten mit dem Phänotyp Sterilität 4-5 Kandidatengene ermittelt. Zeng und Koautoren ziehen daraus den Schluss, dass eine relativ kleine Anzahl von Genen, die die Gametogenese kontrollieren, ausreichend ist, um die Entwicklung von Hybrid-Sterilität zu initiieren.

Da für unsere Mäusekreuzung kein artentrennender Phänotyp bekannt ist, der sich in Gehirnproteinen zeigen würde – wie etwa unterschiedliches Paarungsverhalten – und dessen Segregation für die Rückkreuzungstiere ermittelt werden könnte, kann hier keine Unterscheidung zwischen

119

Polymorphismen getroffen werden, die bei der Artbildung eine Rolle spielten und solchen, die nicht daran beteiligt sind.

Zur Erklärung der Artenbildung existieren verschiedene Modelle, die sich darin unterscheiden, dass sie einige wenige, oder aber eine grosse Anzahl von beteiligten Genen postulieren. Klassische Theorien, die geographische Isolation als Ursache für die Artbildung annehmen, postulieren, dass die Bildung einer neuen Art in der Ansammlung vieler allelischer Unterschiede mit kleinem Effekt besteht (allopatrisches Modell) (Charlesworth et al. 1982). Das peripatrische Modell fordert den Austausch einer grossen Anzahl von Allelen und eine weitreichende qualitative Veränderung des Genpools (Mayr 1982). Sympatrische Modelle nehmen an, dass nur wenige Gene mit grossem Effekt an der Artbildung beteiligt sind (Tauber et al. 1977) oder wenige allelische Substitutionen mit grossem Effekt auf das Erreichen von reproduktiver Isolation, so wie es Zeng für die beiden Drosophila-Arten zeigte.

Das letztgenannte Sympatrische Modell wurde von Untersuchungen von Römer et al. unterstützt. Die Arbeit von Römer et al. (Roemer et al. 1997) erbrachte, dass das Auftreten zweier Mutationen in den Proteinen "Major Urinary Protein" (MUP) und "Olfactory Marker Protein" (OMP) ausreichend ist, um zwischen Mitgliedern derselben Mausspezies eine Fortpflanzungsbarriere zu errichten und Paarung zu verhindern.

Lassen sich die 409 Varianten, die in dieser Arbeit kartiert wurden, auf wenige regulierende, modifizierende Proteine zurückführen, in denen sich die beiden Arten unterscheiden und die damit eine Schlüsselrolle bei der Artbildung einnehmen könnten?

Hierzu müssten mehrere Spots derselben Kartierungsposition identifiziert worden sein, deren Kartierung in der Literatur unzweifelhaft bestätigt ist und von der unsere Kartierung abweicht. Dies würde den Nachweis eines Regulator- /Modifikatorgenortes erbringen, der die identifizierten Proteine alle mit einer Modifikation versieht und der eine Schlüsselrolle bei der Artentrennung einnehmen könnte.

Zwar wurden in dieser Arbeit 25 heterogene Spotfamilien identifiziert: d.h. Spots wurden als Isoformen desselben Proteins identifiziert; die Spots zeigen jedoch verschiedene Polymorphismen, die an verschiedene genomische Positionen kartieren. Aber es wurde kein Locus identifiziert, auf den eine grosser Anzahl der beobachteten Polymorphismen kartiert haben, die sich alle nach der Identifizierung als Proteine erwiesen, die abweichend auf verschiedene andere Locus kartierten. Von einer Häufung abweichend kartierender Polymorphismen müsste angenomnen werden, dass ihr Polymorphismus von einem globalen Modifikator- oder Regulatorgen verursacht wurde, das speziesspezifisch agiert.

Auflösung der genetischen Karte - warum kartieren viele Proteine auf gleiche Positionen?

