

GSF – Forschungszentrum für Umwelt und Gesundheit
Institut für Experimentelle Genetik
Neuherberg

Ein neuartiges Konzept zur genomweiten Kartierung von ENU-induzierten Mausmutanten im Hochdurchsatz

Matthias Klafren

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.rer.nat. Alfons Gierl

Prüfer der Dissertation:

1. Univ.-Prof. Dr.rer.nat. Martin Hrabé de Angelis
2. Univ.-Prof. Dr.sc.techn.ETH, Dr.agr.habil. Hans-Rudolf Fries
3. apl. Prof. Dr.rer.nat., Dr.rer.biol.hum.habil. Jerzy Adamski

Die Dissertation wurde am 27.12.2006 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 27.2.2007 angenommen.

Für meine Eltern

Für Nadine und Fabian

Inhaltsverzeichnis

Zusammenfassung	1
Summary	3
1. Einleitung	5
1.1. Die Maus als Modellorganismus für die funktionelle Genomanalyse	5
1.2. Techniken zur Manipulation des Mausgenoms	5
1.2.1. Herstellung gentechnisch veränderter Mäuse	5
1.2.2. Erzeugung von Mausmutanten mittels Gene-Trapping	6
1.2.3. Erzeugung von Mausmutanten durch ENU Mutagenese	6
1.3. Aufbau des Münchner ENU Mausmutagenese-Projekts (MEP)	8
1.4. Kopplungsanalyse bei Mäusen	10
1.4.1. Inzuchtstämme	11
1.4.2. Rekombinante Inzuchtstämme	11
1.4.3. QTL Analyse	11
1.4.4. Marker zur Genomkartierung	12
1.4.5. Sequenzierung des Mausgenoms und die großen SNP-Discovery Projekte	13
1.4.6. Informationsgehalt von SNPs und Mikrosatelliten	13
1.5. Theorie zur genomweiten Kopplungsanalyse	14
1.5.1. Nachweis von Kopplung in einem Satz von genomweit verteilten Markern	15
1.5.2. Sonderfall: Keine Rekombination	16
1.6. SNP Genotypisierung im Hochdurchsatz	17
1.6.1. Überblick über molekularbiologische Methoden	17
1.6.2. SNP Genotypisierung mittels MALDI-TOF MS	18
1.7. Der Phänotyp der Mauslinie SMA004	19
1.7.1. Struktur und Klassifizierung von Zmpste24	19
1.7.2. Struktur und Klassifizierung der Lamine	20
1.7.3. Rolle der Lamine im Zellkern	20
1.7.4. Definition von Laminopathien	21
1.7.5. Expression und posttranslationale Modifikationen der Lamine	22
1.7.6. Fehlprozessierung von Prälamina A	23
1.7.7. Mausmodelle zur Untersuchung von Lamin A und Progerin	24
1.7.8. Molekulare Ursache	24
1.7.9. Therapieansätze	25
1.8. Ziele der Arbeit	25
2. Material und Methoden	27
2.1. Struktogramme nach Nassi-Shneiderman	27
2.2. Automatisierte DNA Extraktion	28
2.2.1. Vorbereitung der Schwanzbiopsien	28

2.2.2.	Steuerung des Extraktionsvorgangs	28
2.2.3.	DNA Quantifikation mittels des Pegasus-Skripts	33
2.2.4.	DNA Quantifikation mittels des Skriptes dilution.pl	33
2.2.5.	Protokollierung der DNA Extraktion	36
2.3.	SNP Genotypisierung mittels MALDI-TOF MS	36
2.3.1.	Homologe Massenextension (hME)	37
2.3.2.	Einzelprimer-Massenextension iPlex	39
2.4.	Erstellung der Markersätze	41
2.4.1.	Markersätze für die hME Methode	41
2.4.2.	iPlex	42
2.5.	Standardisierung der Genotypisierung	43
2.5.1.	Vorbereitung der Genotypisierung	43
2.5.2.	Auswertung der Genotypisierungsergebnisse	43
2.6.	Arbeiten mit der MyGenotype Datenbank	44
2.6.1.	MyImporter	44
2.6.2.	MyMaintenance	44
2.6.3.	MyLine	45
2.6.4.	MyGenotype	45
2.7.	Molekularbiologische Methoden	45
2.7.1.	Manuelle DNA Extraktion	45
2.7.2.	RNA-Extraktion und Reverse Transkription	46
2.7.3.	Sequenzierung	46
2.7.4.	RT-PCR	47
2.7.5.	Quantitative RT-PCR	47
2.7.6.	Proteinextraktion und Western Blot	48
2.7.7.	Genotypisierung mittels Tetra-Primer ARMS PCR	48
2.8.	Röntgenmethoden	49
2.8.1.	Röntgenabsorptiometrie	49
2.8.2.	Computertomographie (CT)	49
2.9.	Material	50
2.9.1.	Puffer, Lösungen und Reagenzien	50
2.9.2.	Enzyme und Antikörper	52
2.9.3.	Geräte	52
2.9.4.	Reaktionskits	53
2.10.	Anwendungen, Programme und Datensätze	54
3.	Ergebnisse	57
3.1.	Advanced Retrieval Tool for SNPs (ARTS)	57
3.1.1.	Das Programm create.pl	59
3.1.2.	Das Frontend arts.cgi	62
3.2.	Eigenschaften des Markersatzes für die hME Methode	66
3.2.1.	Chromosomale Verteilung und Polymorphie der Marker	66
3.2.2.	Heterozygotie des Markersatzes für Kreuzungen anderer Mausstämme	68
3.3.	Eigenschaften des Markersatzes für die iPlex Methode	68
3.3.1.	Chromosomale Verteilung und Polymorphie der Marker	68
3.3.2.	Heterozygotie des Markersatzes für Kreuzungen anderer Mausstämme	69
3.4.	Die Erstellung der MyGenotype Datenbank	72

3.4.1.	Das Programm MyImporter	73
3.4.2.	Das Programm MyMaintenance	76
3.4.3.	Das Programm MyLine	77
3.4.4.	Das Programm MyGenotype	79
3.5.	Automatisierte DNA Extraktion	85
3.5.1.	Anwendung	85
3.6.	Ergebnisse Mauslinien	86
3.7.	Charakterisierung der Mausmutante SMA004	88
3.7.1.	Verpaarungsmodus	88
3.7.2.	Kartierung der Mutation	89
3.7.3.	Untersuchung der genomischen DNA-Sequenz von <i>Zmpste24</i>	91
3.7.4.	Untersuchung der mRNA von SMA004	92
3.7.5.	Western Blot Lamin A	95
3.7.6.	Röntgenologische Untersuchung von SMA004	96
3.7.7.	Gewichtskurve	97
3.7.8.	Quantifizierung von Zielgenen von Lamin A Interaktionspartnern	98
3.8.	Charakterisierung der Mausmutante ALI029	100
3.8.1.	Kartierung der Mutation	100
3.8.2.	Untersuchung der genomischen DNA Sequenz von <i>Gli3</i>	101
4.	Diskussion	103
4.1.	Design von Markersätzen für die Kopplungsanalyse	103
4.1.1.	Design einfacher Markersätze	103
4.1.2.	Design von Markersätzen mit hoher Heterozygotie	103
4.1.3.	Design einer Mischform	104
4.1.4.	Schwierigkeiten bei der Auffindung geeigneter Marker	106
4.2.	SNP Genotypisierung im Hochdurchsatz	107
4.2.1.	Verwaltung und Auswertung der generierten Daten	107
4.2.2.	Konzeptionsbeweise	108
4.3.	Gesamtkonzeption	110
4.3.1.	Schlussfolgerungen	110
4.3.2.	Ausblick	111
4.4.	SMA004 - Ein Modell für <i>Zmpste24</i> Defizienz	112
4.4.1.	Auffindung und Charakterisierung der Mutation	112
4.4.2.	Einfluß der Mutation auf den Phänotyp	113
4.4.3.	Ausblick	115
5.	Liste eigener Publikationen	117
6.	Abkürzungsverzeichnis	119
7.	Literaturverzeichnis	121
A.	Liste der verwendeten Primer	137
B.	Danksagung	147

Zusammenfassung

Im Rahmen des Münchner ENU Mausmutagenese-Projekts (MEP) konnten bis heute über 750 Mauslinien mit klinisch relevanten Phänotypen etabliert werden. Genomweite Kartierungsprojekte wurden bisher bei annähernd 100 Mauslinien durch die Genotypisierung von Mikrosatellitenmarkern durchgeführt.

Durch den Aufbau eines Systems zur Genotypisierung von Single Nucleotide Polymorphisms (SNPs) im Hochdurchsatz sollte es im Rahmen dieser Promotion ermöglicht werden, diese Kartierung effizienter und genauer durchzuführen. Es sollte erreicht werden, bereits in einem Schritt eine sehr genaue Eingrenzung des subchromosomalen Locus einer ENU-induzierten Mutation zu erhalten.

Die Erstellung eines Satzes von SNP-Markern, der zur effektiven Auswertung aller Kreuzungsvarianten des MEP geeignet war, gestaltete sich jedoch mit den zu Beginn der Arbeit vorhandenen Suchmaschinen schwierig. Während dieser Arbeit wurde daher ein Advanced Retrieval Tool for SNPs (ARTS) entwickelt, mit dem gezielt Markersätze für die Genotypisierung von Kreuzungen verschiedener Mausstämmen erstellt werden können. Im Gegensatz zu anderen Suchmaschinen können mit ARTS in wenigen Schritten SNPs und deren flankierende Sequenzen gefunden werden, die spezifische Anforderungen an die gewünschte Polymorphie und die Qualität der Daten erfüllen.

Es war möglich, mit Hilfe dieser Suchmaschine Markersätze zu generieren, bei denen bis zu 100% der Marker für die im MEP eingesetzten Mausstämmen informativ sind, ein hoher Anteil dieser Marker aber auch informativ für Kreuzungen anderer Mausstämmen ist.

Die SNP-Genotypisierung von 384 DNA-Proben wird mit bis zu 160 Markern in zwei bis drei Tagen durchgeführt und dadurch über 60.000 Genotypdaten erhoben. Ein System musste entwickelt werden, das den Benutzer bei der Validierung und Auswertung dieser Daten unterstützt und speziell für die Aufbereitung der Genotypdaten und die Kopplungsanalyse von Maus-Inzuchtstämmen ausgelegt ist. Die gewonnenen Datensätze können direkt mit diesem System analysiert oder zur Weiterverarbeitung in externen Programmen gespeichert werden.

Innerhalb des Zeitraums der Arbeit wurde mit Hilfe des hier vorgestellten Genotypisierungssystems eine Kopplung bei insgesamt 55 mutanten Mauslinien nachgewiesen. Aufgrund der Dichte der Markersätze konnte bei den meisten Mauslinien eine starke Kopplung mit einem der Marker festgestellt werden, wodurch der betroffene Locus statistisch auf einen Bereich von 3 bis 4 cM eingegrenzt werden konnte. Dies konnte im Rahmen von laufenden oder zwischenzeitlich abgeschlossenen Phänotypisierungsprojekten bei zwölf der untersuchten Mauslinien bestätigt werden. Der durchschnittliche Abstand des am stärksten gekoppelten Markers zu dem durch ENU Mutagenese betroffenen Gen betrug 5,9 Mb, was einer genetischen Distanz von etwa 3 cM entspricht.

Durch die Kopplungsanalyse mit SNPs konnte in der Mauslinie SMA004 eine Mutation in dem Gen *Zmpste24* gefunden werden. Diese Linie ist ein weiteres Modell für das Hutchinson-Gilford-Progerie-Syndrom (HGPS) und die Mandibuloakrale Dysplasie (MAD) und wurde weiterführend untersucht.

Als ein weiteres Beispiel für die gezielte Eingrenzung von Kandidatengenomen anhand der Kopplungsanalyse wird die Entdeckung einer *Gli3* Mutation bei der Polydaktylie-Mauslinie ALI029 angeführt.

Die im Rahmen dieser Arbeit etablierte SNP-basierte Genotypisierung ist mittlerweile ein integraler Bestandteil des MEP und kann im Zusammenspiel mit den entwickelten informatischen Werkzeugen zukünftig auch für die Analyse komplexer Kreuzungsschemata eingesetzt werden.

Summary

Within the Munich ENU Mouse Mutagenesis Project over 750 mouse lines with a clinically relevant phenotype have currently been established. Up to date, genome-wide mapping projects of approximately 100 mouse lines have been performed by typing of microsatellite markers.

A high throughput system for the genotyping of single nucleotide polymorphisms (SNPs) was established to perform these mapping projects in a more precise and efficient way during the course of this thesis.

The aim was to identify the subchromosomal locus of an ENU-induced mutation in a single step.

However, it has proven difficult to generate a suitable SNP marker set for the analysis of all crossing variants of the MEP with the search engines that were available at the beginning of this PhD thesis.

Therefore the Advanced Retrieval Tool for SNPs (ARTS) was developed to generate specific marker sets for the genotyping of crosses of different mouse strains during this thesis. In contrast to other search engines, ARTS can be used to identify SNPs and flanking sequences that fulfill specific requirements in terms of polymorphism and quality of the data.

By means of this search engine it was possible to generate marker sets for which up to 100% of the markers were informative for the mouse strains used in the MEP. A high portion of these markers was also informative for crosses of other strains.

SNP genotyping of 384 DNA samples with up to 160 markers is performed in two to three days and thus data on 60,000 genotypes is collected. A system was developed in this thesis that assists researchers in the validation and analysis of this data and that is specifically designed for the preparation of the genotype data for linkage analysis of mouse inbred strains. Datasets can either be analyzed directly by the system, or can be saved for further processing with external programs.

During this thesis linkage in 55 mutant mouse lines was demonstrated by means of the genotyping system introduced in here. Due to the density of these marker sets, strong linkage with one marker could be detected in most mouse lines, whereby the affected locus could be delimited statistically to a region of 3 to 4 cM. This could be validated for twelve of these mouse lines in ongoing or completed phenotyping projects. The average distance between the most strongly linked marker and the gene affected by ENU mutagenesis was 5.9 Mb, which complies to an approximate genetic distance of 3 cM.

By linkage analysis with SNPs, a mutation in the *Zmpste24* gene could be found in mouse line SMA004. This line is another model for the Hutchinson-Gilford progeria-syndrome (HGPS) and mandibuloacral dysplasia (MAD) and was further studied.

A further example for the systematic reduction of candidate genes by means of this linkage analysis is the identification of a *Gli3* mutation in the mouse line ALI029 having a polydactyly phenotype.

The SNP genotyping pipeline that has been established in this thesis has become an integral component of the MEP and can be used in the future for the analysis of more complex crossing schemes.

1. Einleitung

1.1. Die Maus als Modellorganismus für die funktionelle Genomanalyse

Seit Etablierung der ersten Inzucht-Mauslinie DBA zu Beginn des 20. Jahrhunderts wird die Maus als Modellorganismus zur Studie verschiedenster Aspekte der Biologie und Medizin eingesetzt. Insbesondere in der Tumor- und Transplantationsforschung spielten Mäuse eine entscheidende Rolle. So trug beispielsweise die wiederholte Rückkreuzung von histoinkompatiblen Mäusen auf einen ansonsten genetisch identischen Stamm letztlich entscheidend zur Aufklärung des „Major Histocompatibility Complex“ (MHC) bei [178].

Ungefähr 99% der im Menschen bekannten Gene können auch im Genom der Maus gefunden werden, wie sich bei der Analyse der Nukleotidsequenz dieser beiden Organismen herausgestellt hat [204]. Im Vergleich zu anderen Tiermodellen, mit denen die menschliche Physiologie unter Umständen besser abgebildet werden kann (z. B. *Rattus norvegicus*), konnte nur in der Maus in den letzten Jahrzehnten eine Vielzahl von einfachen Methoden zur Manipulation des Mausgenoms entwickelt werden. Wegen dieser Tatsache und aufgrund von weiteren Faktoren, wie der kurzen Reproduktionszeiten, der relativen Einfachheit der Haltung, und der hohen genetischen Verwandtschaft zum Menschen, ist die Maus der Modellorganismus, der am meisten zur Entschlüsselung der Funktion humaner Gene beiträgt.

1.2. Techniken zur Manipulation des Mausgenoms

1.2.1. Herstellung gentechnisch veränderter Mäuse

Vor ca. 30 Jahren wurden die ersten Experimente vorgestellt, in denen ein Transgen in die Keimbahn von Mäusen eingebracht wurde [82]. Seither wurden viele weitere Methoden zur Herstellung gentechnisch veränderter Mäuse entwickelt. Die interessanteste Option bietet die homologe Rekombination von Vektoren in das Genom von murinen embryonalen Stammzellen (ES Zellen) zur Herstellung von „Knockout“ Mäusen. Mit diesem Verfahren werden Gene entweder vollständig [192, 41] oder, z. B. mit Hilfe des Cre/lox Systems, gewebe- oder zeitspezifisch ausgeschaltet oder in ihrer Funktion verändert [140, 67]. Manipulationen dieser Art sind von großem Wert, wenn es darum geht, die Funktion eines bekannten Gens besser zu verstehen. Dieser gengetriebene „Reverse Genetics“ Ansatz setzt allerdings, im Unterschied zum klassischen, phänotypgetriebenen „Forward Genetics“ Ansatz, die genaue Kenntnis des zu manipulierenden Locus voraus.

1. EINLEITUNG

Die vollständige Sequenzierung des Mausgenoms [204] hat die Herstellung gentechnisch veränderter Mauslinien deutlich vereinfacht. Die durch diese Sequenzierung gewonnenen Erkenntnisse werden zukünftig nicht nur bei der Planung eines Experiments, sondern auch bei dessen Durchführung, insbesondere der Auffindung von korrekt rekombinierten ES Zellklonen eine größere Rolle spielen.

1.2.2. Erzeugung von Mausmutanten mittels Gene-Trapping

Das Gene-trapping ist eine Abwandlung der Herstellung von transgenen Tieren und eine effektive Methode zur Identifizierung und Funktionsanalyse von unbekanntem Genen. In diesem Verfahren integriert ein Vektor mit einer Kassetten, die beispielsweise aus einer Splice-Akzeptorstelle mit einem Reportergen und einer Polyadenylierungssequenz besteht, zufällig in das Genom von murinen ES Zellen [208, 183]. Anhand des Reportergens kann überprüft werden, ob der Vektor in das Intron eines Gens integriert wurde und die Transkription dieses Gens somit durch die Splice-Akzeptorstelle vorzeitig gestoppt wird.

Aufgrund der Tatsache, dass die Sequenz der integrierten Kassetten bekannt ist, kann das durch die Mutagenese betroffene Gen relativ einfach durch Sequenzierung bestimmt werden. Aus den Zellen der so etablierten mutagenisierten Stammzelllinien können nachfolgend Tiere erzeugt werden, in denen eine *in vivo* Analyse des Phänotyps der Mutation möglich ist. Es hat sich jedoch herausgestellt, dass es im Genom eine gewisse Zahl von „Hotspots“ gibt, also Loci, an denen bevorzugt eine Rekombination mit einem Gene-Trap Vektor stattfindet. Zur Kompensation dieses Problems wurde eine Vielzahl von speziellen Vektoren entwickelt (z. B. der FlEx-Vektor [170]), mit deren Hilfe erreicht werden soll, eine Gene-Trap Mutante für die meisten oder alle Gene zu erzeugen. Ein weiterer Nachteil dieser Methode ist, dass die eingeführte Mutation immer zur unvollständigen Transkription eines Gens führt. Subtile Mutationen, die sich beispielsweise nur auf ein Nukleotid beschränken, können auf diese Weise nicht erzeugt werden.

1.2.3. Erzeugung von Mausmutanten durch ENU Mutagenese

Bereits Mitte der 60er Jahre des 20. Jahrhunderts wurde *N*-Ethyl-*N*-nitrosoharnstoff (ethylnitrosourea, ENU) (vgl. Abbildung 1.1) als potentes Mutagen erkannt [43]. Die ursprünglichen Versuche dienten allerdings hauptsächlich den Studien von Tumorbildung und teratogenen Effekten in adulten Tieren.

Der Wirkmechanismus beruht auf der Alkylierung von Nukleinsäuren ohne metabolische Prozessierung [177]. Die Ethylgruppe des ENU kann auf Sauerstoff- oder Stickstoffradikale der verschiedenen Desoxyribonukleinsäuren übertragen werden [174]. Diese Veränderungen sind für sich genommen noch keine Mutationen, die Anwesenheit der Ethylgruppe kann jedoch zu Fehlern während der DNA-Replikation führen, was zu einer Fehlpaarung und letztendlich zur Manifestation einer Punktmutation führt (vgl. Abbildung 1.2 und [137]).

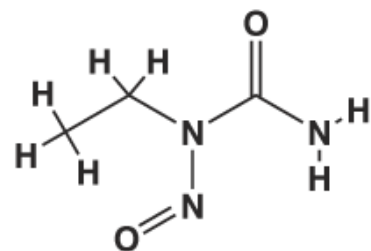


Abbildung 1.1:
Strukturformel von *N*-Ethyl-*N*-nitrosoharnstoff

Der gezielte Einsatz von ENU und anderen Reagenzien wie Ethylmethansulfonat (EMS) zur Erzeugung von Mutanten in *Drosophila melanogaster* wurde erstmals Mitte der 70er Jahre beschrieben [6]. Die genomweite Suche von Mutanten unter Anwendung von chemischer Mutagenese führte unter anderem zum Verständnis des Segmentierungsprozesses in *D. melanogaster* [138]. Die Eignung von ENU zur effektiven Mutagenese von Mäusen wurde von Russell *et al.* bewiesen [161, 158]. Über die gezielte Züchtung von ENU-induzierten Mausmutanten wurde erstmals 1984 berichtet [17]. Die Versuche beschränkten sich jedoch hauptsächlich auf die Identifizierung neuer Allele von drei bereits bekannten Loci in der T Region auf Chromosom 17.

Der erste großangelegte ENU-Mutagenescreen zur genomweiten Identifikation von Mutationen in *Danio rerio* wurde 1994 beschrieben [126]. Kurze Zeit später starteten die groß angelegten genomweiten ENU-Mutagenescreens in *Mus musculus* mit dem Ziel, neue funktionelle Elemente des Genoms zu entdecken, bzw. die Auswirkungen neuer Allele bereits bekannter Gene auf den Phänotyp zu studieren [135, 78].

Im allgemeinen Aufbau solcher Screens werden männliche Mäuse mit ENU mutagenisiert, was einige Vorteile für die Zucht und die Effizienz der Mutagenese mit sich bringt. Die mutagene Wirkung von ENU auf weibliche Organismen ist geringer als auf männliche Organismen [159]. Dies beruht auf der Tatsache, dass ENU hauptsächlich auf prämeiotische Stammzellen wirkt. Adulte weibliche Lebewesen haben jedoch keine sich teilenden prämeiotischen Stammzellen mehr, denn alle Oozyten sind bereits vorhanden. Sie ruhen in der Prophase I der ersten meiotischen Teilung und beginnen sich erst ca. drei Wochen vor der Ovulation zum reifen Ovum weiterzuentwickeln [156, Kapitel 2]. Im Gegensatz dazu werden während der Spermatozytogenese kontinuierlich primäre Spermatozyten aus spermatogonischen Stammzellen per Mitose gebildet, die über mehrere Zwischenschritte zu Spermatozoen reifen [156, Kapitel 2]. Während dieses Prozesses wird mit einer durch ENU ausgelösten Mutationsrate von $5 \cdot 10^7$ bis $2 \cdot 10^6$ pro Basenpaar gerechnet [33]. Diese Mutationsrate ist abhängig von der eingesetzten Dosis und dem Injektionsprotokoll [201, 31].

Unter der Annahme einer Länge des haploiden Mausgenoms von $3 \cdot 10^9$ bp und einer Mutationsrate von ca. $1 \cdot 10^{-6}$ ergäbe sich eine Gesamtzahl von ca. 3000 Mutationen pro Genom. Unter der weiteren Annahme, dass in der Maus ca. 30.000 codierende Sequenzen mit einer durchschnittlichen Länge von 1000 bp vorhanden sind, also ca. 1/100 des Genoms für Proteine codierend ist, würden theoretisch 30 codierende Nukleotide pro Genom von einer ENU-induzierten Mutation betroffen sein.

Viele dieser Mutationen werden jedoch „still“ sein, also keinen messbaren Effekt auf den Phänotyp haben. Diese Zahlen können durch die Beobachtung eines bestimmten Phänotyps in einer Anzahl von Nachkommen überprüft und daraus die durchschnittliche Größe der Zielregion eines Phänotyps berechnet werden [33]. Ein Vorteil der ENU Mutagenese gegenüber anderen Manipulationstechniken ist die Tatsache, dass durch die chemisch induzierte Einbringung von Punktmutationen natürlich vorkommende Allele simuliert werden. Gene werden nicht zwangsläufig ganz oder teilweise ausgeschaltet, sondern können in ihrer Funktion modifiziert werden, was zum teilweisen Funktionsverlust (Hypomorphie), einer gegensätzlichen oder dominant-negativen Funktion (Antimorphie) oder auch zu einer verstärkten Funktion (Hypermorphie) des korrespondierenden Proteins führen kann [137].

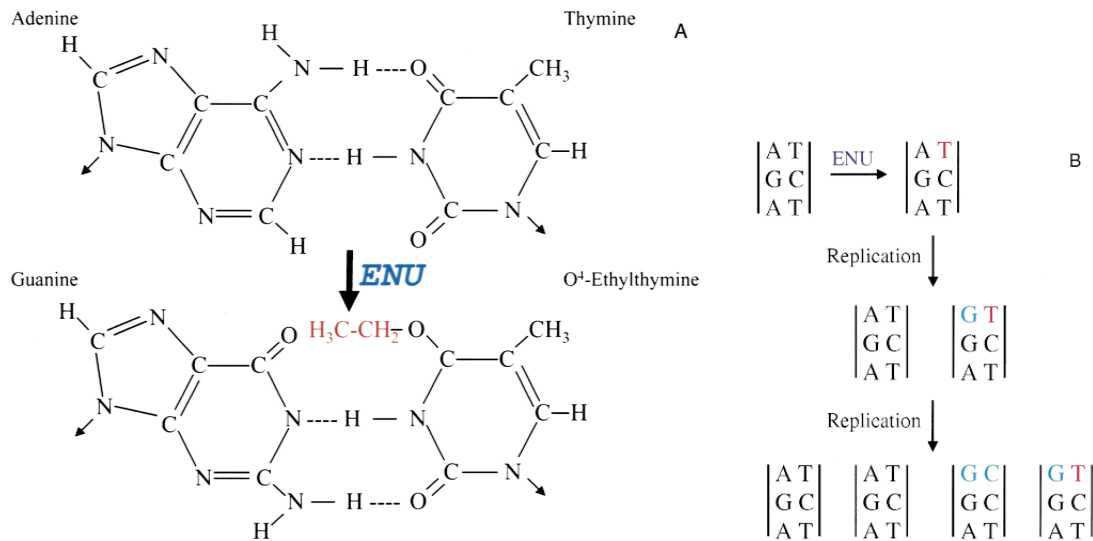


Abbildung 1.2: **A:** Wirkmechanismus von ENU. Die Alkylierung von Thymin führt zur Bildung von O⁴-Ethylthymine, welches als Cytosin fehlinterpretiert wird. **B:** Auswirkung des ENU-induzierten O⁴-Ethylthymins auf nachfolgende Replikationsschritte [137].

1.3. Aufbau des Münchner ENU Mausmutagenese-Projekts (MEP)

Ziel des MEP ist es, durch ENU Mutagenese im großen Maßstab neue Phänotypen zu identifizieren, diese möglichst genau zu charakterisieren und eine Korrelation zu der zugrunde liegenden ENU-induzierten Mutation zu erstellen. Dieser Ansatz hilft dabei, die Ursachen monogener Erkrankungen aufzudecken und neue funktionelle genomische Elemente und Varianten von genomischen Elementen bekannter Funktion zu charakterisieren.

Bis heute wurden im Rahmen des MEP über 750 Mauslinien etabliert, die auf eine ENU-induzierte monogenische Ursache zurückzuführen sind. Die Herleitung einer Genotyp-Phänotyp-Korrelation über die Genotypisierung von Mikrosatellitenmarkern gelang allerdings nur in ca. 100 Fällen. Um die stetig wachsende Kluft zwischen der Anzahl der Mauslinien und der Korrelation zu der genetischen Ursache zu schließen, wurden in dieser Arbeit SNP-basierte genomweite Markersätze zum Einsatz in einem Hochdurchsatzsystem entwickelt. Die Durchführung einer Kopplungsanalyse konnte auf diese Weise von mehreren Wochen für eine Mauslinie auf zwei bis drei Tage für mehrere Mauslinien verkürzt werden.

Injektionsprotokoll und Screening Die Effektivität der ENU Mutagenese ist abhängig vom eingesetzten Mausstamm, dem Injektionsprotokoll und der Dosis des verwendeten Mutagens. So konnte in zahlreichen Studien die Korrelation zwischen der eingesetzten Dosis und der Mutationsfrequenz bei einmaliger intraperitonealer Injektion von 25 bis 250 mg/kg Körpergewicht gezeigt werden [160]. Es hat sich ebenfalls gezeigt, dass die Mutationsfrequenz durch die wöchentliche Injektion kleinerer Dosen gesteigert werden kann [74].

Die Applikation von ENU bedingt eine Sterilitätsphase, in der die mutagenisierten Spermatozoen ihre

Teilungsaktivität vorübergehend einstellen. Die Dauer dieser Phase ist von Individuum zu Individuum unterschiedlich und schwankt zwischen verschiedenen Mausstämmen, vermutlich aufgrund von Polymorphismen im DNA-Reparaturmechanismus. Für die ENU Mutagenese haben sich die Mausstämme C57BL/6J, C3HeB/FeJ und BALB/cR1 als am besten geeignet herausgestellt [88]. Im MEP hat sich eine wöchentliche ENU-Dosis von 90 mg/kg über drei Wochen in Männchen des Mausstammes C3HeB/FeJ als Strategie bewährt, um erfolgreich Mutanten mit den angewendeten Protokollen zu finden [180]. Die so behandelten Tiere werden nach einer Sterilitätsphase von 100 Tagen mit Weibchen des selben Stammes verpaart. Die Nachkommen dieser Verpaarungen werden nach einem von Fuchs *et al.* beschriebenen Protokoll auf verschiedene Parameter, z. B. Auffälligkeiten des Körpers, des allgemeinen physischen Erscheinungsbildes, des Verhaltens und der Interaktion mit der Umwelt untersucht [57]. Im Alter von drei Monaten wird den Mäusen Blut entnommen und auf verschiedene klinisch-chemische Parameter untersucht, wie von Rathkolb *et al.* beschrieben [150]. Auch ein Allergietest wird durchgeführt.

Kreuzungsschema zur genomweiten Identifizierung dominanter Mausmutanten Tiere mit Auffälligkeiten in diesen Screens werden nochmals mit Tieren des C3HeB/FeJ Stammes verpaart um zu überprüfen, ob der beobachtete Phänotyp auf eine reproduzierbare genetische Ursache zurückzuführen ist (vgl. Abbildung 1.3 a). Im Unterschied zu anderen Screens, in denen ENU-behandelte Männchen direkt mit Weibchen eines anderen Stammes verpaart werden [135], findet auf diese Weise eine „Fixierung“ der Mutation durch kontinuierliche Zucht auf den C3HeB/FeJ Hintergrund statt. Dies hat den Vorteil, dass ENU-induzierte Mutationen nicht durch das Einbringen eines veränderten genetischen Hintergrundes beeinflusst werden und die Phänotypisierung vereinfacht wird. Ein Nachteil dieser Fixierung ist, dass gelegentlich ein stammspezifischer Phänotyp erzeugt wird, der während des Auskreuzens auf einem anderen Mausstamm nicht reproduziert werden kann und erst durch die Rückkreuzung auf den C3HeB/FeJ Hintergrund mit teilweise geringer Penetranz wieder aufgefunden wird.

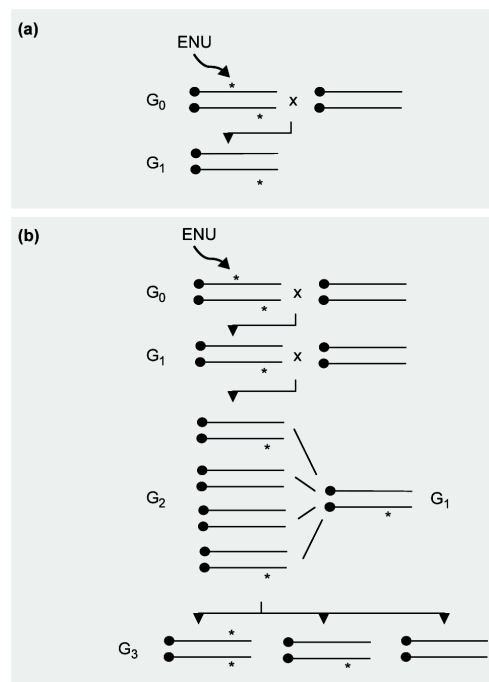


Abbildung 1.3: Zuchtschemata für die ENU Mutagenese nach Brown und Balling (modifiziert) [25]. a: Zuchtschema für genomweite dominante Mutationen, b: Zuchtschema für genomweite rezessive Mutationen

Kreuzungsschema zur genomweiten Identifizierung rezessiver Mausmutanten Ebenso hat sich das Screening auf rezessive Mutationen bewährt. Zunächst werden ENU-mutagenisierte Männchen mit weiblichen C3HeB/FeJ Mäusen verpaart. Männliche Tiere der F₁ Generation, die im ersten Screening keinen Phänotyp zeigen, werden nochmals mit C3HeB/FeJ Weibchen gekreuzt und die Geschwister der resultierenden N₂ Generation entweder wieder mit dem F₁ Männchen oder miteinander verpaart [77]. Die Nachkommen dieser Verpaarungen durchlaufen wiederum die oben beschriebenen Screening- und Bestätigungsprozeduren. Wenn eine ENU-induzierte Mutation vorliegen sollte, wird bei einem Teil der Tiere ein Phänotyp erkennbar sein (vgl. Abbildung 1.3 b).

Kartierung der Mutation Zur Kartierung der Mutation erfolgt gewöhnlich eine Auskreuzung auf den C57BL/6J Mausstamm. Anschließend wird, je nach Erbgang der Mutation, eine Rückkreuzung beziehungsweise eine Geschwisterverpaarung der F₁ Nachkommen vorgenommen. Die Nachkommen dieser Kreuzungen (entweder N₂ oder F₂ genannt) werden wiederum phänotypisiert und Mausschwanzbiopsien dieser Tiere gesammelt. Wenn auf diese Weise eine ausreichende Anzahl an Biopsien gesammelt wurde, wird daraus die DNA extrahiert und eine genomweite Kopplungsanalyse durchgeführt. Ursprünglich wurde die Kopplungsanalyse mit einigen Mikrosatellitenmarkern aus der genetischen Karte des Mausgenoms durchgeführt [39]. Die in dieser Arbeit vorgestellte genomweite Genotypisierung von SNPs im Hochdurchsatz hat diese Methode jedoch abgelöst.

Beispiele für eine erfolgreiche Genotyp-Phänotyp-Korrelation Seit dem Start des MEP sind weit über einhundert Publikationen entweder aus dem Projekt selbst oder über Kooperationen in Fachzeitschriften entstanden. Auf diese Weise konnte gezeigt werden, dass der Ansatz des MEP kontinuierlich neue, von der wissenschaftlichen Gemeinschaft anerkannte Ergebnisse produziert.

Ein sehr bekanntes Beispiel für eine aus dem MEP hervorgegangene Korrelation zwischen einer menschlichen Erkrankung und einer ENU-induzierten semidominanten Mutation ist die Beethoven-Maus (*Bth*), ein Modell für fortschreitenden Hörverlust ausgelöst durch eine T→A Mutation in Exon 13 des Gens *Tmc1* [202].

1.4. Kopplungsanalyse bei Mäusen

Seit der Wiederentdeckung der Arbeiten von Gregor Mendel durch Hugo de Vries und Carl Correns zu Beginn des 20. Jahrhunderts wurde verstanden, dass die Eigenschaften eines Individuums gekoppelt vererbt werden können. Mit der Entdeckung der Struktur und Funktion der DNA wurde auch erkannt, dass die gekoppelte Vererbung zweier oder mehrerer Gene abhängig ist von der räumlichen Nähe ihrer Loci auf einem Chromosom.

Die Verfahren zum Nachweis einer gekoppelten Vererbung zweier Loci auf einem Chromosom haben sich in den letzten Jahren und Jahrzehnten drastisch verändert. Die genomweite Genotypisierung von SNP-Markern im Hochdurchsatz ist der derzeit neueste Stand der Technik und wurde in dieser Arbeit zur Kopplungsanalyse von ENU-induzierten Mausmutanten eingesetzt.

1.4.1. Inzuchtstämme

Mäuse werden schon seit einigen hundert Jahren hauptsächlich nach ästhetischen Gesichtspunkten gezüchtet. Mit der Erkenntnis, dass die Zucht von Mäusen auch für die Wissenschaft nützlich sein kann, wurde die Inzucht von Mausstämmen vorangetrieben, um eine Reihe von genetisch identischen Individuen zu erhalten. Das spontane Auftreten phänotypischer Veränderungen in einzelnen Tieren einer Inzucht-kolonie ermöglichte Forschern die Studie erblich bedingter Erkrankungen am Mausmodell. Teilweise konnte die Kopplung einer neuen Spontanmutation an bereits bekannte Eigenschaften einer Mauslinie bewiesen und weiter erforscht werden.

1.4.2. Rekombinante Inzuchtstämme

Ein Rekombinanter Inzucht- (RI) Stamm wird aus der Kreuzung zweier Inzuchtstämme, der Geschwisterverpaarung der F₁ Generation und der Separation und Kreuzung von Geschwisterpärchen der F₂ Generation gebildet. Auf diese Weise werden die Rekombinationen der F₂ Generation von Generation zu Generation vererbt, bis schließlich (ca. ab Generation F₁₀) der genetische Hintergrund der parentalen Stämme in einem neuen Inzuchtstamm fixiert ist.

Phänotypische Unterschiede der parentalen Stämme können auf diese Weise durch die Genotypisierung des RI-Panels auf Übereinstimmung oder Nichtübereinstimmung getestet werden. Wenn beispielsweise ein BXD (C57BL/6J×DBA) Panel auf eine Eigenschaft des C57BL/6J Stammes getestet wird, kann erwartet werden, dass ein überproportionaler Anteil an RI-Stämmen des BXD-Panels mit derselben Eigenschaft in Übereinstimmung mit dem C57BL/6J an dem fraglichen Locus sein wird.

Ein Vorteil dieser Strategie ist, dass durch die Fixierung der parentalen Allele nach und nach durch Genotypisierung ein großer Vorrat an Genotypdaten akkumuliert wird, der bei Bedarf zur Untersuchung einer neuen Eigenschaft abgerufen werden kann und eine Kartierung erheblich beschleunigt.

Ein laufendes Projekt zur Generierung eines komplexen RI Panels ist der „Collaborative Cross“ [30], in dem insgesamt acht verschiedene Mausstämme als Basis für ein Verpaarungsschema dienen, aus dem insgesamt 1000 neue RI-Stämme generiert werden. Der Vorteil eines solch komplexen Schemas ist, eine genetische Diversität ähnlich der in der humanen Population zu erzeugen, allerdings mit genauer Kenntnis des genetischen Hintergründe und unendlicher Verfügbarkeit der einzelnen Stämme.

1.4.3. QTL Analyse

Die Kartierung von spontanen oder induzierten Mutationen in der Maus hat sehr viel zur Aufdeckung und molekularen Charakterisierung monogener Erkrankungen beigetragen. Jedoch wurde hierbei immer wieder festgestellt, dass einige Erkrankungen mit offensichtlich genetischer Ursache in ihren Ausprägungen variieren können und zu uneindeutigen Ergebnissen bei einer genomweiten Kartierung führen. Ursache hierfür sind komplexe Gen-Gen- oder Gen-Umwelt-Interaktionen. Das heißt, die Kombination verschiedener trans- oder cis-interagierender Gene, eventuell gepaart mit Umwelteinflüssen, kann dafür verantwortlich sein, dass ein Phänotyp in unterschiedlicher, quantifizierbarer Ausprägung beobachtet

wird. Beispiele solcher Interaktionen sind die genetische Komponente der Adipositas, Alkohol- und Drogenabhängigkeit oder Diabetes Typ II.

Die Assoziation von Loci mit quantifizierbaren Eigenschaften, auch als Quantitative Trait Loci (QTL) bezeichnet, ist schwierig und bedarf eines gewissen statistischen Aufwands. In der Vergangenheit wurden einige statistische Verfahren entwickelt, um solche Assoziationen zu beweisen [167, 181]. Mit der Entwicklung von genomweiten Kopplungskarten (z. B. unter Benutzung von Restriction Fragment Length Polymorphisms (RFLPs) in Kreuzungen zwischen *Mus musculus spretus* und C57BL/6 oder BALB/c [18, 9]) war es erstmals möglich, ursprünglich in der Humangenetik eingesetzte statistische Verfahren auf die Kopplungsanalyse in der Maus übertragen [100].

Zusammen mit erweiterten RI Panels [145] und sogenannten Advanced Intercross Lines (AIL) [36] konnte die Auflösung der Kartierung von QTLs verbessert werden. Die Identifikation der ursächlichen Mutation blieb bisher jedoch bis auf wenige Ausnahmen erfolglos. Bis heute gibt es ca. 20 Gene, die als ursächlich innerhalb eines QTL identifiziert werden konnten [51]. Mit der Erstellung von dichten SNP Karten wurde es möglich, anhand einer Haplotypanalyse *in silico* einen gefundenen QTL weiter einzuschränken [40], beziehungsweise einen QTL aufgrund von stammspezifischen Genotyp- und Phänotypdaten möglicherweise vorherzusagen [66]. Der Erfolg eines solchen Ansatzes, zusätzlich gekoppelt mit der Analyse von gewebsspezifischen Expressionsdaten, beispielsweise aus der SymAtlas Expressionsdatenbank [187], stellt ein mächtiges Werkzeug zur Identifikation des genomischen Elementes dar, das einem QTL zugrunde liegt [40].

1.4.4. Marker zur Genomkartierung

Vor der Entdeckung der DNA-Struktur und der Sequenzanalyse wurden verschiedene sichtbare oder biochemisch messbare Parameter, z. B. die Fellfarbe, sowie chromosomale Anomalien zur Kopplungsanalyse genutzt. Diese Methode war sehr hilfreich bei der Erstellung der ersten genetischen Karten des Mausgenoms [102]. Neue Phänotypen konnten so auf eine Reihe von verschiedenen Mausstämmen ausgekreuzt werden und eine Kopplung mit bekannten Markern auf diese Weise festgestellt werden. Die Entdeckung und Nutzung von molekularbiologischen Markern wie RFLPs zur Kartierung eines Genoms hat die Kopplungsanalyse stark vereinfacht [20]. Mit der Einführung der Polymerase Chain Reaction (PCR) Methode [127] konnten diverse neue Längenpolymorphismen auf dem Genom identifiziert werden, was den Arbeitsaufwand für eine Kartierung weiter reduzierte. Variable Number Tandem Repeats (VNTRs) wie Minisatelliten [85] und Mikrosatelliten [205] haben sich als genetische Marker wegen der leichten Amplifizierbarkeit per PCR und Auswertbarkeit mittels Elektrophorese auf einem Agarose- oder Polyacrylamidgel bewährt.

Die Nutzung von SNPs als Marker erfolgte anfangs durch Sequenzierung und polymorphismusspezifische PCR [133] oder durch das spezifische Schneiden eines PCR Produktes durch Restriktionsenzyme. Auch die Analyse von Single Strand Conformation Polymorphisms (SSCP) beruht auf Konformationsänderungen einzelsträngiger DNA durch Unterschiede einzelner Nukleotide in der untersuchten Sequenz.

1.4.5. Sequenzierung des Mausgenoms und die großen SNP-Discovery Projekte

Mit der Veröffentlichung der ersten physikalischen Karten der Genome des Menschen [101, 200] und der Maus [64, 204, 128] wurde es möglich, jede Nukleotidsequenz dieser Organismen *in silico* einer physikalischen Position auf dem Genom zuzuordnen und direkte Homologievergleiche anzustellen. Parallel zu diesen Arbeiten lief die Identifikation von SNPs in verschiedenen Mausstämmen [107, 207], mit deren Hilfe es möglich war, beispielsweise den Stammbaum einiger gängiger Mausstämmen nachzuzeichnen [147] und die Struktur des Mausgenoms aufzuklären [203]. Es stellte sich hierbei heraus, dass alle heutzutage verfügbaren Mausstämme hauptsächlich von den Species *Mus musculus musculus* und *M. m. domesticus* und zu einem geringen Teil auch von *M. m. castaneus* abstammen [203].

Eine Folge dieser Beobachtungen ist, dass auf dem Mausgenom große Bereiche von niedriger Diversität existieren. Es wird geschätzt, dass jede mögliche Kreuzung in 30-60% des Genoms einen gemeinsamen Haplotyp besitzt [207].