Genetische Karten werden erstellt, indem jeweils für zwei Genorte die Anzahl der entkoppelten Vererbung der Allele analysiert wird. Die Zahl der 'Entkopplungen' (= Rekombinationen) zweier Genorte wird 'Rekombinationsfrequenz' genannt und in Centimorgan (cM) angegeben. Tritt in 100

Meiosen zwischen zwei Genorten ein Rekombinationsereignis auf, dann ist ihr genetischer Abstand als 1 cM Centimorgan definiert. Da die Proteinpolymorphismen in 64 B1-Tieren analysiert wurden beträgt die durchschnittliche Auflösung unserer Kartierung 1/64 = 1,56 cM. Findet man zwischen zwei Proteinvarianten in einem der 64 analysierten Rückkreuzungstiere eine einzige Rekombination beträgt ihr durchschnittlicher genetischer Abstand 1,56 cM und einer physikalischen Länge von 3 Megabasen. In einer DNA Sequenz von 3000 Kilobasen kann man jedoch viele Gene vermuten, wenn man damit rechnet, dass Gene durchschnittlich in Intervallen von 40 kb vorkommen, d.h. alle Spots, die auf dieselbe Position der genetischen Karte fallen, können theoretisch durchaus dicht gekoppelte Gene darstellen. Der gefundene Abstand von 1,56 cM entspricht bei 64 analysierten Tieren mit einer statistischen Wahrscheinlichkeit von 95% einem Abstand zwischen 0,4 und 8,8 cM. Mit einer statistischen Wahrscheinlichkeit von 68% trennt die beiden Proteinvarianten der Abstand von 1,2 und 5,3 cM. Die 95% Wahrscheinlichkeit gibt ein Intervall an, das dem Ausschluss von Positionen ausserhalb des Intervalls dient. Tritt zwischen Proteinvarianten in 60 Tieren keine Rekombination auf, so sind die beiden Loci durchschnittlich weniger als 1,5 cM voneinander getrennt. Mit 95% Wahrscheinlichkeit liegen sie nicht weiter als 5,9cM auseinander (Silver 1995). Nachgewiesen werden konnte die geringe Auflösung unserer Karte an folgendem Beispiel: Von der Spotgruppe aus 58 mV-Varianten und einer paV-Variante auf Chromosom 6 wurden 4 Spots als Eno2 identifiziert und 2 als Gapd. Gapd und Eno2 liegen auf der Konsensuskarte der Maus 2.5 cM voneinander entfernt (d.h. ca. 5 Megabasen). Die Kartierungsposition für die varianten Spots ist also korrekt, jedoch sind die beiden Genorte in keiner der 64 Mäuse durch ein Rekombinationsereignis getrennt worden. Unter den 59 Spots dieser Gruppen sind deshalb Spots zu erwarten, die zwischen Gapd und Eno2 liegende Gene repräsentieren. Dies sind neben DNA-Markern etwa 55 Loci auf der genetischen Karte der Maus (2000b).

Für Proteinvarianten, die ohne Rekombination untereinander auf derselben Position der Karte liegen, bestehen also drei Erklärungsmöglichkeiten: (a) die gleichkartierenden Varianten sind Abbauprodukte desselben Proteins, (b) die Varianten reflektieren Polymorphismen von benachbart liegenden Genen, die in einem maximalen Abstand von 5,9cM voneinander entfernt in derselben chromosomalen Region kartieren, (c) eine dritte Möglichkeit besteht darin, dass sie nur deshalb gemeinsam segregieren, weil ein modifizierendes Protein sie alle mit derselben Modifikation versieht, in diesem Falle kartiert man den Modifikator, der in einem Mausstamm vorhanden ist und im anderen nicht. Tab.1 des Anhangs nennt die Centimorganabstände zwischen Varianten der Karte, die in gleiche Regionen fallen. Varianten mit 0 cM Abstand zueinander zeigen keine Rekombination untereinander. Welche der Hypothesen für die einzelnen gleichkartierenden Varianten zutrifft, kann erst nach Identifizierung der Proteinspots mit Massenspektrometrie entschieden werden.