1.4.6. Informationsgehalt von SNPs und Mikrosatelliten

VNTRs, beispielsweise Minisatelliten, sind häufig hochpolymorph, es können also mehrere längenspezifische Allele pro Marker innerhalb einer Population gefunden werden. Diese Eigenschaft wird sich zum Beispiel in der Forensik zunutze gemacht, da es möglich ist, mit wenigen hypervariablen Minisatellitenmarkern humane DNA Proben mit an Sicherheit grenzender Wahrscheinlichkeit einem Individuum zuzuordnen [85]. Die Genotypisierung kurzer locuspezifischer repetitiver Elemente, auch Short Tandem Repeats oder Mikrosatelliten [205] genannt, hat die Genotypisierung von Minisatelliten wegen der leichteren Handhabung und der besseren Automatisierbarkeit [61] in der Forensik weitgehend abgelöst. Auch bei der Kopplungsanalyse werden in der Regel Mikrosatelliten eingesetzt [139].

Demgegenüber steht die in den meisten Fällen biallelische Natur von Single Nucleotide Polymorphisms (SNPs), was ein Nachteil bei der Nutzung als Marker darstellen würde. So wäre es rein rechnerisch erforderlich, 4 verschiedene SNPs zu genotypisieren (also 2^4 unterschiedliche Kombinationen zu erzeugen), um einen Mikrosatelliten mit 16 Allelen zu ersetzen.

Bei der Erstellung eines Markersatzes, mit dem es möglich sein soll, Kreuzungen verschiedener Mausstämmen zu genotypisieren, erscheinen SNPs wegen der enorm vergrößerten Anzahl an zu untersuchenden Markern für den selben Informationsgehalt zunächst den Mikrosatelliten unterlegen. Desweiteren bleibt nicht aus, dass ein Teil der SNP-Marker, abhängig von der Verwandtschaft der untersuchten Mausstämmen, nicht informativ ist, wodurch ein Teil der erhobenen Daten überflüssig wird.

Im Kontext von routinemäßiger Kreuzung und Genotypisierung zweier Mausstämmen ergibt sich jedoch kein Nachteil durch die reduzierte Polymorphie von SNPs, da konstant nur zwischen zwei Allelen unterschieden wird. Daher ist der Informationsgehalt in einem solchen Experiment mit dem von Mikrosatelliten gleichzusetzen und es werden keine überflüssigen Daten erhoben.

SNP Markersätze können also ohne Einschränkung als Ersatz für die Genotypisierung mit Mikrosatelliten im MEP eingesetzt und die Kopplungsanalyse von ENU-induzierten Mauslinien durch die Hochdurchsatzfähigkeit der SNP Genotypisierung erheblich beschleunigt werden.

1.5. Theorie zur genomweiten Kopplungsanalyse

Ein effizientes genomweites Genotypisierungsprojekt startet mit der Festlegung der Anzahl der zu genotypisierenden Tiere und der Anzahl der hierfür benötigten Marker, deren genomische Position bekannt ist (Ankermarker), um eine statistisch gesicherte Kopplungsanalyse zu betreiben. Man versucht also, durch die Untersuchung einer bestimmten Anzahl an Meiosen und Markern eine Kopplung des neu zu kartierenden Locus mit einem oder mehreren der Ankermarker mit hoher Wahrscheinlichkeit zu beweisen. Hierzu ist es notwendig, den Abdeckungsradius („Swept Radius“) eines Markers zu kennen (vgl. Abbildung 1.4).

Nach Carter und Falconer [27] wird dieser Abdeckungsradius definiert als die Länge eines chromosomalen Intervalls auf jeder Seite eines Markerlocus, innerhalb dessen Kopplung mit einer bestimmten Wahrscheinlichkeit detektiert werden kann. Die Wahrscheinlichkeit, mit der Kopplung zwischen zwei Markern bewiesen werden kann, ergibt sich aus Formel 1.1.

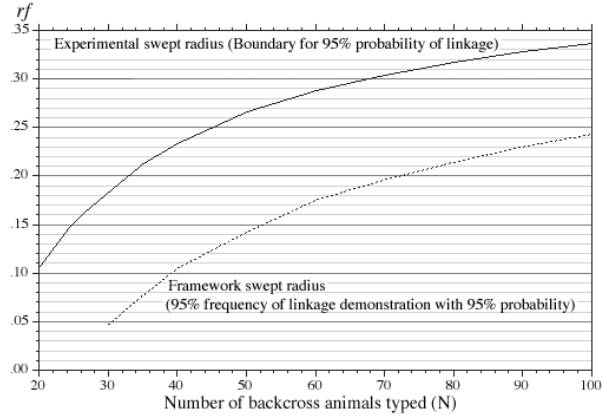


Abbildung 1.4: Darstellung des Framework Swept Radius [175].

Es sei

P : Wahrscheinlichkeit eines χ^2 Testergebnisses
 $swept$: hypothetischer Abdeckungsbereich eines Markers in cM
 $g = 1500$ cM genetische Gesamtlänge des Mausgenoms

Es gilt

$$f_{swept} = \frac{2swept}{g}$$

Wahrscheinlichkeit eines Nachweises von Kopplung

$$P(\text{Kopplung}) = 1 - \left(\frac{P}{P + f_{swept}} \right) \tag{1.1}$$

1.5.1. Nachweis von Kopplung in einem Satz von genomweit verteilten Markern

Durch Einsetzen einer Rekombinationsrate von 0,33 (entspricht einer genetischen Distanz von 39,8 cM nach Formel 1.2) für 90 Tiere aus einem Rückkreuzungsversuch (vgl. Abbildung 1.4) und der Ablehnung der Hypothese H_0 von Nichtkopplung zweier Loci mit einem Schwellwert $P \leq 0,05$ ergäbe sich nach den Einsetzen in Formel 1.1 eine Wahrscheinlichkeit von nur ca. 51%, dass eine beobachtete Kopplung bewiesen ist.

Man muss also einen höheren Schwellwert für P definieren, um eine beobachtete Kopplung auf einem akzeptablen Signifikanzniveau zu beweisen. Durch die Erhöhung des Schwellwertes zur Ablehnung der Hypothese H_0 auf $P \leq 0,001$ ergibt sich $P(\text{Kopplung}) = 1 - \left(\frac{0,001}{0,001+0,044}\right) \approx 0,98$. Falls H_0 eines χ^2 Kopplungstest mit einer Wahrscheinlichkeit von $P \leq 0,001$ abgelehnt wird, ist die Kopplung zweier Loci mit 98%-iger Wahrscheinlichkeit bewiesen.

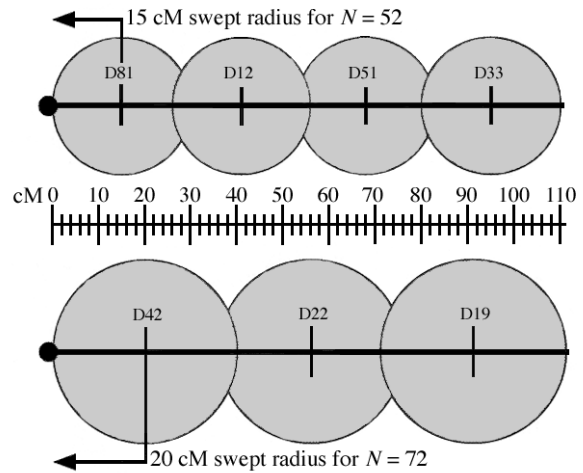


Abbildung 1.5: Anordnung der Marker auf einem Chromosom mit einem Framework Swept Radius (vgl. Abbildung 1.4) von 15 cM [175]

Es sei

- i : Anzahl der Rekombinanten
- N : Gesamtzahl Proben
- r : Rekombinationsfrequenz

Es gilt

$$r = \frac{i}{N}$$

Carter-Falconer Mappingfunktion [27]

$$m_{CF} = \frac{1}{4} \left\{ \frac{1}{2} [\ln(1+2r) - \ln(1-2r)] + \tan^{-1}(2r) \right\} \tag{1.2}$$

Für die Erstellung eines Markersatzes, der für eine Kartierung geeignet ist, müssen jedoch stringendere Werte angenommen werden, denn die tatsächliche Rekombinationsfrequenz zweier bekannter Loci kann von der angenommenen Frequenz stark abweichen [175, Kapitel 9.4]. Es wird daher der kleinste Wert

1. EINLEITUNG

des 95% Vertrauensintervalls der Wahrscheinlichkeitsdichte-Funktion 1.3 als Wert für den Grundabdeckungsradius („Framework Swept Radius“) genommen. Um die Lösung der Funktion 1.3 zu berechnen wird auf ein von Silver entwickeltes Programm zurückgegriffen [175, Appendix D].

Es sei

θ : angenommene Rekombinationsfrequenz

i : Anzahl der Rekombinanten

N : Gesamtzahl Proben

Wahrscheinlichkeitsdichte-Funktion

$$f(\theta) = \theta^i (1 - \theta)^{N-i}$$

Sonderfall $i = 0$

$$f(\theta) = \theta (1 - \theta)^N \tag{1.3}$$

Das heisst, in einem Satz von Markern müssen die Marker einen Abstand zueinander haben, der kleiner ist als das Doppelte des Grundabdeckungsradius (Siehe Abbildung 1.5). Dies würde nach Anwendung der Formel 1.3 mit einer Rekombinationsrate von 0,33 und einer Gesamtzahl von 90 Meiosen ($i = 30, N = 90$) einen Radius von 24,5 cM ergeben.

1.5.2. Sonderfall: Keine Rekombination zwischen Marker und dem zu kartierendem Locus

Aus Abbildung 1.6 und Formel 1.3 (wiederum berechnet mit dem in [175, Appendix D] aufgeführten Programm) lässt sich entnehmen, dass sich bei der Untersuchung von 100 Meiosen mit einer Wahrscheinlichkeit von 95% bei Abwesenheit von Rekombination von einem Abstand von maximal 3,6 cM zwischen einem Marker und dem zu testenden Locus ausgegangen werden kann.

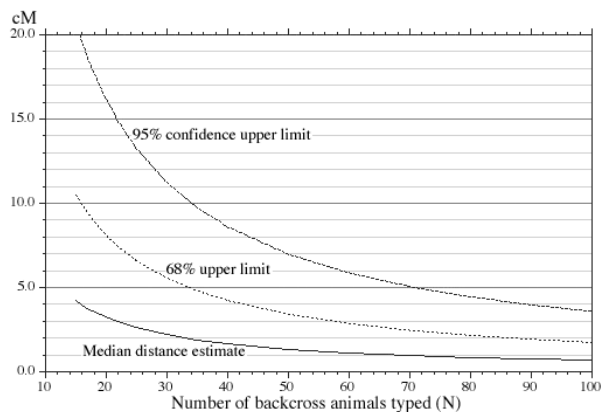


Abbildung 1.6: Der Abstand zweier Loci bei Abwesenheit von Rekombination in einem Backcross-Experiment [175].

1.6. SNP Genotypisierung im Hochdurchsatz

1.6.1. Überblick über molekularbiologische Methoden

Für die Genotypisierung von SNPs im Hochdurchsatz wurden mehrere Methoden und Plattformen entwickelt. Einen guten Überblick hierzu geben beispielsweise die Arbeiten von Gut und von Sobrino [68, 179]. Im Folgenden werden die am weitesten verbreiteten Methoden dargestellt.

Affymetrix, Perlegen Die von diesen Firmen angewandte Methode basiert auf direkter allelspezifischer Hybridisierung fluoreszenzmarkierter Sequenzen mit immobilisierten DNA-Sonden auf einem Microarray und der anschließenden Detektion mit einem Scanner.

Illumina Diese Methode basiert auf der allelspezifischen Extension von Primern mit einer Adresssequenz und der anschließenden Amplifizierung mit fluoreszenzmarkierten Primern. Das erhaltene Produkt wird nach Entfernung ungebundener Primer auf einem Chip hybridisiert, auf dem sich Partikel mit der passenden Adresssequenz befinden. Auch hier erfolgt die Detektion über einen Scanner.

TaqMan SNP Genotyping Ein weiterer Ansatz ist die SNP Genotypisierung mittels allelspezifischer Sonden, die ein Fluoreszenz- und ein Quencher-molekül besitzen. Durch die 5' Nukleaseaktivität der eingesetzten DNA-Polymerase wird das dem Allel entsprechende Fluoreszenzmolekül freigesetzt und kann durch einen Detektor erkannt werden.

SNPlex Eine SNP Genotypisierung kann ebenfalls durch Kapillargelelektrophorese mit einem Sequenziergerät erfolgen. Dieses Verfahren beruht auf der allel- und locus-spezifischen Ligation von Sonden mit einer bestimmten Adresssequenz an die genomische DNA. Nach der anschließenden Amplifikation dieser Sonden bindet ein fluoreszenzmarkierter Primer, der spezifisch für die Adresssequenz ist [195]. Nach der Entfernung von ungebundenen Primern können die Signale der restlichen Primer auf einem Sequenziergerät detektiert werden.

Sequenom Die vorgenannten Methoden benutzen Fluoreszenz zur Detektion allelspezifischer Produkte. Es ist jedoch auch möglich, SNPs ohne Fluoreszenz allein durch die Bestimmung der Masse von allelspezifischen Extensionsprodukten zu genotypisieren. Zur Auswertung der auf diese Weise erhaltenen Produkte wird die matrixunterstützte Laser Desorption/Ionisation (matrix assisted laser desorption ionization, MALDI) gemeinsam mit der Flugzeitmassenspektrometrie (time-of-flight mass spectrometry, TOF-MS) eingesetzt [190].

1.6.2. SNP Genotypisierung mittels MALDI-TOF MS

Allgemeine Funktionsweise In diesem Verfahren wird die zu untersuchende Probe zunächst mit einem 100- bis 100.000-fachen Überschuss einer Matrix versetzt und auf einem Proben-teller kristallisiert. Die Matrix besteht aus organischen Molekülen, die die Eigenschaft haben, Laserenergie einer bestimmten Wellenlänge stark zu absorbieren. Darüber hinaus verhindert die Matrix eine photolytische Beschädigung des Analyten sowie Interaktionen der Analytenmoleküle untereinander [73].

Eine so vorbereitete Probe wird einem intensiven Laserimpuls von wenigen Nanosekunden ausgesetzt, woraufhin Teilchen von der Oberfläche des Matrixkristallgitters explosionsartig verdampfen [89, 94].

Bei der Analyse von DNA Fragmenten kommen als Matrix hauptsächlich organische Säuren, z. B.

3-Hydroxypicolinsäure (3-hydroxypicolinic acid, 3-HPA) zum Einsatz [108, 68]. Der Ablauf des eigentlichen Desorptions-/Ionisationsvorgangs ist bis heute nicht vollständig aufgeklärt. Am wahrscheinlichsten ist, dass bei der Übertragung der Laserenergie auf die in der Matrix eingebetteten Probenmoleküle hauptsächlich einfach geladene Molekülionen entstehen, die in die Gasphase übertreten [73]. Diese Ionen werden in einem elektrischen Feld beschleunigt und durchlaufen anschließend eine feldfreie Driftstrecke, das sog. Flugrohr, wo sie abhängig von ihrer kinetischen Energie nach unterschiedlicher Flugzeit auf den Detektor treffen, es gilt also $E = \frac{1}{2}mv^2$. Die Flugzeit tof ist daher proportional zum Masse/Ladungsverhältnis, also $tof \propto \sqrt{\frac{m}{q}}$, wobei q ein ganzzahliges Vielfaches der Elementarladung ist [65]. In den meisten Fällen gilt $q = z = 1$, daher wird bei der Darstellung der erhaltenen Massenspektren der Wert m/z auf der x-Achse aufgetragen, während der Ausschlag auf der y-Achse proportional zur Anzahl der auf den Detektor treffenden Moleküle ist (vgl. Abbildung 1.7).

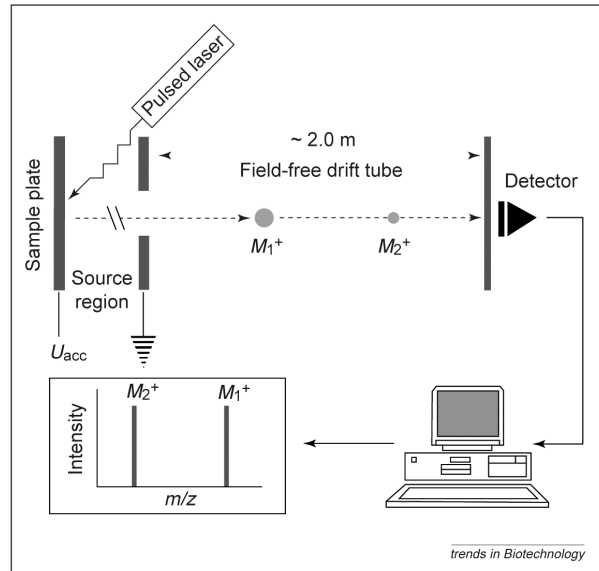


Abbildung 1.7: Schematischer Aufbau eines MALDI-TOF Massenspektrometers [65].

Genotypisierung von SNPs Mit den für die Genotypisierung eingesetzten MALDI-TOF Massenspektrometern ist es möglich, DNA-Fragmente mit einem Molekulargewicht von 1000 Da bis 9000 Da (ca. 3 bis 30 Nukleotide) mit einer Genauigkeit von 0,1% bis 0,01% zu detektieren [190]. Es ist daher möglich, die Extensionsprodukte einer Sequenzierung nach dem Kettenabbruchprinzip [165] durch den Massenunterschied der integrierten Nukleotide eindeutig einer Sequenz zuordnen zu können [185, 148, 103]. Diese Technik kann dazu benutzt werden, durch allelspezifische Primerextension SNPs zu genotypisieren, deren Genotyp und flankierende Sequenzen bekannt sind. Durch den Einsatz von Liquid Handling

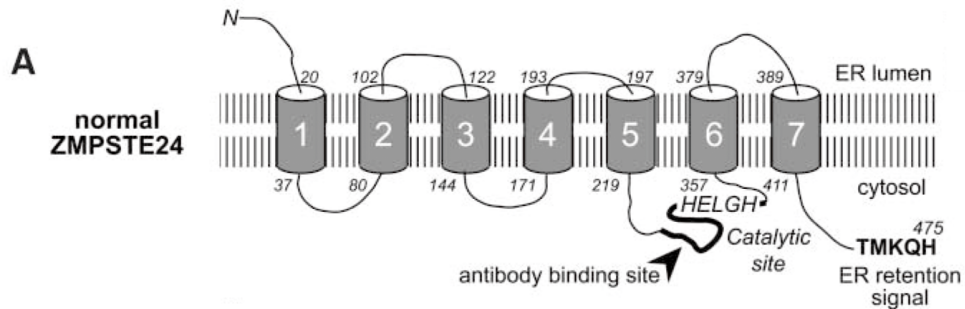


Abbildung 1.8: Struktur des humanen Zmpste24 [131].

Robotern und Chips, auf denen die Kokristallisation stattfindet, läßt sich die Methode desweiteren in hohem Maße automatisieren [104].

1.7. Der Phänotyp der Mauslinie SMA004

Im Zuge dieser Arbeit ist durch die genomweite Genotypisierung von SNP Markern mittels MALDI TOF-MS eine Kopplungsanalyse der Mauslinie SMA004 durchgeführt worden. Es stellte sich heraus, dass diese Mauslinie eine Mutation in dem Gen für Zink-Metalloprotease, STE24 homolog (Zmpste24) trägt. Die Mutation bewirkt einen Funktionsverlust dieser Protease, die maßgeblich an der Prozessierung von Prälamina A zu reifem Lamin A beteiligt ist. Lamin A bildet Filamente, die an der Aufrechterhaltung der Struktur und Funktion des Zellkerns beteiligt sind.

1.7.1. Struktur und Klassifizierung von Zmpste24

Die Zink-Metalloprotease Ste24p wurde zunächst in *Saccharomyces cerevisia* identifiziert. Hier ist sie an der Prozessierung des Pheromons **a**-Faktor beteiligt [58]. Das humane Ortholog wurde durch ein Screening einer humanen Ovar cDNA Bibliothek gefunden [56]. Es wurde zunächst als Farnesyliertes Protein-Converting Enzyme 1 (Face-1) bezeichnet. Dieser Name ist heute noch gebräuchlich. Aufgrund der Ähnlichkeit der proteolytischen Prozessierungsschritte des **a**-Faktors in der Hefe und von Prälamina A in Mammaliern lag die Vermutung nahe, dass die proteolytische Aktivität von Zmpste24 der von Ste24p ähnlich ist. Dies konnte anhand von Studien in Insektenzellen belegt werden [34].

Zmpste24 besteht aus 7 Transmembrandomänen und dem für Zink-Metalloproteasen typischen HEXXH-Motiv nach der fünften Transmembrandomäne (vgl. Abbildung 1.8). Es ist in die Membran des Endoplasmatischen Retikulums integriert und prozessiert dort das farnesylierte Prälamina A [169, 34].

Diese proteolytische Aktivität ist hochspezifisch für farnesyliertes Prälamina A und wird vermutlich über das in Mammaliern hochkonservierte RSYLLG Erkennungsmotiv vermittelt [92].

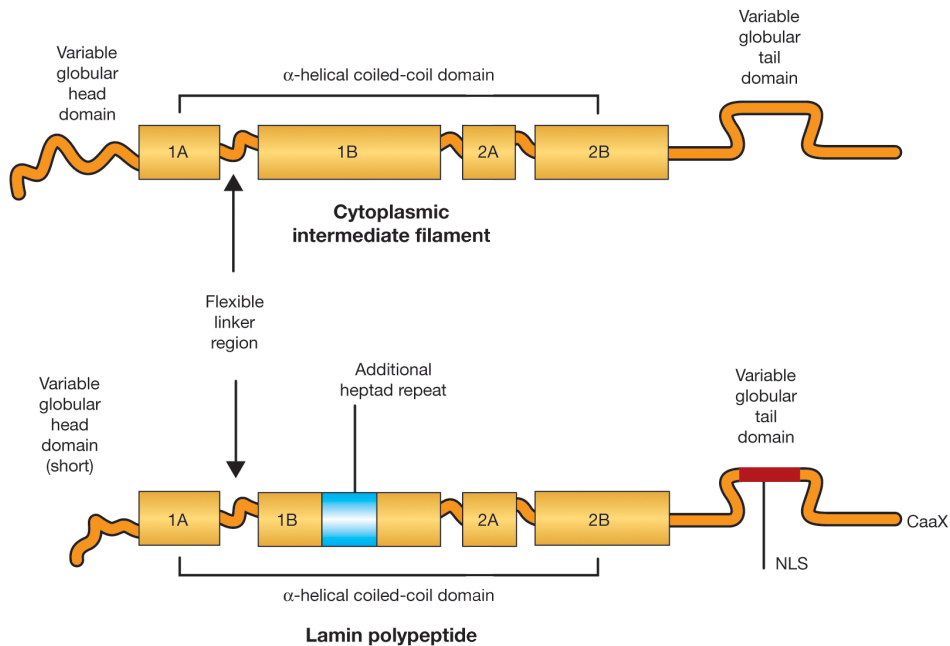


Abbildung 1.9: Allgemeine Struktur von Intermediärfilamentproteinen und Struktur der Lamine [80].

1.7.2. Struktur und Klassifizierung der Lamine

Lamine sind die strukturelle Hauptkomponente der Laminschicht des Zellkerns und werden bei den meisten Vertebraten von den Genen *Lmna*, *Lmnb1*, *Lmnb2* codiert. Splicevarianten des Gens *Lmna* sind Lamin A, AΔ10, C und C2. *Lmnb1* codiert für Lamin B1, und *Lmnb2* für Lamin B2 und B3.

Die Sequenzierung von cDNA Klonen aus verschiedenen Spezies [119, 50] und die Strukturanalyse ergab eine starke Homologie zu den Intermediärfilamenten. Sie wurden in die Klasse V der Intermediärfilamente eingeordnet. Die allgemeine Form dieser Filamente ist dreiteilig: Nach dem globulären Kopfbereich folgt die konservierte rod-Domäne, die aus drei α -helikalen Coiled-coil Bereichen besteht, welche in Heptadenwiederholungen organisiert sind. Der C-Terminus ist wiederum variabel und globulär. Lamine unterscheiden sich von diesem allgemeinen Schema in einigen Merkmalen. Innerhalb der rod Domäne befinden sich 42 zusätzliche Aminosäuren (sechs Heptaden), zusätzlich beherbergen sie ein Nukleuslokalisierungssignal [55] und die carboxyterminale CaaX Box, ein Signal für Isoprenylierung und Carboxymethylierung [75, 96] (vgl. Abbildung 1.9). Die rod Domänen zweier Lamin Monomere bilden eine doppelsträngige coiled-coil α -Helix. Diese Dimere sind die Grundbausteine für die Bildung der übergeordneten Struktur des Laminfilaments [1, 186].

1.7.3. Rolle der Lamine im Zellkern

Die Lamine bilden ein geordnetes filamentöses Netzwerk in den Zellkernen eukaryotischer Zellen und interagieren mit einer Vielzahl von membrangebundenen Proteinen und Transkriptionsfaktoren (TF) (vgl. Abbildung 1.10). Das Netzwerk ist allerdings nicht nur auf die innere Kernmembran beschränkt, sondern

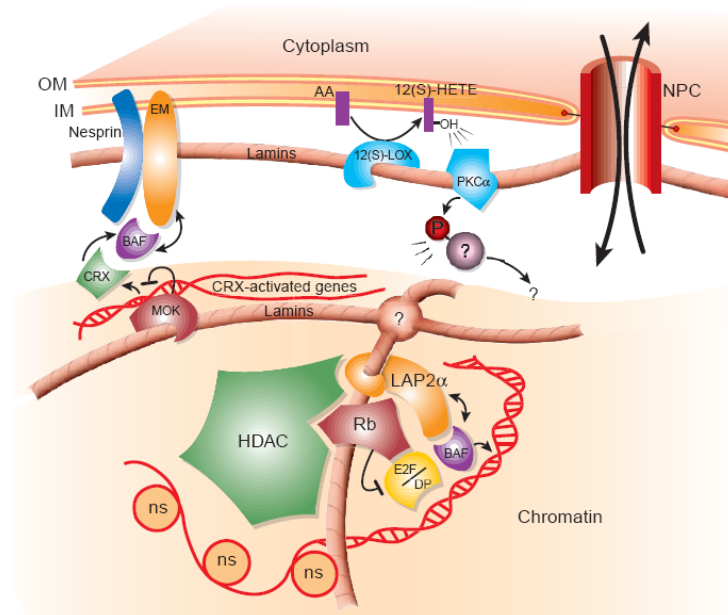


Abbildung 1.10: Mögliches Szenario der Interaktionen der Laminschicht im Nucleus [212].

durchzieht den gesamten Nucleus [76]. In Oozyten von *Xenopus laevis* konnte nachgewiesen werden, dass die räumliche Ausrichtung der Zellkernporenkomplexe durch die Laminfilamente erfolgt [1, 152]. Ebenso können Lamin A und C mit einer Vielzahl von Proteinen und Bestandteilen des Zellkerns interagieren und gemeinsam mit Nesprin und Emerin Kopplungskomplexe für das Cytoskelett bilden [121, 12]. Neben den Architekturproteinen Nesprin, Emerin, LAP1 und LAP2 α interagieren auch Histone [191], Transkriptionsfaktoren wie Mok2, Srebf1 a/c und Rb, Signalproteine wie E1B 19K, 12(s)-Lipoxygenase und PKC α , sowie Narf, ein Protein von unbekannter Funktion [212].

In der Nähe der mitotischen Zellteilungsspindel konnten Lamine vom Typ B nachgewiesen werden [197]. Lamin A und C sind an der intranukleären Bildung von Splicing-Komplexen [83] und der mRNA-Transkription durch RNA Polymerase II beteiligt [182].

1.7.4. Definition von Laminopathien

Unter Laminopathien sind gemeinhin Fehlfunktionen der nukleären Lamina zusammengefasst. Die X-koppelte Emery-Dreifuss muskuläre Dystrophie (EDMD1, OMIM #310300) wird ausgelöst durch Mutationen des nukleären Proteins Emerin [15]. Eine Duplikation des Gens für Lamin B1 führt zu einer Leukodystrophie (ADLD, OMIM #169500) und erworbene teilweise Lipodystrophie (APL, OMIM #608709) wird durch Mutationen in Lamin B2 hervorgerufen [142, 71]).

Momentan lassen sich zehn menschliche Krankheitsbilder auf Mutationen in Lamin A zurückführen. Die Phänotypen sind prominent in vielen verschiedenen Gewebe- und Zelltypen und lassen sich teilweise nicht deutlich voneinander abgrenzen (vgl. Tabelle 1.1).

Tabelle 1.1: Lamin A Mutationen erzeugen eine Vielzahl von unterschiedlichen Krankheitsbildern

Phänotyp	genaue Bezeichnung	Abkürzung	OMIM Nummer	Referenz
Muskeldystrophie	autosomal dominante Emery-Dreifuss Muskuläre Dystrophie	EDMD2	#181350	[19]
	autosomal rezessive Emery-Dreifuss Muskuläre Dystrophie	EDMD3	#604929	[10]
	Gliedergürtelmuskeldystrophie Typ 1B	LGMD1B	#159001	[125]
Kardiomyopathie	dilatative Kardiomyopathie	DCM	#115200	[49]
Lipodystrophie	familiäre partielle Lipodystrophie Typ 2	FPLD2	#151660	[172]
Neuropathie	Charcot-Marie-Tooth Syndrom Typ 2 B1	CMT2B1	#605588	[164]
Progeriesyndrom	Hutchinson-Gilford-Progeriesyndrom	HGPS	#176670	[45]
	atypisches Werner-Syndrom	WRN	#277700	[29]
Dysplasie	Mandibuloakrale Dysplasie	MAD	#248370	[136]
	Restriktive Dermatopathie	RD	#275210	[132]
nicht klassifiziert	hypertrophische Kardiomyopathie mit insulinresistenter Diabetes			[28]
	Arthropathie mit progeroidem Erscheinungsbild			[46]

OMIM: Online Mendelian Inheritance in Man

1.7.5. Expression und posttranslationale Modifikationen der Lamine

Lamin B2 wird in den meisten Zelltypen ubiquitär und unabhängig von der Zellentwicklung exprimiert, während Lamin B1 hauptsächlich in proliferierenden Zellen nachgewiesen wird. Lamin A und C wird hauptsächlich in ausdifferenzierten Zellen exprimiert [22], in *M. musculus* jedoch nicht in der Milz, dem Knochenmark und im Thymus [162]. A- und B-Typ Lamine haben unterschiedliche Funktionen in der Zelle. So sind an der Zellteilung hauptsächlich Lamine des Typs B beteiligt. Genomisch liegt Lamin B2 in direkter Nachbarschaft zu einer Stelle des DNA-Replikationsstarts und wird unmittelbar nach Einleitung der S-Phase der Mitose transkribiert [60]. Die Transkription und der Einbau der A- und C- Lamine startet im Gegensatz dazu während der G₁ Phase der Zellreplikation. Sie sind hauptsächlich für die mechanische Stabilität und korrekte Organisation des Nukleus verantwortlich [98, 80].

Laminen des Typs A und B ist das C-terminale CaaX Motiv gemeinsam, durch das die Prenylierung mit dem Lipid Farnesyldiphosphat veranlasst wird. Farnesylierte Proteine sind in die Plasmamembran oder in Mikrosomen integriert. Dort wird proteolytisch das aaX Tripeptid entfernt und das nun C-terminale Cystein carboxymethyliert [213]. Die letzten 15 Aminosäuren des farnesylierten, carboxymethylierten Carboxy-Terminus von Lamin A werden proteolytisch entfernt, um reifes, nicht farnesyliertes Lamin A zu erzeugen. Die Farnesylierung wird durch Farnesyldiphosphat Farnesyltransferase I (Fdft1) vermittelt und die Carboxymethylierung erfolgt über Isoprenylcysteinylmethyltransferase (Icmt) mit S-Adenosylmethionin als Methylendonator [2]. Bei der Prozessierung von Prälamina A zu reifem Lamin A wird sowohl der erste proteolytische Schritt als auch der zweite proteolytische Schritt von Zmpste24 durchgeführt [34] (vgl. Abbildung 1.11). Lamine des Typs C werden ohne weitere posttranslationale Modifikation in die nukleäre Lamina integriert. Sie sind eine Splicevariante von *Lmna*, in der nicht das CaaX Prenylierungsmotiv enthalten ist.

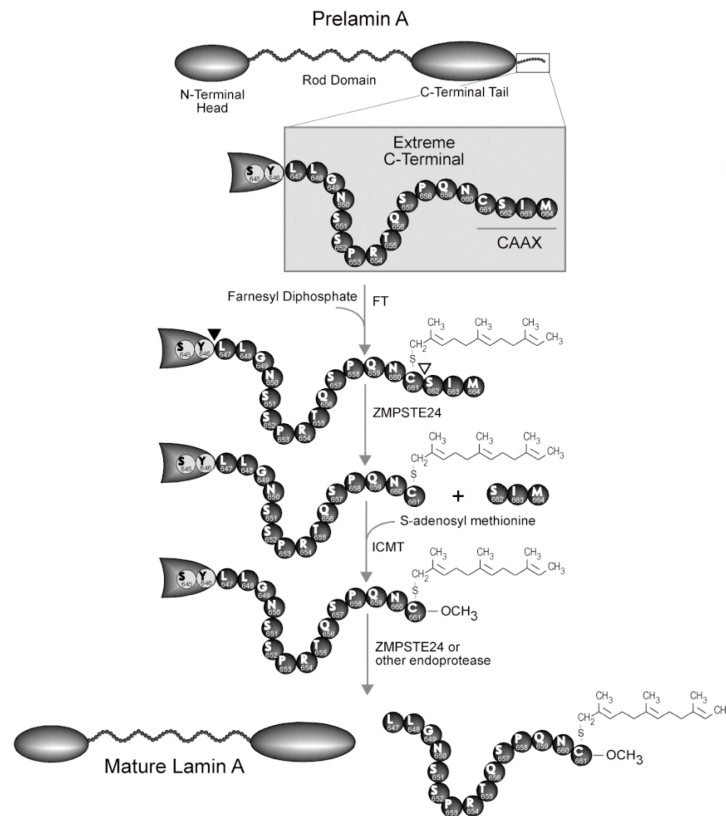


Abbildung 1.11: Schema der Prozessierung von Prälamina A zu reifem Lamin A [2].

1.7.6. Fehlprozessierung von Prälamina A

Mutationen in dem Gen *Lmna*, die zur Aktivierung einer kryptischen Splice-Stelle oder zu einer Störung der Erkennungsmotive für die Proteolyse durch *Zmpste24* führen, werden mit HGPS, MAD und Restriktiver Dermatopathie (RD) in Verbindung gebracht. Speziell im Fall von HGPS gibt es einige Hinweise, dass die Aktivierung einer kryptischen Splice-Stelle auf Exon 11 zu einer $\Delta 150$ *LMNA* mRNA und letztlich zu einem $\Delta 50$ Lamin A Produkt führt, das zwar das CaaX Motiv enthält, farnesyliert und teilweise carboxymethyliert werden kann, jedoch nicht proteolytisch zu reifem Lamin A prozessiert wird [211]. Dies führt zu einer Akkumulation von farnesyliertem Prälamina A in der Kernhülle und zu einer Störung des Aufbaus der Laminschicht [124]. Kürzlich wurde ebenfalls bekannt, dass auch in gesunden Zellen gelegentlich diese kryptische Splice-Stelle genutzt wird, was einen negativen Effekt auf die Struktur der Zellkerne hat [168]. Daraus folgt, dass ohne vollständige posttranslationale proteolytische Prozessierung durch *Zmpste24*, speziell ohne die Abspaltung von 15 Aminosäuren am farnesylierten C-terminalen Ende, ein progeroider Phänotyp ausgelöst wird. Das mutierte farnesylierte Prälamina A wird als Progerin bezeichnet.

Dieselben Effekte können beobachtet werden, wenn die posttranslationale Prozessierung durch Mutationen in *Zmpste24* verhindert wird. Auch in diesem Fall befindet sich Progerin im Zellkern, da nach der Farnesylierung keine weitere Prozessierung durch *Zmpste24* erfolgen kann (vgl. Abbildung 1.11).

1.7.7. Mausmodelle zur Untersuchung von Lamin A und Progerin

Ein Knockout von *Lmna* bewirkt einen Phänotyp, der an Emery-Dreifuss Muskuläre Dystrophie erinnert [188]. Es zeigte sich, dass die Zellkerne von *Lmna*^{-/-} murinen embryonischen Fibroblasten (MEFs) sehr unregelmäßig waren und die gleichmäßige Verteilung von LAP2 α und den Nukleären Porenkomplexen nicht aufrechterhalten werden konnte. Desweiteren war Emerin nicht mehr in der inneren Kernmembran gebunden, sondern über das gesamte Cytosol verteilt.

Ein Knockout von *Zmpste24* wurde sowohl von Pendás als auch von Bergo vorgestellt [146, 13]. Beide *Zmpste24*^{-/-} Mäuse zeigten einen schweren Phänotyp mit unregelmäßig geformten Zellkernen, multiplen Knochenbrüchen, Kardiomyopathie und Lipodystrophie. Dieser Phänotyp überschneidet sich stark mit dem der *Lmna* Knockout Maus.

Die Phänotypisierung von Doppelmutanten (*Zmpste24*^{-/-}*Lmna*^{+/-}) ergab, dass eine um die Hälfte reduzierte Dosis von farnesyliertem Prälamina A ausreicht, um den Phänotyp zu eliminieren [54, 199]. Dies führte zu dem Schluß, dass das farnesylierte Prälamina A der eigentliche Auslöser der progerieartigen Symptome ist. Diese Hypothese wurde erhärtet durch die Generierung von Knockout-Mäusen, die nur Lamin C exprimieren (*Lmna*^{LCO/LCO}). Diese Mäuse zeigen keinen sichtbaren Phänotyp und die Form der Zellkerne erscheint normal. Ebenso reduziert sich die Anzahl der deformierten Zellkerne in Zellen von *Zmpste24*^{-/-}*Lmna*^{LCO/+} und *Zmpste24*^{-/-}*Lmna*^{LCO/LCO} Mäusen [53].

1.7.8. Molekulare Ursache

Auswirkungen von Lamin A und Zmpste24 Defekten auf die Zellkernstruktur Mehrfach konnte gezeigt werden, dass alle Arten von Mutationen des *Lmna* und des *Zmpste24* Gens einen instabilen und unregelmäßig geformten Zellkern zur Folge haben [188, 45, 32, 124]. Es gibt mannigfaltige Beweise dafür, dass der progeroide Phänotyp hauptsächlich durch die Instabilität und Verformung des Zellkerns verursacht wird. Von dieser Instabilität werden verschiedene Effekte, wie die erhöhte Anfälligkeit für mechanischen Stress [99], die gestörte Lokalisation der nukleären Poren [62] oder eine Fehlverteilung von peripherem Heterochromatin abgeleitet [32, 38]. *Zmpste24*^{-/-} MEFs zeigen außerdem erhöhte Beschädigungen der DNA und chromosomale Anomalien [109].

Störungen der Transkription durch Lamin A Defekte Lamin A interagiert mit einer Vielzahl von nukleären Bestandteilen wie Transkriptionsfaktoren, strukturellen Komponenten, Histonen, DNA und Polymerasen. Vielfach wurde bereits die Wichtigkeit von Lamin A für die Verteilung von Hetero- und Euchromatin beschrieben und dessen Beteiligung an Transkriptions- und Splicing-Prozessen [83, 23]. Es liegt daher nahe anzunehmen, dass Lamin A eine wichtige Rolle in der zeitlichen und räumlichen Organisation von Transkriptionsprozessen spielt. Insbesondere die Interaktion zwischen Lamin A und dem Retinoblastomprotein (pRB) ist hierfür als ein Beweis anzusehen [141, 86]. Das Retinoblastom Protein spielt eine wichtige Rolle in der Kontrolle des Zellzyklus, indem es den Faktor E2F, der die S-Phase des Zellzyklus einleitet, reprimiert. Während der G₁ Phase des Zellzyklus ist es an die Lamina gebunden, wo es vor der Degradation durch das Proteasom geschützt ist [80]. Es interagiert über LAP2 α

mit Lamin A [115]. Johnson *et al.* mutmaßen, dass Lamin A für die korrekte Platzierung von pRB/E2F Komplexen notwendig ist [86].

Desweiteren interagiert der Sterol regulatory element binding transcription factor 1 (Srebf1) mit Lamin A. Dieser Transkriptionsfaktor spielt eine wichtige Rolle in der Cholesterinbiosynthese sowie in der Differenzierung von Adipozyten [93] und eine Fehlregulation durch eine gestörte Interaktion zwischen Lamin A und Srebf1 könnte eine der Ursachen für die Lipodystrophie von HGPS und MAD Patienten sein [110].

1.7.9. Therapieansätze

Mehrfach konnte gezeigt werden, dass farnesyliertes Prälamina A der eigentliche Auslöser der progerieartigen Symptome in HGPS und MAD Patienten ist. Es wurde daher versucht, mit Inhibitoren die Farnesylierung von Prälamina A zu unterbinden.

Mehrere Studien belegen, dass das nicht farnesylierte Prälamina A anstatt in der Kernmembran im Nukleoplasma akkumuliert und sich *in vitro* die Form der Zellkerne deutlich verbessert [209, 196, 113]. Kürzlich konnte ebenfalls gezeigt werden, dass der progeroide Phänotyp von Zmpste24-defizienten Mäusen *in vivo* deutlich durch die Gabe des Farnesyltransferase-Inhibitors ABT-100 gemildert wird [52]. Parameter wie die Gewichtszunahme, die Sterblichkeit und die Griffstärke von so behandelten Mäusen konnten gegenüber der Zmpste24-defizienten Kontrollgruppe verbessert werden, erreichten allerdings nicht die Parameter der unbehandelten bzw. mit ABT-100 behandelten Wildtypgruppe, was auf die eventuell zu niedrige Dosierung von ABT-100 zurückgeführt wurde.

1.8. Ziele der Arbeit

Bis heute wurden im Rahmen des MEP über 750 Mauslinien etabliert, die veränderte morphologische, immunologische, klinisch-chemische Parameter oder sonstige Auffälligkeiten aufweisen. In diesem Zeitraum konnte über die Genotypisierung von Mikrosatelliten nur bei ca. 100 Mauslinien eine Kopplungsanalyse durchgeführt werden.

Diese Methode erwies sich daher als zu langsam und zu arbeitsaufwendig, um mit der stetig wachsenden Zahl von Mauslinien Schritt zu halten. Zur Beseitigung dieses Ungleichgewichts sollte ein System etabliert werden, mit dem in kurzer Zeit eine genomweite Kopplungsanalyse durchgeführt werden konnte. Trotz der Einfachheit in ihrer Handhabung ist die Genotypisierung von Mikrosatelliten im Hochdurchsatz zwar prinzipiell möglich [81], aber durch den massiven Preisverfall bei der Genotypisierung von SNPs in der routinemäßigen Genotypisierung von Kreuzungen weniger Mausstämmen nicht mehr rentabel.

Ein weiterer Vorteil der SNP Genotypisierung ist die hohe Verfügbarkeit von genotypisierbaren Loci. Die derzeit größte öffentlich zugängliche Datenbank für Maus-SNPs, die dbSNP, hat momentan Einträge an knapp 5 Millionen Loci des Mausgenoms. Bezogen auf eine Gesamtlänge des Mausgenoms von ca. 3 Milliarden bp [204] entspricht dies einem SNP pro 600 bp. In dieser Arbeit sollte (1) ein hochdurchsatz-

fähiges Genotypisierungs-System basierend auf SNPs entwickelt werden und **(2)** ein Markersatz etabliert werden, der bereits nach der genomweiten Genotypisierung die Eingrenzung auf ein möglichst kleines subchromosomales Intervall erlaubt.

(3) Geeignete Marker mussten gefunden und validiert werden. Die zu Beginn des Projekts vorhandenen öffentlichen und proprietären Suchmaschinen waren allerdings nur bedingt zur Erstellung von Markersätzen geeignet, daher musste eine neuartige Suchmaschine entwickelt werden, mit der geeignete SNP Marker für ein genomweites Genotypisierungsprojekt auf komfortable Weise gefunden werden konnten. **(4)** Die erhaltenen Genotypdaten mussten gesammelt, ihren parentalen Allelen und passenden Loci zugeordnet und für die Kopplungsanalyse aufbereitet werden. Zu diesem Zweck musste ein System geschaffen werden, das diese Anforderungen erfüllt.

Zur Unterstützung der routinemäßigen Genotypisierung sollte **(5)** ein automatisiertes DNA-Extraktionssystem etabliert werden. **(6)** Konzeptionsbeweise für die Effektivität der genomweiten Kopplungsanalyse sollten erbracht werden. Je eine Mauslinie mit einem dominant und einem rezessiv vererbten Phänotyp wurde hierfür weitergehend untersucht.

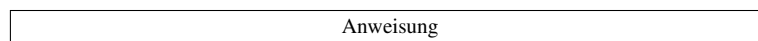
(7) Die Mauslinie SMA004 mit einem rezessiv vererbten Phänotyp sollte kartiert, und der zugrunde liegende Genotyp möglichst anhand der Daten aus der Kopplungsanalyse gefunden werden. Parallel zu diesen Arbeiten sollte der Phänotyp dieser Mauslinie eingehender untersucht werden. **(8)** Für die Auffindung einer dominanten Mutation aufgrund der Genotypdaten wurde die Mauslinie ALI029 ausgewählt.