Analyse der Spotgruppen identischer, genomischer Position hinsichtlich der Abschätzung der Anzahl verschiedener Proteine auf einem 2DE- Gel

Eine wichtige Frage im Zusammenhang mit der 2-DE Technologie betrifft die Anzahl der verschiedenen Proteine einer Probe, die auf einem 2D-Gel getrennt werden können. Diese Anzahl hängt zum einen von der Probe und zum anderen von der verwendeten 2DE Technik ab:

Proteine von Zellen aus Zellkultur lassen sich unkompliziert in Urea-Chaps(-Detergens)-haltigem Puffer solubilisieren und ergeben übersichtliche, klare Proteinmuster, während für die Analyse das jeweilige Gewebe komplexer Gewebe ausgefeilte, auf zugeschnittene Proteinextraktionsmethoden ermittelt wurden. Der zweite entscheidende Faktor ist die Separationslänge des pH-Gradienten im Isoelektrischen Fokussierungsgel. Während mit Ampholyten, die an Acrylamidmoleküle gebunden sind (Immobilinen) der pH-Gradient in das Gel eingegossen wird (Gorg et al. 2000), baut sich der pH-Gradient bei Verwendung von freibeweglichen Trägerampholyten unter angelegter Spannung erst während der Elektrophorese auf. Der begrenzte Gehalt an unterschiedlichen Ampholyten in den käuflichen Immobilin-Mischungen macht eine Streckung des pH-Gradienten über eine bestimmte Gellänge hinaus nicht sinnvoll (Proteine werden in länglichen Spots fokussiert), während mit Trägerampholytmixturen ein pH Gradient von 3-10 über eine Gellänge von 40 cm aufgebaut werden kann. Die bisher höchste Auflösung, was die Anzahl von Proteinpots betrifft, wurde für Mausgewebe von Klose beschrieben (Klose und Kobalz 1995). Für eine in Cytoplasma-, Membran- und Kernfraktion getrennte Proteinextraktion von Mausgehirn konnten etwa 8500, 1700 bzw. 50 verschiedene Spots auf 40x 30 cm grossen 2D-Gelen aufgetrennt werden (s. Abb.4.6).

Gewebefraktionen	Zahl der Protein Spots			
Gewebenaktionen	Leber	Gehirn	Herz	
cytoplasmatische/ pufferlösliche Fraktion	9 204	8 458	4 790	
Harnstoff/CHAPS- lösliche Fraktion	1 975	1 692	1 470	
DNA-verdaute Restsuspension	73	50	40	

Abb.4.6 Zahl der Proteinspots je Maus-Gewebe und Fraktion nach Klose (Klose 1999b). Für die Fraktionen b) und c) wurden nur die jeweils zusätzlich auftretenden Spots gezählt

Wenn ein Säugetierorganismus ca. 30000-60000 Gene besitzt (Hattori et al. 2000; Liang et al. 2000) und wenn ein Zelltyp oder ein Gewebe ca. 20 000 Gene exprimiert, was stellen die 10 000 mit der 2D-Technik darstellbaren Proteinspots dar? Wieviele Spots des 2D-Gels repräsentieren ein einzelnes Protein, wieviele Spots sind Modifikationen oder Abbauprodukte desselben Proteins?

Die wichtigste Methode zur Beantwortung dieser Frage ist die Identifizierung der Proteinspots mit der Massenspektrometrie. Bisher wurden 341 Spots der cytosolischen Fraktion des Mausgehirn 2D-Gels

massenspektrometrisch identifiziert (Tab.1 im Anhang). Sie repräsentieren 175 verschiedene Proteine, d.h. ein Protein ist durch etwa 2 Spots auf dem Gel vertreten.

Welchen Beitrag kann die genetische Kartierung von Polymorphismen zur Beantwortung dieser Frage liefern? Die genetische Kartierung von 409 Varianten führte zu dem Ergebnis: 360 Varianten wurden auf den Chromosomen lokalisiert und 49 zeigten Kopplung zu je einem Chromosom. Von den 360 Varianten kartieren 201 Varianten in 82 grösseren und kleineren Gruppen von 2 bis zu 59 Spots. 159 Varianten kartieren separat und können als 1-Varianten-Gruppe gewertet werden. D.h. die 360 Varianten kartieren in 241 (82+159) Gruppen. Da für die genetische Kartierung Spots mit gleichem Variationstyp und identischer Segregation schon unter demselben Namen zusammengefasst worden waren (Klose unveröffentlicht), verstecken sich hinter den 360 positionskartierten Varianten 610 polymorphe Spots. Hieraus kann man den Schluss ziehen, dass durchschnittlich 2,5 Spots (610/241) von einem Protein abstammen. Massenspektrometrische Identifizierung von Spots und genetische Kartierung von polymorphen Spots kommen also zu ganz ähnlichen Ergebnissen.