2. Material und Methoden

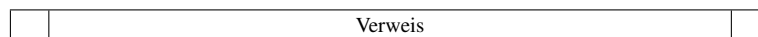
2.1. Struktogramme nach Nassi-Shneiderman

Im Folgenden werden die Symbole zu den in dieser Arbeit benutzten Nassi-Shneiderman Struktogrammen abgebildet [130]. Mit dieser Methode ist es möglich, einen zeilenbasierten Programmablauf in kompakter Form darzustellen.

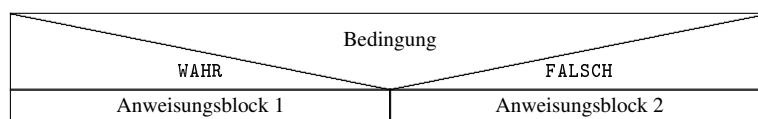
Linearer Ablauf



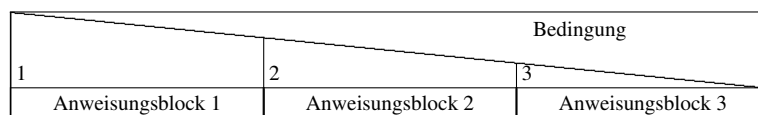
Verweis auf Unterverarbeitung oder externe Verarbeitung mit Rückkehr



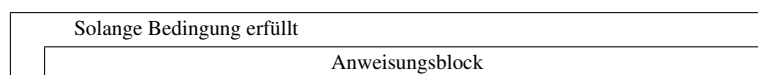
Einfache Verzweigung



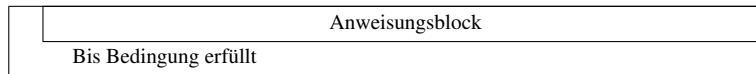
Mehrfache Verzweigung



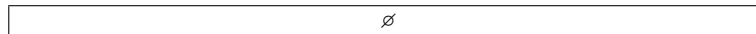
Wiederholung mit vorausgehender Bedingungsprüfung



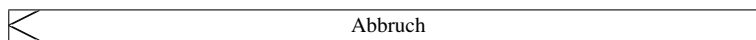
Wiederholung mit nachfolgender Bedingungsprüfung



Überspringen bis zur nächsten Anweisung



Abbruch des Programms / Sprung zur übergeordneten Schleife



2.2. Automatisierte DNA Extraktion

Für die vollautomatisierte DNA-Extraktion mittels Liquid Handling Roboter wurde das Gerät Freedom Evo 150 (Tecan) mit einem robotischen Arm (RoMa), einem 4-Kanal Liquid Handler (LiHa) mit einem Maximalvolumen von 1000 μl pro Spitze, einem Magnetseparator (TeMagS) und einem GENios Plattenabsorptionsmessgerät etabliert. Als Extraktionskit kam das AGOWAmag Maxi DNA-Kit zum Einsatz. Die endgültige Version des Skripts wurde so gestaltet, dass es dem Benutzer auf Wunsch möglich ist, zusätzlich zur DNA-Extraktion eine automatisierte Quantifizierung und Einstellung einer gegebenen Endkonzentration auszuführen.

2.2.1. Vorbereitung der Schwanzbiopsien

Proben von maximal 1 cm Länge werden in einer Deep Well Plate (DWP) mit einer Mischung aus 197 μl Cell Lysis Buffer und 3 μl Proteinase K (20 mg/ml) versetzt und über Nacht bei 57°C inkubiert. Die Proben werden anschließend für 10 min bei 2500 rpm in einer Eppendorf-Zentrifuge mit Ausschwingrotor zentrifugiert. Der Überstand wird in eine neue DWP überführt, mit 200 μl BLM Lysis Buffer (AGOWA) und 200 μl 100% Ethanol versetzt und gründlich gemischt.

2.2.2. Steuerung des Extraktionsvorgangs

Mit dem Programm Gemini (Tecan) wurde ein Skript erstellt, welches eine für das AGOWA DNA-Extraktionskit optimierte DNA Extraktion auf dem beschriebenen System gewährleistet. Mit Hilfe dieses Skripts können bis zu 96 Gewebeproben extrahiert werden. Es können maximal 48 Proben gleichzeitig mit dem TeMagS bearbeitet werden. Es sind also zwei Durchgänge notwendig, um eine komplett bestückte Mikrotiterplatte mit 96 Proben zu bearbeiten. Tabelle 2.1 enthält eine Auflistung aller zum Verständnis des Skripts relevanten Variablen. In Abbildung 2.1 ist das allgemeine Vorgehen in einem Flussdiagramm dargestellt.

Benutzereingaben Das Skript benötigt zunächst Eingaben zur Anzahl der zu extrahierenden Proben (Variable `Probenanzahl`) und dem Versatz (Variable `Kavitätenversatz`) der ersten Probe relativ zur ersten Position.

Anschließend kann der Benutzer durch Setzen der Variablen `Messen` und `Verdünnen` wählen, ob eine Messung der DNA Konzentration und Verdünnung der extrahierten Proben gewünscht ist. Falls eine Messung gewünscht wurde, wird die Kommunikation mit Pegasus hergestellt und der Benutzer durch das Pegasus-Skript `selectExcel.pts` zur Auswahl einer zum Gewebetyp passenden Excel-Vorlage aufgefordert, in der sämtliche Parameter zur Messung und Verdünnung enthalten sind.

Verarbeitung durch das Extraktionskript Eine allgemeine Übersicht über das Vorgehen wird in Abbildung 2.1 dargestellt. Zunächst wird die Separation vorbereitet, indem alle internen und externen Variablen gesammelt und initialisiert werden (vgl. Tabelle 2.1), die Arbeitsfläche durch den RoMA beladen wird (Abbildung 2.5) und die Lysate auf den Magnetseparator überführt werden (vgl. Abbildung 2.2).

Anschließend erfolgt die Magnetseparation und das Aufreinigen der an die paramagnetischen Partikel gebundenen DNA (vgl. Abbildung 2.3). Zunächst werden 300 μl der Waschpuffer BLM1 oder BLM2 zugegeben und der Mischvorgang gestartet. Ein weiteres Volumen von 420 μl BLM1/BLM2 wird während des Mischens zugegeben, um das von Hersteller vorgeschriebene Volumen von 420 μl zu erreichen. Diese Vorgehensweise führt zu einem deutlichen Geschwindigkeitsgewinn gegenüber der einmaligen Zugabe von 720 μl , da von der wesentlich schnelleren Mehrfach-Pipettierung des Liquid Handlers Gebrauch gemacht wird. Zusätzlich wird ein Mischeffekt während der Zugabe des zweiten Volumenanteils erreicht. Nach der zweiten Zugabe von BLM2 werden die Magnetpartikel zur Evaporation des Acetons aus dem Puffer BLM2 bei 60°C inkubiert.

Die getrockneten Magnetpartikel werden in 10 mM Tris pH=8,0 eluiert. Während der Elution wird die Arbeitsfläche durch den RoMA umkonfiguriert (vgl. Abbildung 2.6). Je nach Status der Variablen `Messen` und `Verdünnen` wird eine Eluat-, eine Meß- und eine Arbeitsplatte auf die Arbeitsfläche gestellt und unterschiedliche Volumina auf diese Platten überführt (vgl. Tabelle 2.2). Wenn mehr als 48 Proben separiert werden sollen, wird der zweite Durchgang eingeleitet (vgl. Abbildung 2.4).

Nach der Separation aller Proben wird, ebenfalls abhängig von `Messen` und `Verdünnen`, die Quantifikation und gegebenenfalls die Verdünnung eingeleitet und alle Utensilien von der Arbeitsfläche entfernt.

2. MATERIAL UND METHODEN

Tabelle 2.1: Übersicht über die wichtigsten Variablen des automatisierten DNA Extraktionsprotokolls

Name	Typ	Funktion	Kontrolle	Eigenschaften
Probenanzahl	Integer	Gesamtzahl der zu extrahierenden Proben	Benutzer, Skript	$\text{Probenanzahl} \leq 96$, während eines Durchgangs $\text{Probenanzahl} \leq 48$
Kavitätenversatz	Integer	Versatz der ersten Probe relativ zur ersten Plattenposition	Benutzer, Skript	$\text{Kavitätenversatz} \leq 96$, nach Durchgang 1 $\text{Kavitätenversatz} = 48$
Messen	Integer/ Boolean	Regelt, ob DNA quantifiziert werden soll	Benutzer	$\text{Messen} = 1$ wenn Quantifikation gewünscht
Verdünnen	Integer/ Boolean	Regelt, ob DNA verdünnt werden soll	Benutzer	$\text{Verdünnen} = 1$ wenn Verdünnung gewünscht, bedingt $\text{Messen} = 1$
Durchgang2	Integer/ Boolean	Gibt die Nummer des Durchgangs an	Skript	$\text{Durchgang2} = 0$, nach erstem Durchgang $\text{Durchgang2} = 1$
Probenanzahl1	Integer	Erhält Wert von Probenanzahl	Skript	Wenn $\text{Probenanzahl} > 48$, dann $\text{Probenanzahl1} = \text{Probenanzahl} - 48$
b1m2	Integer/ Boolean	Regelt die Wiederholung der BLM2 Zugabe	Skript	$\text{b1m2} = 0$, nach erstem Waschschrift $\text{b1m2} = 1$

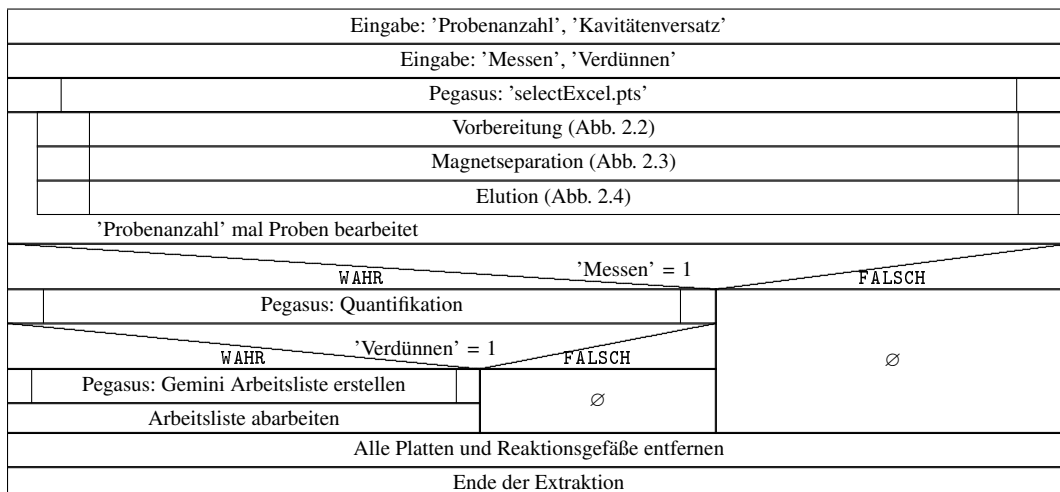


Abbildung 2.1: Allgemeines Ablaufschema der DNA Extraktion

Tabelle 2.2: Elutionsvolumina pro Platte in Abhängigkeit von dem Status der Variablen „messen“ und „verdünnen“.

	messen = 0 verdünnen = 0	messen = 1 verdünnen = 0	messen = 1 verdünnen = 1
Elutionsplatte	180 μl	160 μl	120 μl
Messplatte		20 μl	20 μl
Verdünnungsplatte			40 μl

WAHR		'Probenanzahl' > 48	FALSCH	
'Probenanzahl1' = 'Probenanzahl'			∅	
'Probenanzahl' = 48				
Positionieren der Lysatplatte und der Reaktionsgefäße				
WAHR		'Durchgang2' = 0	FALSCH	
Positionieren der DiTis				
Vormischen der Magnetpartikel und Verteilen von je 20 µl				
'Probenanzahl' mal Lysat auf Magnetseparator überführen				
Inkubation				
WAHR		'Probenanzahl1' < 49	FALSCH	
			Mischen der Magnetpartikel	
Separation				
Überstand abnehmen				

Abbildung 2.2: Ablaufschema Vorbereitung der Arbeitsplatte und Überführung der Lysate

'Probenanzahl' mal Zugabe von BLM1				
Mischen				
Separieren				
'Probenanzahl' mal Überstand abnehmen				
'blm2' = 0				
'blm2' < 2				
'blm2' = 'blm2' + 1				
'Probenanzahl' mal Zugabe von BLM2				
Mischen				
WAHR		'Probenanzahl1' < 49 und 'blm2' = 1	FALSCH	
			Mischen der Magnetpartikel	
Separieren				
'Probenanzahl' mal Überstand abnehmen				
Inkubation bei 60°C				
WAHR		'Probenanzahl1' < 49 und 'blm2' = 1	FALSCH	
			Mischen der Magnetpartikel	

Abbildung 2.3: Ablaufschema Aufreinigung und Trocknen der Magnetpartikel

2. MATERIAL UND METHODEN

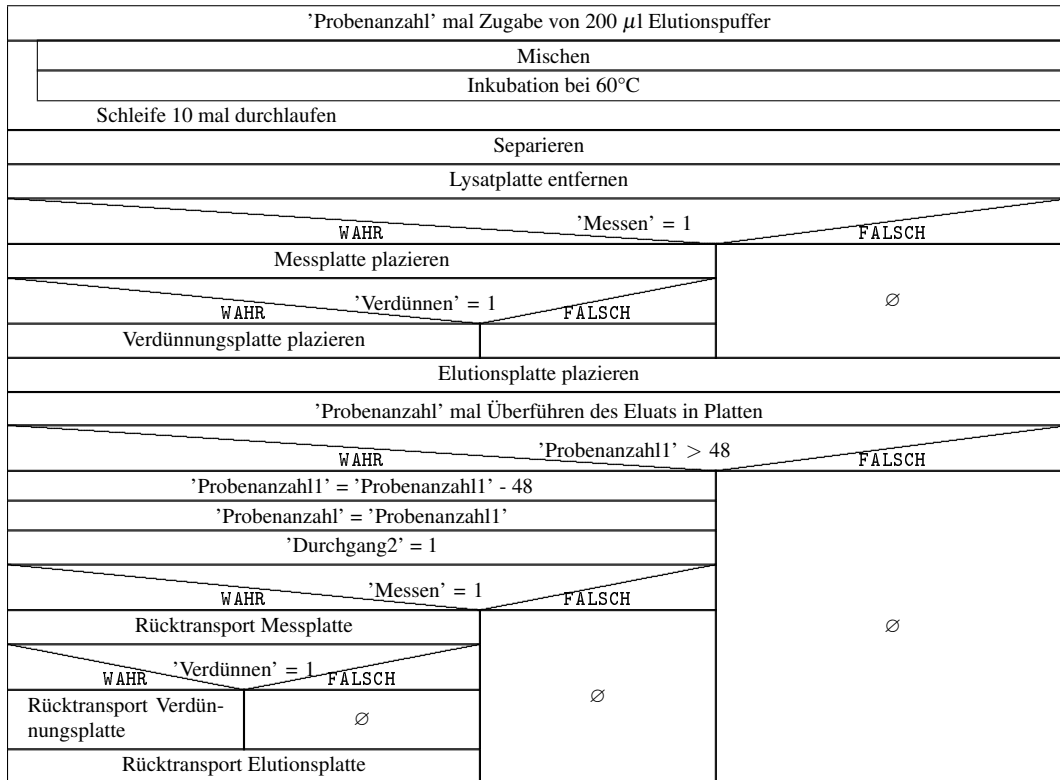


Abbildung 2.4: Ablaufschema Elution und Vorbereitung des zweiten Durchgangs

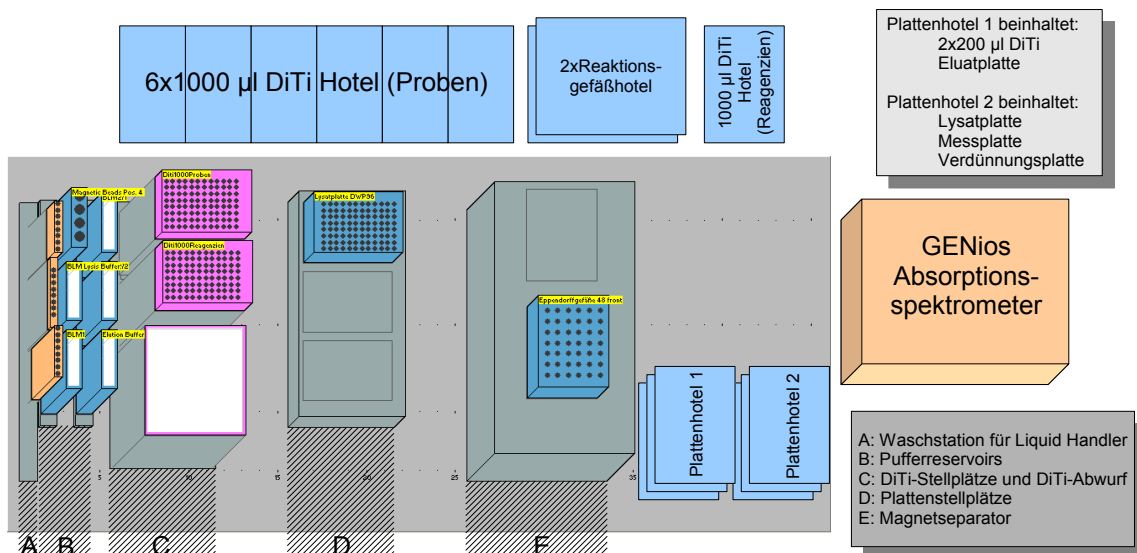


Abbildung 2.5: Bestückung des Arbeitstisches beim Start der Extraktion (DiTi: Box mit Pipettenspitzen). Es sind 1000 μ l DiTis auf Stellplatz C, die Lysatplatte auf Stellplatz D, sowie maximal 48 Reaktionsgefäße auf Stellplatz E positioniert. A-E: Träger auf der Arbeitsfläche

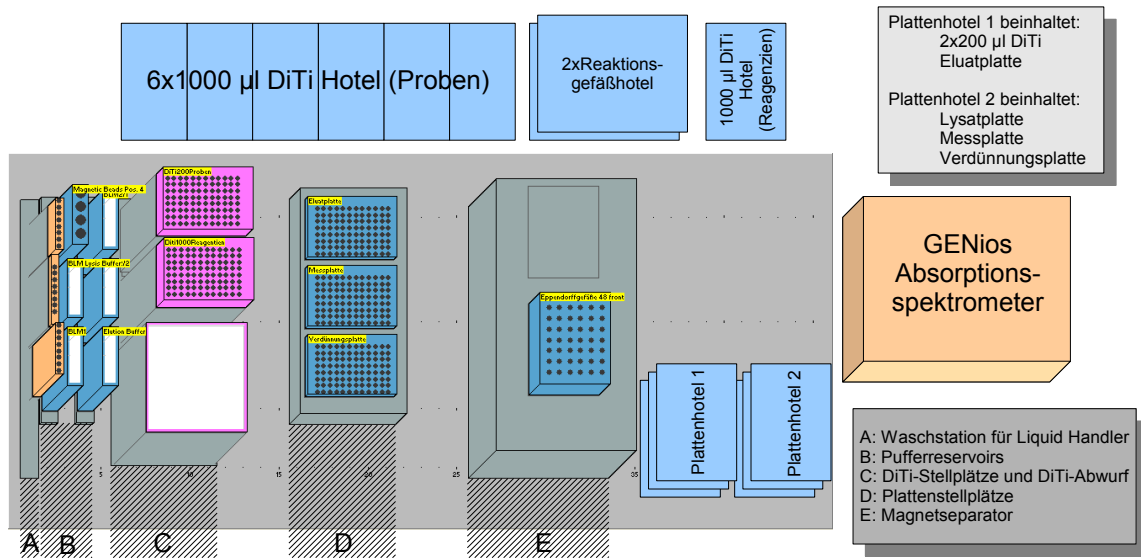


Abbildung 2.6: Bestückung des Arbeitstisches bei der Elution (DiTi: Box mit Pipettenspitzen). Es sind 200 µl DiTis auf Stellplatz C, eine Elutionsplatte, eine Messplatte und eine Verdünnungsplatte auf Stellplatz D und maximal 48 Reaktionsgefäße auf Stellplatz E positioniert. A-E: Träger auf der Arbeitsfläche

2.2.3. DNA Quantifikation mittels des Pegasus-Skripts

Die Quantifikation der DNA erfolgt durch ein Skript für die Software Pegasus, welches das GENios Absorptionsspektrometer ansteuert und eine Absorptionssmessung veranlasst. Sämtliche Parameter der Verdünnung können in der Excel-Vorlage, die durch das Skript `selectExcel.pts` in Gemini ausgewählt wird, eingestellt werden.

Nachdem alle Proben extrahiert wurden, wird das GENios Absorptionssmessgerät über die Pegasus-Module `InitExcel.pts` und `OpenDoor.pts` initialisiert und dessen Einzug geöffnet. Die Messplatte wird mit 80 µl Elutionspuffer versetzt, um 180° gedreht und zum Messgerät transportiert. Anschließend wird eine Absorptionssmessung bei einer Wellenlänge von 260 nm, 280 nm und 340 nm durchgeführt. Die erhaltenen Messwerte werden nach Formel 2.1 in DNA Konzentrationen umgerechnet und in eine Gemini Steuerdatei umgewandelt.

Die Messplatte wird an ihren ursprünglichen Platz gestellt und wieder um 180° gedreht. Danach wird die Gemini Steuerdatei abgearbeitet und alle Platten vom Arbeitstisch entfernt (siehe Abbildung 2.1).

2.2.4. DNA Quantifikation mittels des Skriptes `dilution.pl`

Vorbereitungen Bevor das Pegasus Skript vollständig einsatzbereit war, wurde die DNA Konzentration mittels des Excel-Makros XFluor4 Version 4.50 ermittelt. Die erhaltenen Rohwerte wurden in DNA Konzentrationen umgerechnet, wie in Formel 2.1 dargestellt.