Was tragen diese Ergebnisse zur Abschätzung der auf einem 2D-Gel darstellbaren exprimierten Gene dar? Gesetzt die Annahme, alle Spots, die auf eine Position kartieren, sind entweder Abbauprodukte oder Modifikationen desselben Proteins, so könnte man mit 8700/2,5 etwa 3480 verschiedenen Proteinen rechnen. Damit muss man annehmen, dass auf einem 2D-Gel der cytosolischen Fraktion von Mausgehirn von den 8700 Proteinspots nur ca 1/6 der geschätzten 20 000 exprimierten RNAs als Proteinspots sichtbar sind. Für jedes Protein treten jedoch einige oder wie bei LDH und Gamma Enolase eine Vielzahl von Modifikationen auf (heterogene Spotfamilien), die wie in Tab.2 des Anhangs gezeigt, auf der Wirkung anderer Geneorte beruhen. Damit erweist sich die 2D-Elektrophorese über die blosse Darstellung eines meoglichst umfassenden Katalogs von Proteinen hinaus vielmehr als ein Werkzeug zur Analyse von Protein-Interaktionen und Regulationbeziehungen zwischen den Proteinen eines Gewebes.

4.7 CDNA-Kartierungen – die physikalische Alternative

2D-Elektrophorese wird oft als Screeningtechnik eingesetzt, um Organismen, Organe und Zellen in verschiedenen Entwicklungs- oder Krankheitsstadien miteinander zu vergleichen. Als Ergebnis erhält man Proteinspots, die eine Volumen- oder Positionsveränderung aufweisen. Ein Anteil dieser Veränderungen werden – besonders, wenn man menschliche Gewebe untersucht – auf natürlichen Polymorphismen beruhen, die zwischen Individuen in grosser Zahl zu erwarten sind. Ein anderer Teil der beobachteten Veränderungen an Proteinspots sind der experimentellen Fragestellung zuzuschreiben, d.h. sie sind entwicklungs-, krankheits-, organabhängig und die eigentlich gesuchten Veränderungen.

Wir haben daher nach einer Möglichkeit gesucht speziell für krankheitsbezogene Studien, zwischen unwichtigen natürlichen und krankheitsbezogenen Spotveränderungen zu unterscheiden. Denn ist ein Krankheitsphänotyp kartiert: z.B. wurden für Hypertonie auf verschiedenen Chromosomen QTLs kartiert oder für Alzheimer durch Familienanalyse mehreren chromosomalen Regionen gefunden, die an der Krankheit beteiligt sind, so zeigt die chromosomale Kartierung eines veränderten Proteinspots, ob er als Kandidat für den beobachteten Phänotyp gelten kann.

Weil viele Proteinspots keinen Polymorphismus aufweisen und daher nicht genetisch kartiert werden können und da zudem die genetische Kartierung von Proteinspots von 2D-Mustern aufwendig ist, wurde nach einer Methode gesucht, die das Problem löst. Folgende Strategie wurde angewendet, um jedem beliebigen Spot eines 2D-Gels chromosomal zu lokalisieren:

Fünf identifizierte Proteinspots eines 2D-Gels der cytosolischen Fraktion des Gehirns von *Mus musculus* B6 wurden als Modellproteine für die exemplarische, physikalische Kartierung nichtpolymorpher Proteine eingesetzt. Für die identifizierten Proteinsequenzen oder deren DNA-Sequenz wurden Homologievergleiche mit der Datenbank dbEST (http://www.ncbi.nlm.nih.gov/dbEST/ (Boguski et al. 1993)) durchgeführt, um eine EST-Sequenz (expressed sequence tags) zu erhalten. EST-Klone wurden im Rahmen des I.M.A.G.E Konsortiums der Allgemeinheit zugänglich gemacht und sind über vier Verteiler weltweit erhältlich (Lennon et al. 1996). Die Kartierung der Protein-spezifischen EST-Klone erfolgte in zwei Schritten. Im ersten Schritt wurden die EST-Sequenzen radioaktiv markiert und gegen eine Maus BAC-Klonbank hybridisiert, die 8 Genomäquivalenten entspricht. Im 2. Schritt wurden die BAC-Klone als Substrat für IRS-PCR eingesetzt. Die erhaltenen Produkte wurden gegen YAC-Filter hybridisiert, die IRS-PCR Amplifikate mehrerer YAC-Klonbanken enthielten. Ebendiese Klonbanken wurden bereits vom Whitehead Institute (URL:http://carbon.wi.mit.edu:8000/cgibin/mouse/index) sowie von L.C. Schalkwyk und H.R. Himmelbauer am Max-Planck-Institut für Molekulare Genetik, Berlin, dazu benutzt, eine physikalische Karte des Mausgenoms zu erstellen (Schalkwyk et al., eingereicht bei Genome Research und http://www.molgen.mpg.de/~rodent/). Identifizierte YAC-Klone, die bereits einer chromosomalen Region oder einem YAC-Kontig zugeordnet worden waren, ergeben sehr genaue Kartierungspositionen für die cDNA und damit den Proteinspot des 2DE-Gels (s. Abb.3.5.3.d)).

Dieser Ansatz hängt von der Identifizierbarkeit des gewählten Proteinspots mit Massenspektrometrie ab. Die ausgewählten Spots waren mit MALDI Massenspektrometrie iedentifiziert worden (Gauss et al. 1999). Da zur Identifizierung des Proteins die Massen von Trypsin-verdauten Peptide verwendet werden können nur bekannte Proteine oder translatierte Gene identifiziert werden. Andere Massenspektrometrische Methoden erlauben jedoch auch die Gewinnung von Aminosäuresequenzinformation von Trypsin-verdauten Peptiden und deshalb die Identifizierung von cDNA-oder EST-Sequenzen, ohne dass das gesamte Gen oder Protein bekannt sein muss.

Eindeutig identifizierte Proteine wurden benutzt, um wie oben beschrieben, cDNA- oder EST-Sequenzen zu ermitteln. Die Begriffe cDNA-Sequenz und EST-Sequenz werden hier als Synonyme benutzt. Es handelt sich bei den cDNA-Klonen, die man als I.M.A.G.E. Klone bestellen kann, um zumeist unvollständige cDNA-Klone, für die Sequenzinformation vorliegt.

Die cDNA-Hybridisierung gegen YAC-Klone ist durchführbar, aber technisch nicht trivial (nicht gezeigt). Dieser Weg wurde getestet und erwies sich aus mehreren Gründen als nicht praktikabel: bringt man einen YAC-Klon auf eine Nylonmembran auf, so ist nur wenig YAC-DNA pro Kolonie vorhanden. Das Hybridisierungssignal ist daher oft sehr schwach. Zusätzlich ist die mehrmalige Verwendung von YAC-Koloniefiltern problematisch: weil ein YAC-Koloniefilter nach der Hybridisierung mit einer radioaktiven cDNA nicht dehybridisiert werden kann - es würde zuviel DNA abgewaschen - muss mit der nächsten Hybridisierung einige Wochen lang gewartet werden, bis das radioaktive Signal der vorigen Hybridisierung abgeklungen ist.