2. MATERIAL UND METHODEN

Tabelle 2.3: Übersicht über die wichtigsten Variablen des Skripts 'dilution.pl'

Name	Typ	Funktion	Kontrolle	Eigenschaften
start	integer	enthält das Volumen der zu verdünnenden Proben	Benutzer	
maxp	integer	enthält das Maximalvolumen der Mikrotiterplatte	Skript	eingestellt auf 200
vol	float / integer	enthält die hinzuzugebenen Volumina	Tabelle, Skript	wird auf Ganzzahl gerundet, wenn $> \text{max} - \text{start}$ dann $\text{max} - \text{start}$
maxd	integer	Maximalvolumen des DiTi	Skript	eingestellt auf 1000
cal	integer	Kalibrationsvolumen des DiTi	Skript	eingestellt auf 50
aspmx	integer	Maximal aufzunehmendes Volumen	Skript	ist gleich $\text{maxd} - \text{cal}$
aspmx1	integer	Maximal aufzunehmendes Volumen abzüglich des höchsten hinzuzufügenden Wertes für vol	Skript	ist gleich $\text{maxd} - \text{cal} - \text{start} - \text{maxp}$
asp	integer	Startwert = 0, enthält Summe von vol	Skript	ist $< \text{aspmx}$

Konstanten

$bl = 0,05$: Grundabsorption (Plateblank) der eingesetzten Mikrotiterplatte bei 260 nm

$\alpha = 50 \text{ ng}/\mu\text{l}$: Der Extinktionskoeffizient für DNA in H_2O : $1A_{260\text{nm}} = 50 \text{ g/ml}$

$p = 0,29$: Die Pfadlänge des Lichts in Lösung (in mm)

Variablen

c : DNA-Konzentration einer Probe (in $\text{ng}/\mu\text{l}$)

v_{org} : Ausgangsvolumen der zu vermessenden Probe

v_{dil} : Volumen des Verdünnungspuffers

DNA-Konzentrationsberechnung

$$c = (A_{260\text{nm}} - bl) \cdot \alpha \cdot \frac{1}{\frac{v_{org}}{v_{dil}}} \cdot \frac{1}{p} \quad (2.1)$$

Die so ermittelten Konzentrationen c_{start} wurden in ein Excel-Arbeitsblatt kopiert, in der das Ausgangsvolumen v_{start} und die Endkonzentration c_{end} angegeben werden konnten. In dieser Liste wurde das Additionsvolumen v_{add} berechnet (siehe Formel 2.2).

$$v_{add} = \frac{v_{start} \cdot c_{start}}{c_{end}} - v_{start} \quad (2.2)$$

Wenn das errechnete Additionsvolumen kleiner 0 war (d.h. $c_{start} < c_{end}$), wurde das Additionsvolumen in dem Excel-Arbeitsblatt auf 0 gesetzt. Das Arbeitsblatt wurde als kommagetrennte Textdatei gespeichert.

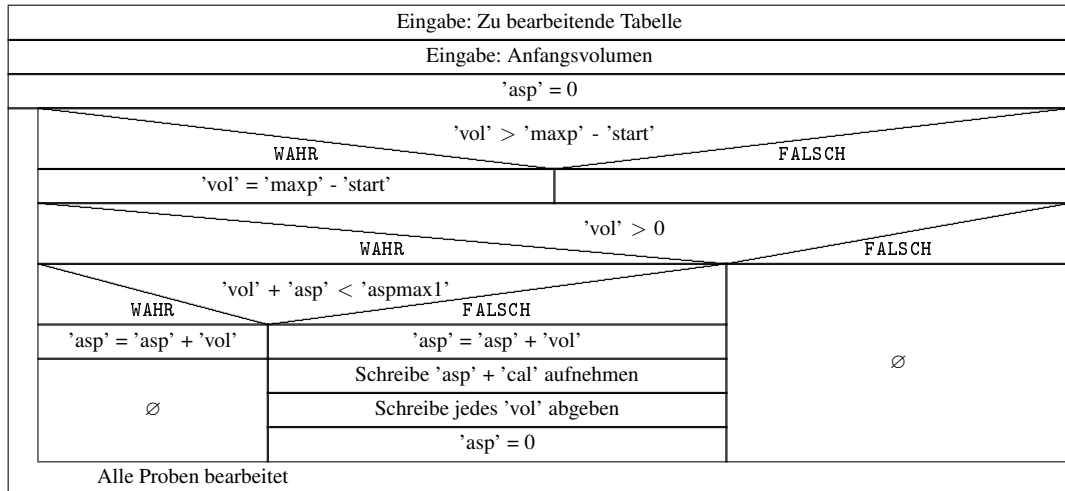


Abbildung 2.7: Struktogramm des Programmes 'dilution.pl'

Bearbeitung durch dilution.pl Mittels des Perl Skripts `dilution.pl` wurde die erhaltene Ausgabedatei in eine Steuerdatei für die Gemini Software übersetzt. Die zum Verständnis relevanten Variablen sind in Tabelle 2.3 aufgelistet. Der Benutzer hatte die Möglichkeit, das Volumen `start` der zu verdünnenden Proben anzugeben. Das Maximalvolumen `maxp` der standardmäßig benutzten Mikrotiterplatten wurde auf $200 \mu\text{l}$ festgesetzt.

Die abzugebenden Volumina `vol` wurden auf den Wert `maxp - start` reduziert, falls sie größer als `maxp - start` waren. Anschließend wurden sie intern gespeichert und bis maximal `aspmax1` und `asp` addiert. Werte für `vol`, die kleiner als $5 \mu\text{l}$ (Minimalvolumen des Liquid Handlers) waren, wurden übersprungen. Wenn dieser Wert übertroffen wurde, wurde der auslösende Wert von `vol` intern gespeichert und zu `asp` hinzuaddiert. Das aufzunehmende Volumen konnte also maximal `aspmax` betragen. Ein Kommando zum Aufnehmen von `asp + cal` wurde geschrieben. Danach wurden die Kommandos für die intern gespeicherten Werte von `vol` geschrieben gefolgt von einem Kommando, nach der letzten Abgabe das DiTi zu wechseln. Ein Struktogramm ist in Abbildung 2.7 dargestellt.

2.2.5. Protokollierung der DNA Extraktion

Um eine nachvollziehbare DNA Extraktion zu gewährleisten, wurde das Formular Arbeitsblatt_DNA_Extrakt1-2.xls eingeführt. Auf der ersten Seite wird eine Zusammenfassung des Auftrags geschrieben. Auf der zweiten Seite werden die individuellen 8-stelligen Maus-IDs aus der Mousenet Datenbank [143] aufgelistet. Dies geschieht in einer Weise, die die Positionen der Proben auf der 96-er DWP repräsentieren.

Dieser Auftrag wird sowohl in Papierform abgelegt als auch elektronisch gespeichert. Außer zur Dokumentation dient dieses Formular auch als Vorlage für die Probenbezeichnungen in der SNP Genotypisierung.

2.3. SNP Genotypisierung mittels MALDI-TOF MS

Der allgemeine Ablauf der SNP Genotypisierung gliedert sich in sechs Teile.

1. Amplifikation des zu analysierenden Locus mittels PCR [127]
2. Inaktivierung nicht verbrauchter dNTPs mittels Shrimp Alkalischer Phosphatase (SAP) [163]
3. Minisequenzierung durch allelspezifische Primerextension [26]
4. Deionisierung des Reaktionsgemisches
5. Spotten der Proben auf speziell präparierte Chips
6. MALDI-TOF Massenspektrometrie

Die Minisequenzierung kann auf zwei Arten erfolgen. Zum Einen gibt es die Möglichkeit, einen SNP durch die Zugabe eines Gemisches von Didesoxyribonukleosidtriphosphaten (ddNTP) und Desoxyribonukleosidtriphosphaten (dNTP) zu analysieren, sodass unter Zugabe eines locuspezifischen Extensionsprimers ein allelspezifischer Kettenabbruch entsteht. Die so erhaltenen Extensionsprodukte unterscheiden sich um die Masse mindestens eines Nukleotids (ca. 300 Da) [87].

Zum Anderen kann die allelspezifische Extension auch unter der ausschließlichen Zugabe von ddNTPs erfolgen. Die erhaltenen Extensionsprodukte unterscheiden sich dann lediglich in der Masse des allelspezifisch integrierten ddNTPs [153]. Der Massenunterschied reicht dabei von 9 Da (ddATP/ddTTP) bis 40 Da (ddCTP/ddGTP). Allerdings ist speziell der Unterschied von 9 Da zwischen ddATP und ddTTP zu gering, um im Hochdurchsatz zuverlässig erkannt zu werden. Es werden daher modifizierte ddNTPs eingesetzt, welche den Massenunterschied zwischen ddATP und ddTTP um 46,9 Da auf 55,9 Da erhöhen. Mit diesem Ansatz ist es momentan möglich, bis zu 30 SNPs gleichzeitig zu genotypisieren [149].

2.3.1. Homologe Massenextension (hME)

PCR Amplifikation Die PCR-Reaktionen fanden in 384 Well Mikrotiterplatten mit einem Reaktionsvolumen von 6 μl statt (Tabelle 2.4). In diesen Mikrotiterplatten befanden sich 5 ng eingetrocknete DNA je Probe (vgl. Abschnitt 2.5.1). Der Ansatz für 384 Proben wurde mittels der Liquid Handling Station Genesis RSP 150 (Tecan) verteilt und die PCR auf MJ Research Tetrad Thermocyclern (BioRad) mit dem in Tabelle 2.5 angegebenen Programm durchgeführt.

Tabelle 2.4: Einfacher PCR-Ansatz für die hME-Methode im 384 Well Format

Volumen	Reagenz	Ausgangskonzentration
4,52 μl	H ₂ O	
0,6 μl	10× Puffer	
0,6 μl	dNTPs	2mM
0,01 μl	Sense Primer	je 100 μM
0,01 μl	Antisense Primer	je 100 μM
0,24 μl	MgCl ₂	25 mM
0,02 μl	HotStar Taq	5 U/ μl

Das angegebene H₂O-Volumen ist bezogen auf einen Ansatz mit einem Primerpaar

Tabelle 2.5: Programm für die PCR im 384 Well Format

Temperatur	Dauer	Beschreibung	Zyklen
94°C	15 min	Denaturierung der DNA, Aktivierung der Polymerase	1×
94°C	20 sec	Denaturierung	45×
56°C	30 sec	Anlagerung	
72°C	1 min	Extension	
72°C	3 min	Extension	1×
20°C	1 sec	Abkühlung auf RT	

SAP Addition Zu jeder Reaktion wurden 2 μl des in Tabelle 2.6 angegebenen Gemisches gegeben, um unverbrauchte dNTPs zu inaktivieren. Der Ansatz für 384 Proben wurde durch die Multimek 96 Liquid Handling Station verteilt. Das in Tabelle 2.7 angegebene Programm wurde auf MJ Tetrad Thermocyclern durchgeführt.

Tabelle 2.6: Ansatz für die Inaktivierung der dNTPs für die hME-Methode im 384 Well Format

Volumen	Reagenz
1,53 μl	H ₂ O
0,17 μl	hME Puffer
0,3 μl	SAP

2. MATERIAL UND METHODEN

Tabelle 2.7: Programm für die Inaktivierung überschüssiger dNTPs im 384 Well Format

Temperatur	Dauer	Beschreibung
37°C	30 min	dNTP-Inaktivierung
85°C	10 min	Inaktivierung der SAP

Allelspezifische Primerextension Jede Reaktion erhielt 2 μl des in Tabelle 2.8 angegebenen Ansatzes, um die allelspezifische Primerextension durchzuführen. Der Ansatz für 384 Proben wurde durch die Multimek 96 Liquid Handling Station verteilt. Das in Tabelle 2.9 angegebene Programm wurde auf MJ Tetrad Thermocyclern durchgeführt.

Tabelle 2.8: Ansatz für die Extensionsreaktion im 384 Well Format mit der hME Methode

Volumen	Reagenz	Ausgangskonzentration
4,34 μl	H ₂ O	
0,6 μl	hME Extension Mix	
0,1 μl	Extensionsprimer	je 10 μM
0,02 μl	hME Enzym	1 U/ μl

Das angegebene H₂O-Volumen ist bezogen auf einen Ansatz mit einem Primerpaar.

Tabelle 2.9: Programm für die Extensionsreaktion im 384 Well Format für das hME-Protokoll

Temperatur	Dauer	Beschreibung	Zyklen
94°C	2 min	Denaturierung	75x
94°C	5 sec	Denaturierung	
52°C	5 sec	Anlagerung	
72°C	5 min	Extension	
20°C	1 sec	Abkühlung auf RT	

Ionenaustausch Es werden 3 mg eines Ionentauscher-Harzes (SpectroCLEAN, Sequenom) zu jeder Probe zugegeben. Anschließend werden 25 μl LiChrosolv H₂O (Merck) zu jeder Probe mit der Multimek 96 Liquid Handling Station zugegeben. Das Gemisch kommt für mindestens 20 min auf einen Rotator, und wird 10 min bei 3000 rpm abzentrifugiert.

Spotten der Proben Es werden 15 nl des Gemisches mit dem Spotter auf einen SpectroCHIP 384 (Sequenom) zusammen mit einem 3-Punkt Kalibranten (Sequenom) aufgetragen.

MALDI-TOF MS Die mit der Matrix der SpectroCHIPS kokristallisierten Proben werden mit einem auf das Format der SpectroCHIPS angepassten MALDI-TOF Massenspektrometer (Bruker Biflex, Bruker Autoflex, Sequenom MassARRAY Compact Analyzer) analysiert.

2.3.2. Einzelprimer-Massenextension iPlex

PCR Amplifikation Die PCR-Reaktionen fanden in 384 Well Mikrotiterplatten mit einem Reaktionsvolumen von 5 μl statt (Tabelle 2.10). In diesen Mikrotiterplatten befanden sich 5 ng eingetrocknete DNA der zu genotypisierenden Proben (vgl. Abschnitt 2.5.1). Der Ansatz für 384 Proben wurde mittels der Liquid Handling Station Genesis RSP 150 verteilt. Das eingesetzte PCR Programm ist analog zu dem in Tabelle 2.5.

Tabelle 2.10: Einfacher PCR-Ansatz für die iPlex-Methode im 384 Well Format

Volumen	Reagenz	Ausgangskonzentration
2,85 μl	H ₂ O	
0,625 μl	10xPuffer	
0,325 μl	MgCl ₂	25 mM
0,1 μl	dNTPs	2mM
1 μl	Primermix	je 10 μM
0,1 μl	HotStar Taq	5 U/ μl

Der Primermix besteht aus Sense und Antisense Primern mit einer Endkonzentration von 0,5 μM .

SAP Addition Zu jeder Reaktion wurden 2 μl des in Tabelle 2.11 angegebenen Gemisches gegeben, um unverbrauchte dNTPs zu inaktivieren. Der Ansatz für 384 Proben wurde durch die Tecan Aquarius Liquid Handling Station verteilt. Das in Tabelle 2.12 angegebene Programm wurde auf MJ Tetrad Thermocyclern durchgeführt.

Tabelle 2.11: Ansatz für die Inaktivierung überschüssiger dNTPs im 384 Well Format

Volumen	Reagenz	Ausgangskonzentration
1,53 μl	H ₂ O	
0,17 μl	SAP Puffer	
0,3 μl	SAP	

Tabelle 2.12: Programm für die PCR im 384 Well Format

Temperatur	Dauer	Beschreibung
37°C	20 min	dNTP-Inaktivierung
85°C	5 min	Inaktivierung der SAP

Allelspezifische Primerextension Jede Reaktion erhielt 2 μl des in Tabelle 2.13 angegebenen Ansatzes, um die allelspezifische Primerextension durchzuführen. Der Ansatz für 384 Proben wurde durch die Tecan Aquarius Liquid Handling Station verteilt. Das in Tabelle 2.14 angegebene Programm wurde auf MJ Tetrad Thermocyclern durchgeführt.

2. MATERIAL UND METHODEN

Tabelle 2.13: Ansatz für die Extensionsreaktion im 384 Well Format nach der iPlex Methode

Volumen	Reagenz	Ausgangskonzentration
0,755 μ l	H ₂ O	
0,3 μ l	10× iPlex Puffer	
0,2 μ l	iPlex Termination Mix	
0,804 μ l	Extensionsprimermix	
0,041 μ l	iPlex Enzym	1 U/ μ l

Der Primermix besteht aus Extensionsprimern in unterschiedlichen Konzentrationen und wird mit H₂O auf ein festes Volumen eingestellt.

Tabelle 2.14: Programm für die iPlex-Extensionsreaktion im 384 Well Format

Temperatur	Dauer	Beschreibung	Zyklen
94°C	30 sec	Denaturierung	
94°C	5 sec	Denaturierung	
52°C	5 sec	Anlagerung	5× 40×
80°C	5 sec	Extension	
72°C	3 min	Extension	
20°C	1 sec	Abkühlung auf RT	

Ionenaustausch Es werden 6 mg eines Ionentauscher-Harzes (SpectroCLEAN, Sequenom) zu jeder Probe zugegeben. Anschließend werden 25 μ l H₂O zu jeder Probe mit der Multimek 96 Liquid Handling Station zugegeben. Das Gemisch kommt für mindestens 20 min auf einen Rotator und wird 10 min bei 3000 rpm in einer Zentrifuge mit Ausschwingrotor abzentrifugiert.

Spotten der Proben Es werden 15 nl des Gemisches mit dem Spotter auf einen SpectroCHIP 384 (Sequenom) zusammen mit einem 3-Punkt Kalibranten (Sequenom) aufgetragen.

MALDI-TOF MS Die mit der Matrix der SpectroCHIPS kokristallisierten Proben werden mit einem auf das Format der SpectroCHIPS angepassten MALDI-TOF Massenspektrometer mit ausreichender Auflösung (Bruker Autoflex, MassARRAY Compact Analyzer) analysiert.

2.4. Erstellung der Markersätze

Es wurden Markersätze für die zwei in Abschnitt 2.3 erwähnten Protokolle erstellt. Der erste Markersatz für die hME Methode wurde so generiert, dass ein ungefährender Abstand von 30 Mb (15 cM) zwischen den Markern besteht. Der Abstand der Marker zueinander ist also wesentlich kleiner als der Grundabdeckungsradius (vgl. 1.4.6) und eine beobachtete Kopplung ist zu annähernd 100% signifikant. Die Anzahl der untersuchten Meiosen wurde auf 48 festgesetzt, um eventuellen Ausfällen von Markern und Proben entgegenzuwirken. Dies entspricht einer Anzahl von 48 Proben bei der Auswertung eines Rückkreuzungs-Experiments und einer Anzahl von 24 Proben bei der Auswertung einer Geschwisterverpaarung.

Dieser Markersatz wurde mit weiteren Markern ergänzt, die zwischen den Loci des ersten Markersatzes positioniert wurden. Mit dieser Kombination war es möglich, nicht nur eine Grobkartierung auf eine relativ große chromosomale Region zu erreichen, sondern in einem Schritt bereits eine sehr genaue Eingrenzung der betroffenen subchromosomalen Region auf 3 bis 4 cM zu erhalten. Der Abstand der Marker wurde so auf ca. 15 Mb (7,5 cM) verkleinert.

Der Markersatz für die iPlex Methode wurde von vorneherein so konstruiert, dass er eine Mischform zwischen einem Grobkartierungsansatz und einem Ansatz für die Feinkartierung darstellt (vgl. Abschnitt 1.5). Der Abstand der Marker betrug 15 Mb (7,5 cM). Zusätzlich wurden die Marker relativ zum Centromer um ca. 7,5 Mb (3,7 cM) versetzt, um eine Anordnung der Marker ähnlich der Abbildung 1.5 zu erhalten. Tabelle A.6 enthält die Sequenzen aller eingesetzten Primer.

2.4.1. Markersätze für die hME Methode

Unter Berücksichtigung der in Abschnitt 1.4.6 dargestellten statistischen Voraussetzungen wurden zunächst 69 Assays generiert, die das gesamte Mausgenom abdeckten. Der zweite Satz von 87 Markern wurde als Erweiterung des ersten Satzes konzipiert. Es wurden hierbei möglichst nur solche SNPs ausgewählt, die nicht nur zwischen den Stämmen C3HeB/FeJ und C57BL/6J, sondern auch zwischen C3HeB/FeJ und BALB/cByJ Heterozygotie zeigen.

Negative Validierungsergebnisse und der Ausfall von Assays im Routinebetrieb machten es notwendig, beide Markersätze gelegentlich mit neuen Assays zu ergänzen und vorhandene Assays neu zusammenzustellen, um eine möglichst hohe Multiplexrate beizubehalten.

Mangels Daten für den im ENU Screen verwendeten Mausstamm C3HeB/FeJ wurden zur Identifizierung von Markern stets die Genotypdaten der Stämme C3HHe/J und C3HHe benutzt. Aufgrund des hohen Verwandtschaftsgrads der Stämme [147] war nur mit einem geringen Fehler bei einem Test auf Heterozygotie zum Mausstamm C57BL/6J zu rechnen.

Erster Markersatz Eine Liste von SNPs, die Heterozygotie zwischen den Mausstämmen C57BL/6J und C3H/HeJ aufwiesen, wurde vom Centre National de Génotypage heruntergeladen. Es wurden SNPs und flankierende Sequenzen ausgewählt, die einen Abstand von ca. 30 Mb zueinander hatten. Insgesamt wurden 69 Marker ausgewählt.

2. MATERIAL UND METHODEN

Die Berechnung der Primer erfolgte mit dem Programm Spectrodesigner Version 1.3.4 (Sequenom) mit den vorgegebenen Standardeinstellungen und einer Multiplexrate zwischen 3 und 6 Assays pro Experiment.

Zweiter Markersatz Für die zweite Version des Markersatzes wurden SNP-Marker genutzt, die nicht nur zwischen dem Referenzstamm C3HeB/FeJ und dem Auskreuzungsstamm C57BL/6J polymorph sind, sondern zusätzlich noch Polymorphismen mit dem alternativen Auskreuzungsstamm BALB/cByJ zeigen. Diese Marker wurden mit einer Vorversion der Vergleichsroutine des ARTS Programmes (siehe Abschnitt 3.1) und den über das EnSMART Webinterface (vgl. Tabelle 2.21) heruntergeladenen Genotypdaten der Mausstämmen identifiziert. Die Berechnung der Primer erfolgte mit dem Assay Design Programm Version 2.0.0.17 (Sequenom) mit den Standardeinstellungen und ebenfalls mit einer maximalen Multiplexrate von 6 Assays pro Experiment.

Ergänzende Marker Gelegentlich kam es vor, dass ein Marker nicht die gewünschte Heterozygotie zeigte oder ein generierter Assay keine verlässlichen Ergebnisse lieferte. Um durch nicht funktionierende Assays auftretende Lücken im ersten und zweiten Markersatz zu kompensieren, wurden neue Multiplex-Experimente berechnet. Es wurden dieselben Kriterien wie bei dem zweiten Markersatz angewandt.

2.4.2. iPlex

Markersets für die iPlex Methode

Geeignete Marker-SNPs für die dritte Version des Markersatzes wurden mit Hilfe der aktuellsten Version der ARTS Suchmaschine identifiziert. Zusätzlich zu den in Abschnitt 2.4.1 erwähnten Stämmen wurde der Mausstamm AKR/J bei der Erstellung berücksichtigt. ENU Injektionen werden teilweise ebenso auf dem C57BL/6J Hintergrund ausgeführt und ein Auskreuzen mit C3HeB/FeJ ist nicht immer sinnvoll, da Mäuse dieses Stammes zu retinaler Degeneration [129] neigen, was sie unbrauchbar für ein Screening auf ENU-induzierte Mutationen des visuellen Systems macht.

Geeignete SNPs wurden zunächst mit den Standardeinstellungen von ARTS (vgl. [95] und Abschnitt 3.1) und dem Panel Generator identifiziert. Es wurden SNPs gesucht, die der Bedingung $AKR/J = C3HeB/FeJ \neq C57BL/6J = BALB/cByJ$ entsprachen und einen Abstand von 15 Mb (± 3 Mb) voneinander hatten. Der erste Marker sollte einen Abstand von 8 Mb (± 3 Mb) zum Centromer jedes Chromosoms haben. Zusätzlich sollten die identifizierten SNPs einen Anteil des häufiger vorkommenden Allels von maximal 0,6 haben.

Wenn keine geeigneten SNPs gefunden werden konnten, wurde die Suchparameter schrittweise gelockert. Zunächst wurde in einem Radius von 10 Mb um eine Zielregion nach Markern gesucht, dann wurde die Nummer der Berichte pro SNP auf 1 reduziert, anschließend wurde die Nummer der bekannten Genotypen pro SNP ebenfalls auf 1 gesetzt. Falls immer noch keine SNPs gefunden werden konnten, wurde der Anteil des am häufigsten vorkommenden Allels auf 1 erhöht. Zuletzt wurde der Filter für

unbekannte oder maskierte Nukleotide in den flankierenden Sequenzen auf 100% gesetzt und damit deaktiviert. Nahe verwandte Stämme wie der Stamm AKR und die verschiedenen BALB/c Unterstämme wurden mit in die Abfrage einbezogen. Als letzte Option wurde entweder der AKR/J oder der BALB/c-ByJ Stamm aus der Abfrage entfernt und die Suche zunächst mit den Standardparametern von Neuem begonnen und die Parameter wieder schrittweise gelockert. PCR- und Minisequenzierungsprimer für die Genotypisierung mit dem iPlex Protokoll wurden mit MassARRAY AssayDesign V3.0.1 (Sequenom) berechnet.

2.5. Standardisierung der Genotypisierung

Die gewählte Genotypisierungsmethode arbeitet mit Platten und Chips im 384 Well Format. Da unter Berücksichtigung der in Abschnitt 1.4.6 aufgeführten statistischen Gegebenheiten eine Genotypisierung von 384 Proben pro Mauslinie bei dem gewählten Markerabstand nicht notwendig ist, wurden mehrere Mauslinien pro Runde genotypisiert. Dies hat zur Folge, dass die Platzierung der Proben der verschiedenen Mauslinien in einer geordneten und reproduzierbaren Weise erfolgen muss.

Die Auswertung der Genotypisierungsergebnisse muss ebenfalls in einer geordneten Weise erfolgen, um ein falsch-positives oder falsch-negatives Ergebnis durch eine fehlerhafte Interpretation der Daten zu vermeiden.

2.5.1. Vorbereitung der Genotypisierung

Eine neue Genotypisierungsrunde wird mit Hilfe des Arbeitsblatts `Ansatz_Tochterplatten.doc` geplant. Hier wird die Verteilung der einzelnen Mauslinien und die Koordinaten der Kontrollen festgelegt. Die genaue Zuordnung der Proben zu den einzelnen Wells der 384 Well PCR Platte erfolgt unter Zuhilfenahme des Excel-Arbeitsblattes `384-erVorlage.xls`. Aus den DNA-Extraktionsprotokollen (vgl. Abschnitt 2.2.5) werden die Platzierungen der entsprechenden Proben übernommen und Kontrollen eingefügt. Mittels eines Excel-Makros werden die Proben in die einer 384 Well Platte entsprechende Reihenfolge gebracht und anschließend mit dem Plate Editor (vgl. Tabelle 2.22) in die Sequenom Datenbank importiert.

DNA-Proben mit einer Ausgangskonzentration von $50 \text{ ng}/\mu\text{l}$ werden in 96 Well Deep Well Platten (Qiagen) auf eine Endkonzentration von $2 \text{ ng}/\mu\text{l}$ gebracht. Von diesen Proben werden jeweils $5 \mu\text{l}$ (entsprechend einer Menge von $10 \mu\text{g}$) mittels eines Tecan Te-Mo Roboters in 384-Well PCR-Platten (Abgene) überführt. Die Proben in den 384 Well Platten werden danach über Nacht bei RT getrocknet.

Die Genotypisierung wird wie in Abschnitt 2.3 beschrieben durchgeführt.

2.5.2. Auswertung der Genotypisierungsergebnisse

Jeder Assay wird mit dem Programm Genotyper Version 3.4 (Sequenom) auf die allgemeine Gültigkeit der Ergebnisse überprüft. Bevor ein Assay zur Auswertung herangezogen wird, müssen folgende Bedingungen erfüllt sein:

2. MATERIAL UND METHODEN

1. Es muss annähernd eine Genotypverteilung vorliegen, die der erwarteten entspricht. Wenn zum Beispiel nur F₂ Tiere genotypisiert werden, sollte ein Genotyp-Verhältnis von ca. 1:2:1 vorliegen.
2. Es müssen klar voneinander abgrenzbare Genotyp-Cluster vorliegen.
3. Positiv- und Negativ-Kontrollen müssen das erwartete Ergebnis zeigen.

Wenn diese Bedingungen für einen Assay zutreffen, werden die nicht von der automatischen Genotypisierung erfassten Ergebnisse vom Benutzer nachbearbeitet. Hierbei müssen folgende Regeln beachtet werden:

1. Das Ergebnis muss in der Nähe eines Clusters von bekannten Genotypen liegen.
2. Das Maximum der Peaks muss in den Bereich der erwarteten Peaks fallen.
3. Die Peaks müssen sich deutlich von ihrer Umgebung abheben.

2.6. Arbeiten mit der MyGenotype Datenbank

2.6.1. MyImporter

Die Genotypisierungsergebnisse werden mittels des Skripts `MyImporter` (siehe 3.4.1) in die MyGenotype Datenbank eingelesen.

Daten, die während des Einlesens nicht die automatisierte Validierung bestehen, werden nochmals vom Benutzer kontrolliert und in die Kategorien „Valid / Problems / Invalid“ eingeteilt, bei den letzteren beiden Kategorien muss eine Begründung ausgewählt werden.

Eine ausführliche Beschreibung der Vorgänge befindet sich in Abschnitt 3.4.1.

2.6.2. MyMaintenance

Mit dem Programm `MyMaintenance` werden allgemeine Aufgaben der Datenbank geregelt (vgl. Abschnitt 3.4.2). Das Programm kann momentan folgende Aufgaben ausführen:

- Sicherheitskopie der Datenbank erstellen
- Import und Export von Assaylisten
- Aktualisierung der Markerloci über ARTS
- Aktualisierung von Genotypen für verschiedene Mausstämme
- Erstellung eines Berichts für einen Genotypisierungslauf

2.6.3. MyLine

Vor einem Genotypisierungslauf wird die Belegung der Platte festgelegt, wie in 2.5.1 dargestellt. An die Sequenom-Datenbank werden nur Proben-IDs ohne weiterführende Informationen übermittelt, was zur Folge hat, dass nach der Auswertung der Genotypisierung und dem Import der Ergebnisse mit MyImporter eine neuerliche Zuordnung der Proben zu den einzelnen Mauslinien stattfinden muss. Dies geschieht mit dem Programm MyLine. Neben der Zuordnung von Proben-IDs kann festgelegt werden, ob es sich um einen Wildtyp, eine Mutante oder ein Individuum mit unbekanntem Phänotyp handelt. In Abschnitt 3.4.3 wird das Vorgehen ausführlich behandelt.

2.6.4. MyGenotype

Die Daten der neu angelegten Linien können über das MyGenotype-Frontend analysiert oder zur Weiterverarbeitung mit anderen Programmen exportiert werden. Es stehen folgende Auswertungen zur Verfügung:

- Auswertung im HTML-Format mit χ^2 Kopplungsanalyse.
- Darstellung der Ergebnisse im R/QTL Importformat [24].
- Darstellung der Ergebnisse zur Weiterverarbeitung im MapManager [114].
- Darstellung der Ergebnisse zum Einlesen in dChip [106].

Das genaue Vorgehen des Programms wird in Abschnitt 3.4.4 dargestellt.

2.7. Molekularbiologische Methoden

2.7.1. Manuelle DNA Extraktion

Zur manuellen Extraktion von DNA aus Schwanzbiopsien von ca. 1 cm Länge wurde das auf Präzipitation basierende PUREGENE Kit (GENTRA) benutzt. Das Protokoll des Herstellers wurde modifiziert. 300 μ l Cell Lysis Solution wurde mit 1,5 μ l 20 mg/ml Proteinase K (Sigma) versetzt und über Nacht bei 55°C inkubiert. Das Lysat wurde 3 min mit einer Tischzentrifuge (TZ) bei voller Geschwindigkeit zentrifugiert und 250 μ l des Überstands wurden abgenommen. 100 μ l Protein Precipitation Solution wurden zugegeben, für 15 sec gevortext und auf Eis für 5 min inkubiert. Es wurde 3 min mit einer TZ bei voller Geschwindigkeit zentrifugiert und 300 μ l in ein neues Eppendorf Reaktionsgefäß überführt. 300 μ l 100% Isopropanol wurden zugegeben und durch Invertieren des Reaktionsgefäßes gemischt. Nach Zentrifugation der Probe (1 min bei voller Geschwindigkeit der TZ) wurde der Überstand verworfen, das gebildete Pellet kurz getrocknet und 300 μ l 70% Ethanol zugegeben und das Reaktionsgefäß wiederum invertiert. Nach Abziehen des Überstands und Trocknung des Pellets wurde mit 100 μ l DNA Hydration Solution die DNA über Nacht bei Raumtemperatur (RT) rehydriert.

2.7.2. RNA-Extraktion und Reverse Transkription

Die Extraktion der RNA erfolgte aus der Oberschenkelmuskulatur oder der Leber mit dem RNeasy Midi Kit (Qiagen) nach Herstellervorgabe.

Die erhaltene RNA wurde mit einem BioPhotometer (Eppendorf) per Absorptionsmessung bei 260 nm quantifiziert und mit einem 1%-igen Agarosegel mit Formaldehyd auf Degradation überprüft.

Die RNA wurde für die Analyse mittels einer qualitativen und einer quantitativen Reverse Transkriptase PCR (RT-PCR) präpariert.

Je 10 µg der gesamt-RNA aus dem Muskel von fünf Wildtyptieren wurde zu einem Pool zusammengestellt. 2 µg der gesamt-RNA von fünf *Zmpste24^{SMA004/SMA004}* Tieren und des Pools wurden mit den Komponenten des SSRT II Kits (Invitrogen) unter Benutzung von Oligo(dT)_{15–25} Primern (MWG Biotech) zur spezifischen reversen Transkription von mRNA in cDNA revers transkribiert (vgl. Tabelle 2.15 und Tabelle 2.16).

Auf diese Weise wurden auch je 2 µg der gesamt-RNA aus der Leber von fünf Wildtyptieren und fünf *Zmpste24^{SMA004/SMA004}* Tieren revers transkribiert.

Tabelle 2.15: Mastermixe für reverse Transkription

Mix	Volumen/Menge	Reagenz	Ausgangskonzentration
1	1 µl	Oligo(dT) _{15–25}	
	1 µl	dNTP	20 mM
	2 µg	gesamt-RNA	2 µg
	bis 12 µ	H ₂ O	
2	4,5 µl	First-Strand Buffer	
	2 µl	DTT	0,1 M
	1 µl	Superscript II	200 U/µl

Tabelle 2.16: Protokoll zur reversen Transkription

Mix	Temperatur	Dauer	Funktion
1	65°C	10 min	Denaturieren der RNA
	0°C (Eis)	≥ 1 min	Verhindern der Renaturierung
2	42°C	≥ 1 h	reverse Transkription
	70°C	15 min	Denaturierung des RT-Enzyms

2.7.3. Sequenzierung

Eine Standard PCR Reaktion wurde mit gepoolter genomischer DNA von jeweils fünf betroffenen Tieren einer Mutantenlinie und fünf Tieren aus der C3HeB/FeJ Stammzucht durchgeführt. Die eingesetzten Primer sind in Abschnitt A.1 und A.2 aufgeführt. Die Aufreinigung der PCR Produkte erfolgte mit AMPure (GC biotech). Nach der Sequenzierungsreaktion mit dem Big Dye Kit (Invitrogen) nach den Hersteller-

vorgaben wurden ungebundene fluoreszenzmarkierte ddNTPs mit CleanSEQ (GC biotech) entfernt. Die Sequenzierung fand auf einem 3730 DNA Analyzer (Applied Biosystems) statt.

2.7.4. RT-PCR

Die cDNA aus der Leber von einem WildtypTier und einem *Zmpste24*^{SMA004/SMA004} Tier wurde 1:10 verdünnt und als Vorlage für eine Standard PCR mit den Primern für die Amplifikation möglicher Splicevarianten benutzt (vgl. Tabelle A.4). Die Auswertung erfolgte auf einem 3%-igen TBE-Agarosegel (Biozym). Anschließend wurden die erhaltenen Produkte wie in Abschnitt 2.7.3 beschrieben sequenziert.

2.7.5. Quantitative RT-PCR

Tabelle 2.17: Programm für die quantitative RT-PCR

Temperatur	Dauer	Beschreibung	Zyklen
95°C	15 min	Denaturierung der DNA, Aktivierung der Polymerase	1×
95°C	15 sec	Denaturierung	45×
55°C	20 sec	Anlagerung	
72°C	20 sec	Extension	
53°C-95°C	0,5°C/sec	Schmelzkurve	1×

Primer für die Quantifizierung wurden mit Hilfe von Primer3 und der cDNA Sequenzen der in Tabelle 3.11 aufgelisteten Gene generiert. Um eine lineare Amplifizierung von genomischer DNA zu vermeiden wurden die Primer so gewählt, dass sie möglichst am Übergang zweier Exons lagen. Die maximale Produktlänge der PCR-Produkte betrug 200 bp und die Vorgabe für die Annealingtemperatur betrug 60°C.

Die cDNA aus der reversen Transkription wurde 1:10 verdünnt und als Vorlage für die Quantifizierung benutzt. Die relative Quantifizierung erfolgte mit dem LightCycler Real-Time PCR System (Roche) unter Zuhilfenahme des QuantiTect SYBR Green PCR Kits (Qiagen) nach den Herstellervorgaben mit Primern in einer Konzentration von 10 mM. Zur Identifikation von differenziell exprimierten Kandidatengenen wurden die Proben im Duplikat auf differenzielle Expression der in Tabelle genannten Gene untersucht. Eine Effizienzkurve wurde mit Hilfe von unverdünnter, 1:10, 1:100, 1:1000 verdünnter cDNA einer der Wildtyp-cDNAs für die Gene *Hprt1* und *Pcsk9* erstellt. Diese Probe diente als Kalibrator-cDNA. Die Expression von *Pcsk9* der 1:10 verdünnten cDNAs der fünf *Zmpste24*^{SMA004/SMA004} und der fünf Wildtyp-Proben, aus denen der Pool gebildet wurde, wurde auf diese Weise in Relation zu *Hprt1* quantifiziert. Das Vorgehen entspricht der Beschreibung im Handbuch zu RelQuant V1.01 (Roche).

2.7.6. Proteinextraktion und Western Blot

Proteine wurden aus den Nieren von zwei *Zmpste24*^{+/+} Tieren, einem *Zmpste24*^{SMA004/+} Tier und einem *Zmpste24*^{SMA004/SMA004} Tier mit dem RNA/Protein Extraction Kit (Macherey-Nagel) nach den Herstellervorgaben extrahiert. Extrahierte Proteine wurden nach ihrer Masse mittels Sodium Dodecyl Sulfate (SDS) Polyacrylamide Gel Electrophoresis (PAGE) unter reduzierenden Bedingungen aufgetrennt. Hierfür wurden NuPAGE Novex 4/12% Bis-Tris Gele (Invitrogen) und eine XCell SureLock Kammer (Invitrogen) eingesetzt. Die Auftrennung erfolgte nach den Herstellervorgaben mit NuPAGE MES SDS Laufpuffer (Invitrogen).

Die nach ihrer Masse aufgetrennten Proteine wurden von den Polyacrylamidgelen in der XCell SureLock Kammer und dem zugehörigen Blot Modul auf PVDF Membranen (Poll) mit einem Tris Glycin Transferpuffer überführt. Der Western Blot wurde mit dem Primär-Antikörper (PAK) anti-Lamin A (SC-6214, Santa Cruz Biotechnology) in einer Verdünnung von 1:200 durchgeführt. Als Sekundär-Antikörper (SAK) kam ein mit Meerrettich Peroxidase markierter Maus-Anti-Ziege Antikörper (Dianova) zum Einsatz. Der Western Blot wurde wie bei Ausubel *et al.* beschrieben durchgeführt [8, Kapitel 10-5].

Tabelle 2.18: Protokoll für den Western Blot

Dauer	Schritt	Reagenzien	Wiederholungen
1 h	Blocken	TTBS mit 5% Magermilchpulver	
1 h	Bindung des PAK	1:200 PAK in TTBS mit 0,5% Magermilchpulver	
15 min	Waschen	TTBS	4x
1 h	Bindung des SAK	1:2500 SAK in TTBS mit 0,5% Magermilchpulver	
15 min	Waschen	TTBS	4x

2.7.7. Genotypisierung mittels Tetra-Primer ARMS PCR

Da eine Genotypisierung der G→A Mutation durch die Analyse eines Restriktionsfragmentlängenpoly-morphismus (RFLP) mangels einer geeigneten Restriktionsschnittstelle nicht infrage kam, wurde eine PCR-basierte Genotypisierung entwickelt, die auf der von Ye *et al.* vorgestellten Tetra-primer Amplification Rerfractory Mutation System (ARMS)-PCR Methode [210] beruhte. Mit dem von Ye *et al.* entwickelten Programm zur Erstellung von Primern für die Tetra Primer ARMS PCR (vgl. Tabelle 2.21) wurde ein Satz von vier Primern (vgl. A.3) gefunden, mit dem fortan die Genotypisierung der genomischen DNA möglich war.

Das äußere Primerpaar dient der Amplifikation des Locus, während das innere Primerpaar mit je einem der äußeren Primer ein PCR-Produkt bildet. Diese PCR-Produkte haben unterschiedliche Größen und können auf einem Polyacrylamid- oder einem Agarosegel aufgetrennt werden. Zur Erhöhung der Spezifität der Reaktion besitzen die inneren Primer je ein falsches Nukleotid an Position 2 nach dem allelspezifischen Nukleotid. Zusätzlich wird eine Touchdown-PCR durchgeführt, die Annealingtemperatur

wird also von 72°C bei jeder Wiederholung solange um ein 1°C gesenkt, bis die endgültige Annealingtemperatur erreicht ist.

Tabelle 2.19: Einfacher Ansatz für die Tetra Primer ARMS PCR

Volumen	Reagenz	Ausgangskonzentration
5,6	μl H ₂ O	
1	μl 10× HotStar Puffer	
0,1	μl dNTPs	20mM
<i>je</i> 0,1	μl Äußere Primer	je 10 μM
<i>je</i> 1	μl Innere Primer	je 10 μM
0,1	μl HotStar Taq	5 U/ μl
1	μl gDNA	50 ng/ μl

Tabelle 2.20: Programm für die Tetra Primer ARMS PCR

Temperatur	Dauer	Beschreibung	Zyklen
94°C	15 min	Denaturierung der DNA, Aktivierung der Polymerase	1×
94°C	1 min	Denaturierung	7× (-1°C/Zyklus)
72°C-66°C	1 min	Anlagerung	
72°C	1 min	Extension	
94°C	1 min	Denaturierung	28×
66°C	1 min	Anlagerung	
72°C	1 min	Extension	
72°C	3 min	Extension	
20°C	1 sec	Abkühlung auf RT	1×

2.8. Röntgenmethoden

2.8.1. Röntgenabsorptiometrie

Je fünf weibliche Mäuse der Gruppen *Zmpste24^{+/+}* und *Zmpste24^{SMA004/SMA004}* und vier Mäuse aus der *Zmpste24^{SMA004/+}* Gruppe im Alter von 14 bis 16 Wochen wurden mittels peripheral Dual Energy X-Ray Absorptiometry (pDEXA, Norland Medical Systems) mit je zwei Messungen gemäß Protokoll der Deutschen Mauslinik untersucht [59].

2.8.2. Computertomographie (CT)

Je zwei 14-16 Wochen alte Tiere der Gruppen *Zmpste24^{+/+}*, *Zmpste24^{SMA004/+}*, *Zmpste24^{SMA004/SMA004}* wurden mittels eines Computertomographen (Vamp) gemäß Protokoll der Deutschen Mauslinik untersucht [59]. Mäuse wurden pro Aufnahme um 0,5° gedreht. Der erhaltene Datensatz aus 720 Röntgenaufnahmen wurde in die syngo Darstellungssoftware (Siemens) importiert und dort rekonstruiert und bearbeitet.