Aus diesem Grund wurde der Weg über die Anwendung der IRS-PCR beschritten. Die Hybridisierung der cDNA gegen die BAC-Koloniefilter liefert einige BAC-Klone mit genomischer DNA, die neben der cDNA-Sequenz weitere benachbarte DNA enthalten - ein BAC ist ca. 150kb gross. Die BAC-Klone dienen nun als Substart für IRS-PCR: da das repetitive Element B1 der Maus sehr häufig ist, konnte mit dem B1-spezifischen Primer B1R von 80% aller BAC-Klone mindestens ein und bis zu drei IRS-PCR Produkte amplifiziert und über ein Agarosegel gereinigt und ausgeschnitten werden. Diese dem cDNA Locus benachbarte Sequenz wurden radioaktiv markiert und - nicht gegen YAC-Klonfilter, sondern gegen IRS-PCR-Amplifikate von YAC-Klon-Pools hybridisiert. Mittels IRS-PCR werden YAC-Sequenzen amplifiziert und die Menge an YAC-DNA, die für die Hybridisierung auf dem Filter auf gebracht wurde, ist um ein vielfaches höher als auf Koloniefiltern. Die YAC-IRS-PCR-Amplifikate sind deshalb einfacher zu hybridisieren, man erhält viel stärkere Hybridisierungssignale, die radioaktiv markierte Hybridisierungsprobe (das BAC-IRS-PCR-Produkt) lässt sich mit salzhaltiger

Dehybridisierungslösung abwaschen und die YAC-IRS-PCR-Produkt-Filter lassen sich viele Male für Hybridisierungen verwenden.

YAC-Vektoren sind die Vektoren, in denen die grössten genomischen DNA-Stücke kloniert werden können (500-800kb DNA). Sie sind deshalb die Vektoren der Wahl, wenn von der Erstellung einer physikalischen Karte die Rede ist. Sie werden mit DNA-Markern (Mikrosatelitenmarker, IRS-PCR-Marker) typisiert und darauf basierend zu Kontigs zusammengefasst. Beim Kontig-Bauen aus YAC-Klonen wird die DNA-Sequenz eines Chromosoms in Form einer geordneten Reihe von YAC-Klonen nachgebaut. Dies nennt man die physikalische Karte eines Chromosoms. Das Auffinden von YAC-Klonen ausgehend von der cDNA bedeutet eine Lokalisierung der cDNA auf der physikalischen Karte der Maus. Da die YAC-Klone jedoch noch nicht sequenziert sind, werden Positionen dieser physikalischen Karte immer noch anhand der auf den YACs lokalisierten genetischen Marker angegeben. Die Kartierungspositionen, die hier für die fünf cDNA-Kartierungen präsentiert werden, sind die Positionen der YAC-Klone auf der Karte des Whitehead Instituts. Die YAC-Klone wurden mit Mikrosatelliten- und STS-Markern charakterisiert, diese Information ist auf der Webseite http://genome.wi.mit.edu abrufbar:

Mit dieser Kartierungsstrategie existieren zwei Probleme:

(a) 50% der YAC-Klone sind chimär, d.h. sie enthalten DNA von verschiedenen Chromosomen.

Um sichere Lokalisierungen der cDNA zu erhalten, müssen mehrere verschiedene YAC-Klone identifiziert werden, die demselben Kontig angehören.

(b) Ein anderes Problem stellt die Möglichkeit dar, dass über Hybridisierungsexperimente ein Pseudogenort entdeckt wurde anstatt des funktionellen Genortes. Die cDNA hybridisiert mit mehreren genomischen BAC-Klonen. Diese Klone wurden aufgrund ihrer IRS-PCR-Produkte und ihres Restriktionsfragmentmusters, das sich bei Rückhybridisierung mit der cDNA zeigt, in Klassen ähnlicher/verwandter Klone eingeteilt. In unserem Beispiel "Albumin" haben wir unter den BAC-Klonen zwei Gruppen von Restriktionsfragmenten gefunden (Klon 1,3,4,5 und Klon 2, siehe Kap.3.5.3) während diese fünf BACs sich aufgrund ihrer IRS-PCR-Produkte anders einteilen liessen: Gruppe 1 enthielt Klone 1,3,4 und Gruppe 2 Klone 2 und 5. Dies zeigte, dass alle BAC-Klone untereinander verwandt waren, ähnliche DNA-Sequenzen enthielten und daher aus derselben genomischen Region stammen.

Die Hybridisierung der IRS-PCR-Produkte gegen YAC-Filter identifizierte 26 YAC-Klone. Diese enthielten Mikrosatellitenmarker von 10 verschiedenen Chromosomen, fünf davon fielen auf Chromosom 5, Kontig WC5.38.

Um auszuschliessen, dass wir homologe genomische Regionen (Pseudogene oder auch das funktionelle Gen) übersehen haben, wurde ein Southernblot von genomische DNA der Maus angefertigt: genomische DNA der Maus wurde mit Restriktionsenzymen verdaut, auf einem Agarosegel getrennt, auf eine Membran geblottet und mit der jeweiligen cDNA hybridisiert. Die

126

Hybridisierungssignale wurden mit denen der BAC-Klon Blots (Abb.3.5.3.b) verglichen und zeigten weder für Albumin noch das Neurofilament triplet L Protein zusätzliche Banden.

Die Kreuzhybridisierung der cDNA mit verwandten Sequenzen kann also mit folgenden Überlegungen getestet werden:

- (a) wie oben beschrieben werden die BAC-Klone aufgrund ihres in Restriktionsfragmentmusters und ihrer IRS-PCR-Produkte in Gruppen eingeteilt: jede Gruppe kann aus einer anderen genomischen Region stammen
- (b) die Hybridisierung der cDNA mit einem weiteren genomischen Locus würde sich an der Zahl hybridisierender BAC-Klone zeigen: unserer BAC-Klon-Bibliothek enthielt 8-10 Genomäquivalente DNA. Pro cDNA wurden 10-15 BAC-Klone gefunden.
- (c) mittels Southernblot von genomischer DNA und von BAC-Klonen konnte gepr
 üft werden, ob die Hybridisierung der cDNA f
 ür beide ein vergleichbares Bandenmuster aufweist oder zus
 ätzliche Banden zeigt: zus
 ätzliche Banden im Southernblot von genomische DNA w
 ären ein Hinweis darauf, dass ein weiterer,
 ähnlicher Locus im Genom existiert.
- (d) die erhaltene Kartierungsposition der cDNA kann mit Kartierungsdaten orthologer Gene in anderen Spezies verglichen werden. Die cDNA-Kartierung wird bestätigt, wenn beide Kartierungspositionen in bekannten syntenen Regionen liegen (DeBry und Seldin 1996).

Die hier dargestellte Kartierungsstrategie muss gegen die Kartierung von cDNAs auf Maus-Strahlungshybriden abgewogen werden (Cox et al. 1990; Van Etten et al. 1999). Die vollständige Charakterisierung von Strahlungshybridzelllinien erfordert die Typisierung mit Markerzahlen im Bereich der 1000de. Bei Durchführung der in dieser Arbeit beschriebenen cDNA-Kartierungen war das käuflich erwerbbare Maus-Strahlungshybridlinien nur unzureichend mit ca. 300 Markern genomweit (McCarthy et al. 1997), 200 zusätzlichen Markern für Chromosom 17 (Schalkwyk et al. 1998) und zusätzlichen Markern für Chromosom 13 typisiert (Elliott et al. 1999). Seit Durchführung der hier präsentierten und teilweise veröffentlichten Arbeiten (Nock et al. 1999) ist die umfassende Charakterisierung eines Maus-Strahlungshybriden veröffentlicht worden (Van Etten et al. 1999). Für beide cDNA-Kartierungsmethoden, die hier beschriebene sowie die Strahlungshybrid-basierte stehen nun Resourcen (mit 2486 Markern vollständig charakterisierte, käuflich erwerbbare Maus-Strahlungshybridlinien sowie BAC- und YAC-Klonbanken, siehe auch http://www.rzpd.de) der Allgemeinheit zur Verfügung. Für die Kartierung einer cDNA auf Strahlungshybriden müssen spezifische Primer generiert werden: Primerdesign ist mit einer Misserfolgsrate von 50-75% behaftet. Für die Kartierung einer cDNA über die hier vorgestellte Hybridisierung von IRS-PCR-Produkten müssen YAC-Pool-Filter hergestellt werden. Bei beiden Methoden kann die irrtümliche Kartierung einer verwandten genomischen Sequenz nicht ausgeschlossen werden.