2.9. Material

2.9.1. Puffer, Lösungen und Reagenzien

Reagenzien

- 37% Formaldehyd (Sigma, Seelze)
- DNA Agarose (Biozym, Hessisch-Oldendorf)
- Ampuwa H₂O (Fresenius Kabi, Bad Homburg)
- Borsäure (Sigma, Seelze)
- Ethanol (Merck, Darmstadt)
- Ethylene-diamine-tetraacetic acid (EDTA) (Sigma, Seelze)
- Glycin (Sigma, Seelze)
- Isopropanol (Merck, Darmstadt)
- LiChrosolv H₂O (Merck, Darmstadt)
- Magermilchpulver (Sigma, Seelze)
- Natrimchlorid (Sigma, Seelze)
- SDS (Sigma, Seelze)
- SpectroCLEAN (Sequenom, Hamburg)
- Tris·HCl (Sigma, Seelze)
- Tween 20 (Sigma, Seelze)

Puffer und Lösungen von Drittherstellern

- 3-Punkt Kalibrant (Sequenom, Hamburg)
- 10× Puffer für PCR (Qiagen, Hilden)
- 25 mM MgCl₂ (Qiagen, Hilden)
- dNTP Mix (Fermentas, St. Leon-Rot)
- Oligo(dT)_{15–25} Primermix (MWG Biotech, Ebersberg)
- hME MassEXTEND Nucleotide Tri-Mix Pack (Sequenom, Hamburg)
- hME Puffer (Sequenom, Hamburg)
- iPlex Pack (Sequenom, Hamburg)

- NuPAGE MES SDS Running Buffer (Invitrogen, Karlsruhe)
- NuPAGE Novex 4/12% Bis-Tris Gel (Invitrogen, Karlsruhe)
- SAP Puffer (Sequenom, Hamburg)

Längenstandards für die Analyse von DNA und Proteinen

- GeneRuler 100 bp DNA Ladder Plus (Fermentas, St. Leon-Rot)
- pUC Mix Marker, 8 (Fermentas, St. Leon-Rot)
- SeeBlue Plus2 Pre-Stained Standard (Invitrogen, Karlsruhe)
- SmartLadder 200 bp to 10 kb (Eurogentec, Seraing, Belgien)

Selbst hergestellte Puffer und Lösungen

- 10× Cell Lysis Buffer
 - 10 mM Tris-HCl, pH=8,0
 - 50 mM EDTA
 - 1% SDS
- 10× Tris Glycin Puffer
 - 25 mM Tris-HCl, pH=8,3
 - 192 mM Glycin
- 10× TBE Puffer
 - 89 mM Tris-HCl, pH=8,0
 - 89 mM Borsäure
 - 2 mM EDTA
- TTBS Puffer
 - 50 mM Tris-HCl, pH=7,4
 - 500 mM Natriumchlorid
 - 0,05% Tween 20

2.9.2. Enzyme und Antikörper

Enzyme

- Proteinase K (Sigma, Seelze)
- HotStar PCR Taq DNA Polymerase (Qiagen, Hilden)
- Thermosequenase (Sequenom, Hamburg)
- Super Script II Reverse Transkriptase (Invitrogen, Karlsruhe)
- SAP (Sequenom, Hamburg)

Antikörper

- Lamin A (C-20) SC-6214 Antikörper (Santa Cruz Biotechnology, Heidelberg)
- Maus-anti-Ziege Antikörper, HRPO-gekoppelt (Dianova, Hamburg)

2.9.3. Geräte

Liquid Handling Roboter

- Freedom Evo 150 mit Aufbauten (Seriennummer 419, Tecan, Crailsheim)
- Genesis RSP 150 mt Aufbauten (Seriennummer 6621, Tecan, Crailsheim)
- Temo mit 96-Nadelkopf (Seriennummer 1087, Tecan, Crailsheim)
- Multimek 96 Automated 96-Channel Pipettor (Beckman Coulter, Fullerton, USA)
- Aquarius 96 (Tecan, Crailsheim)
- MassARRAY Nanodispenser (Sequenom, Hamburg)

Zentrifugen

- Heraeus Biofuge Pico (Thermo Electron, Langenselbold)
- Heraeus Biofuge Fresco (Thermo Electron, Langenselbold)
- Eppendorf 5810 R mit Ausschwingrotor (Eppendorf, Hamburg)
- Sigma 4K15C (Sigma Laborzentrifugen, Osterode)
- Rotana 46 RS (Hettich, Darmstadt)

Geräte für die PCR

- MJ Research PTC 225 Tetrad Peltier Thermal Cycler (BioRad, München)
- BioRad DNA Engine Tetrad 2 Peltier Thermal Cycler (BioRad, München)
- LightCycler (Roche, Mannheim)

TOF Massenspektrometer

- Biflex (Bruker Daltonics, Bremen)
- Autoflex (Bruker Daltonics, Bremen)
- MassARRAY Compact Analyzer (Sequenom, Hamburg)

Absorptionsphotometer

- NanoDrop ND-1000 (PEQLAB, Erlangen)
- BioPhotometer (Eppendorf, Hamburg)

Rotator

- Roto-Shake Genie (Scientific Industries, Bohemia, USA)

Gerät zur DNA Sequenzierung

- 3730 DNA Analyzer (Applied Biosystems, Darmstadt)

SDS-PAGE und Western Blotting Kammer

- XCell SureLock (Invitrogen, Karlsruhe)

Röntgenabsorptiometrie

- pDEXA Sabre X-ray Bone Densitometer (Norland Medical Systems, Basingstoke, UK)

Computertomographie

- Micro-CT (Vamp, Erlangen)

2.9.4. Reaktionskits

DNA Extraktion

- PUREGENE Genomic DNA Isolation Kit (GENTRA, Minneapolis, USA)
- AGOWAmag Maxi DNA Kit (AGOWA, Berlin)

2. MATERIAL UND METHODEN

RNA Extraktion

- RNeasy Midi Kit (Qiagen, Hilden)

RNA/Protein Extraktion

- RNA/Protein Extraction Kit (Macherey-Nagel, Düren)

Quantitative RT-PCR

- QuantiTect SYBR Green PCR Kit (Qiagen, Hilden)

Sequenzierung

- AMPure PCR Purification Kit (GC biotech, Rotterdam, Niederlande)
- BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Darmstadt)
- CleanSEQ Big Dye Terminator Removal Kit (GC biotech, Rotterdam, Niederlande)

2.10. Anwendungen, Programme und Datensätze

Tabelle 2.21: Liste der benutzten web-basierten Anwendungen

Name	Kurzbeschreibung	URL	Referenz
UCSC Genome Browser	Allgemeines Sequenzwerkzeug	genome.ucsc.edu	[90]
BLAST	Genomweites Alignment	www.ncbi.nlm.nih.gov/blast	[4]
OMIM	Beschreibungen aller menschlichen Erbkrankheiten	www.ncbi.nlm.nih.gov/omim	[120]
Mouse Genome Informatics (MGI)	Mausspezifische Ressourcen	www.informatics.jax.org	[16, 44]
EnsmART	Datamining Tool	ensembl.ctaalliance.org/ Ensmart	[91]
Primer3	Primerdesign Tool	frodo.wi.mit.edu/cgi-bin/ primer3/primer3_www.cgi	[154]
ExonPrimer	Primerdesign Tool	ihg.gsf.de/ihg/ExonPrimer. html	
Tetra primer ARMS PCR	Primerdesign Tool	cedar.genetics.soton.ac. uk/public_html/primer1. html	[210]
NetGene2	Splice Site Vorhersage	www.cbs.dtu.dk/services/ NetGene2	[70]
Bibliosphere	Literaturrecherche	www.genomatix.de/ products/BiblioSphere/ BiblioSpherePE1.html	

Tabelle 2.22: Liste der benutzten lokal installierten Programme

Name	Version	Kurzbeschreibung	Hersteller	URL	Referenz
Perl	5.8.1	Programmiersprache	Open Source	www.perl.org	
MySQL	4.0.1	Datenbank	Open Source	www.mysql.com	
Gemini	4.2 SP 3	Steuerungsprogramm für Liquid Handling Roboter	Tecan, Crailsheim	www.tecan.de	
Pegasus	3.0	Prozesssteuerung für Automation	Tecan, Crailsheim	www.tecan.de	
XFLuor4	4.50	Makro zur Berechnung der DNA Konzentration	Tecan, Crailsheim	www.tecan.de	
MassARRAY AssayDesign	3.0.1	Primerdesign	Sequenom	www.sequenom.com	
Phred	0.020425.c	Nukleotidbenennung (Base calling)	B. Green / P. Ewing	www.phrap.org	[48, 47]
Phrap	0.990329	Sequenzalignment	B. Green / P. Ewing	www.phrap.org	
PolyPhred	4.29	Polymorphismuserkennung	D.A. Nickerson	droog.mbt.washington.edu/PolyPhred.html	[134]
Consed	15.0	Grafische Sequenzdarstellung	D. Gordon	www.phrap.org	[63]
RelQuant	1.01	Realtime PCR Auswertung	Roche Diagnostics	www.roche-applied-science.com	
ClustalX	1.83	Sequenzalignment	Thompson <i>et al.</i>	bips.u-strasbg.fr/fr/Documentation/ClustalX	[193]
MassARRAY	3.4	Genotypisierungssoftware	Sequenom	www.sequenom.com	

Es wird jeweils die hauptsächlich oder zuletzt benutzte Version aufgeführt.

Tabelle 2.23: Liste der benutzten Datensätze

Name des Datensatzes	Inhalt	URL	Referenz
dbSNP	SNP Daten	ftp.ncbi.nlm.gov/snp/organisms/mouse_10090	[173]
SNP Discovery program	SNP Daten	www.cng.fr/download/snp_souris.xls	
LitMiner	Literaturrecherche	andromeda.gsf.de/litminer	[112]

3. Ergebnisse

3.1. Advanced Retrieval Tool for SNPs (ARTS)

Für die Generierung eines Markersatzes für die genomweite Kartierung von Mausmutanten ist es sinnvoll, bereits *in silico* aus der Fülle an verfügbarer Information (die neueste Version von ARTS, basierend auf dbSNP Build 126, enthält ca. 5 Millionen SNP Loci und ca. 47 Millionen Genotypen), diejenige herauszufiltern, die am erfolgversprechendsten ist und zur definierten Fragestellung passt. Es wurde ein Ansatz benötigt, der es dem Benutzer zu ermöglichen soll, möglichst einfach einen Satz von Markern zur Kartierung des Mausgenoms auf vielen verschiedenen genetischen Hintergründen zu erstellen. Die Auswahl der Marker sollte desweiteren so eingeschränkt werden können, dass eine hohe Erfolgswahrscheinlichkeit bei der Anwendung *in vitro* gegeben ist. In Abbildung 3.1 wird das allgemeine Design von ARTS dargestellt.

Zur Erreichung dieser Ziele wurde ein System entwickelt, welches die gegebenen Anforderungen erfüllt [95]. Das Programm `create.pl` bereitet die Datensätze der dbSNP [173] für die Species *Mus musculus* auf und schreibt die erhaltenen Daten in eine MySQL-Datenbank. Mit `arts.cgi` wurde eine komfortable Oberfläche zur Abfrage dieser Datensätze entwickelt. Beide Programme wurden in Perl geschrieben. Passende Datensätze lagen in XML Dateien vor, welche Informationen über SNP-Loci, die analysierten Mausstämme und deren Genotyp enthielten. Sie wurden aus dem Verzeichnis `/genotype` auf dem dbSNP FTP-Server (siehe Tabelle 2.23) heruntergeladen. Dateien mit den flankierenden Sequenzen der SNP-Loci im FASTA Format [144] wurden im Verzeichnis `/rs_fasta` gefunden. Die Daten wurden von dem Importprogramm `create.pl` in die Datenbank eingelesen. Die Daten aller Autosomen, sowie des X und Y Chromosoms und des Mitochondriums wurden berücksichtigt. Zur Vermeidung von Inkonsistenzen wurden die Daten nicht lokalisierbarer SNPs (aus Datei `gt_chrUn.xml`) nicht eingelesen. Diese Datensätze können über das ARTS Frontend (www.gsf.de/cgi-bin/arts.cgi) abgerufen werden. Beide Programme interagieren mit der ARTS Datenbank über das Perl Modul DBI.

3. ERGEBNISSE

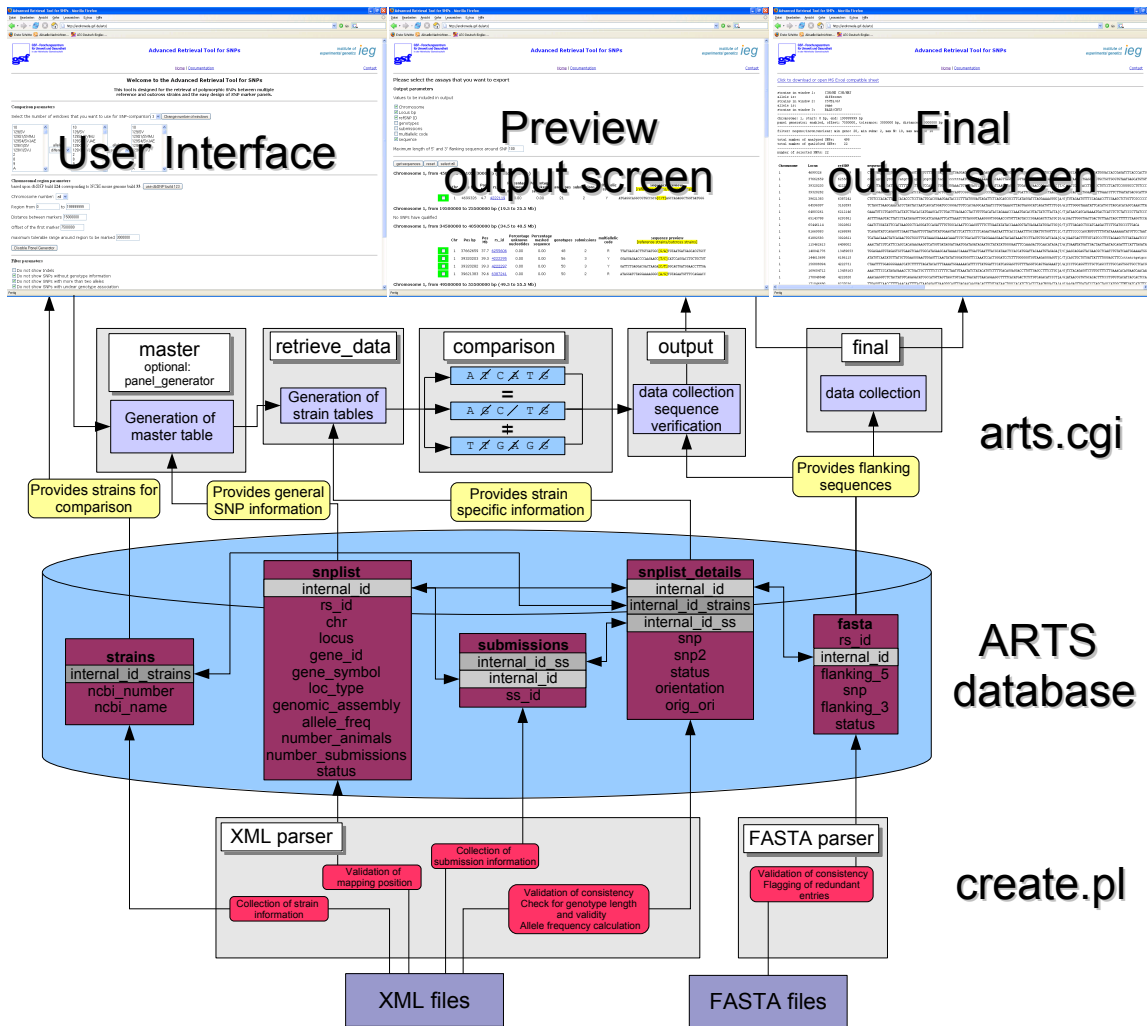


Abbildung 3.1: Flussdiagramm von ARTS. Dargestellt werden die unterschiedlichen Ebenen der Suchmaschine. Über arts.cgi werden sämtliche Benutzerinteraktionen gesteuert und Daten aus der ARTS Datenbank zusammengestellt. Die ARTS Datenbank selbst besteht aus fünf untereinander verknüpften Tabellen. Das Skript create.pl bereitet die Rohdaten der dbSNP auf und schreibt diese in die ARTS Datenbank.

3.1.1. Das Programm create.pl

Aufbereitung der FASTA Dateien

Jede FASTA Datei wird zunächst durch das Skript entkomprimiert und die `rs_id`, die flankierenden Sequenzen und SNP-spezifische Informationen aus den Einträgen extrahiert. Alle Datensätze werden auf Vollständigkeit und Gültigkeit der Genotyp-Angaben überprüft (vgl. Abbildung 3.2). Die Sequenzdaten werden in die Tabelle `fasta` eingetragen. Anschließend werden alle Marker auf Redundanz und Inkonsistenz getestet, `fasta` mit den entsprechenden Einträgen in der Spalte `status` versehen und die ursprüngliche Datei wieder komprimiert.

```

>gnl|dbSNP|rs8281485 rs=8281485 pos=251|len=501|taxid=10090|mol="mito"|class=1|alleles=C/T
build=118
CGCGGTATC CTGACCGTGC AAAGGTAGCA TAATCACTTG TTCCTTAATT AGGGACTAGC ATGAACGGCT AAACGAGGGT
CCAACCTGCT CTTATCTTTA ATCAGTGAAA TTGACCTTTC AGTGAAGAGG CTGAAATATA ATAATAAGAC GAGAAGACCC
TATGGAGCTT AAATTATATA ACTTATCTAT TTAATTTATT AAACCTAATG GCCCAAAAAC TATAGTATAA GTTTGAAATT
TCGGTTGGGG
GACCTCGGAG AATAAAAAAT CCTCCGAATG ATTATAACCT AGACTTACAA GTCAAAAGTAA AATCAACATA TCTTATTGAC
CCAGATATAT TTTGATCAAC GGACCAAGTT ACCCTAGGGA TAACAGCGCA ATCCTATTTA AGAGTTCATA TCGACAATTA
GGGTTTACGA CCTCGATGTT GGATCAGGAC ATCCCAATGG TGTAGAAGCT ATTAATGGTT CGTTTGTTC ACGATTAAAG
TCCTACGTGA
    
```

Legend:

- refSNP ID
- Allele mit monoallelischer / multiallelischer Nomenklatur (müssen identisch sein)
- flankierende Sequenzen

Abbildung 3.2: Die Zerlegung einer Sequenz im FASTA Format durch create.pl.

Aufbereitung der XML Dateien

Der Import der XML Dateien wird über das Perl Modul `XML::Parser` gesteuert. Dieses Modul gibt zwei Subroutinen namens `handle_start` und `handle_end` vor, welche es ermöglichen, beim Start oder Ende eines XML Elements mit einer bestimmten Bezeichnung die dazugehörigen Attribute und den Inhalt auszulesen. Aus Abbildung 3.3 wird das Vorgehen bei den beiden Hauptelementen `Individual` und `SnpInfo` sowie deren Unterelementen ersichtlich.

Eintragung von Daten in die Tabelle „strains“

Bei Auftreten des Elements `Individual` wird überprüft, ob der Wert des Attributs `indId` (enthält die Nummer eines Mausstammes) bereits in der Datenbank vorhanden ist. Wenn dies nicht der Fall ist, so wird die `indId` intern abgelegt und beim Start des in `Individual` eingebetteten `SubmitInfo` Unterelements die dazugehörigen Stammbezeichnungen der verschiedenen Labors (gekennzeichnet mit dem Attribut `submittedIndId`) intern gesammelt. Wenn das Element `Individual` beendet ist, wird die `indId` aus `Individual` zusammen mit der ersten Stammbezeichnung, die Buchstaben — also vermutlich einen realen Namen — enthält in die Tabelle `strains` geschrieben. Zur Vermeidung von Inkonsistenzen wird die Nummer `internal_id_strains` vergeben, welche die Zuordnung von Genotypen zu einem Mausstamm innerhalb der Datenbank eindeutig macht.

3. ERGEBNISSE

```

<Individual indId="4560" taxId="10090" sex="?">
  <SourceInfo source="TheJacksonLaboratory" sourceType="repository"
    ncbiPedId="3587" pedId="3587" indId="001800" mald="0" pald="0" />
  <SourceInfo source="TheJacksonLaboratory" sourceType="repository"
    ncbiPedId="3587" pedId="3587" indId="FVB/NJ" mald="0" pald="0" />
  <SubmitInfo popId="1182" submittedIndId="FVB/NJ" subIndGroup="Mouse" />
  <SubmitInfo popId="1445" submittedIndId="FVB/NJ" />
</Individual>

```

lies *indId*
 lies *submittedIndId*
 lies *submittedIndId*
 schreibe Stammdaten in *strains*

```

<SnpInfo rsId="3156436">
  <SnpLoc genomicAssembly="35:unknown" chrom="19" start="45605182" locType="
    2" rsOrientToChrom="rev"/>
  <SnpLoc genomicAssembly="35:C57BL/6J" chrom="19" start="45611168" locType="
    2" rsOrientToChrom="rev"/>
  <SnpLoc genomicAssembly="35:Celera" chrom="19" start="46476407" locType="
    2" rsOrientToChrom="rev"/>
  <SsInfo sslId="39694787" locSnpId="mCV23137671" ssOrientToRs="rev">
    <ByPop popId="1637" sampleSize="6">
      <AlleleFreq allele="G" freq="0.667"/>
      <AlleleFreq allele="A" freq="0.333"/>
      <GTypeFreq gtype="A/A" freq="0.333"/>
      <GTypeFreq gtype="G/G" freq="0.667"/>
      <GTypeByInd gtype="G/G" indId="2917"/>
      <GTypeByInd gtype="A/A" indId="2921"/>
      <GTypeByInd gtype="G/G" indId="2922"/>
    </ByPop>
  </SsInfo>
  <GTypeFreq gtype="C/C" freq="0.667"/>
  <GTypeFreq gtype="T/T" freq="0.333"/>
</SnpInfo>

```

lies *rsId*
 lies Locusdaten und Orientierung
 lies Locusdaten und Orientierung
 lies Locusdaten und Orientierung
 schreibe Locusdaten in *snplist*,
 lies *sslId* und *ssOrientToRs*,
 schreibe *sslId* in *submissions*
 lies *gtype* und *indId*
 lies *gtype* und *indId*
 lies *gtype* und *indId*
 schreibe Genotypdaten in
snplist_details, schreibe
 Anzahl Übermittlungen und
 Anzahl Stämme in *snplist*

Abbildung 3.3: Ein Beispiel für das Auslesen der XML-Elemente **Individual** und **SnpInfo** durch `create.pl`.

Eintragung von Daten in die Tabellen „snplist“ und „snplist_details“

Beim Start des Elements `SnpInfo` wird versucht, den SNP-Eintrag für die entsprechende `rs_id` aus der `fasta` Tabelle zu holen, welches impliziert, dass zu dem Eintrag auch eine Sequenz vorhanden ist. Wenn dies nicht der Fall ist, wird der SNP-Eintrag auf `XX` gesetzt, um alle nachfolgenden Beschreibungen dieser `rs_id` ordnungsgemäß prozessieren zu können, letztendlich aber dem Frontend unzugänglich zu machen. Das Unterelement `SnpLoc` enthält Informationen zum Locus des SNPs (im Attribut `start`) sowie Angaben zur Verknüpfung mit Genen oder anderen interessanten Loci. Es kann mehrere Unterelemente dieser Bezeichnung pro `SnpInfo` geben, um die Kartierungen des Locus auf verschiedenen Kartenkonstrukten (z. B. von Celera [116] oder des Mouse Genome Sequencing Consortiums [204]) wiederzugeben, die Namen derer finden sich im Attribut `GenomicAssembly`. All diese Angaben werden intern zwischengespeichert.

Bestimmung des korrekten Locus durch „write_genomic_data“

Beim Auftreten des ersten Unterelements `SsInfo` (gleichzeitig sind damit alle `SnpLoc` Unterelemente abgearbeitet), wird das Unterprogramm `write_genomic_data` gestartet. Mit diesem Unterprogramm wird überprüft, ob ein oder mehrere `SnpLoc` Unterelemente für das `Snpinfo` Element existieren. Wenn nur eines existiert, werden die Angaben direkt in die dem Chromosom entsprechende Untertabelle von `snplist` geschrieben. Falls der Ausdruck `C57BL` nicht in dem Attribut `GenomicAssembly` vorkommt, bekommt der Eintrag den Status `no_map`. Bei Existenz von mehreren `SnpLoc` Unterelementen werden die `GenomicAssembly` Attribute herausgefiltert, in denen der Ausdruck `C57BL` zusammen mit einer gültigen Positionsangabe vorkommt. Wenn dieser Ausdruck nicht vorkommt, werden die Angaben des letzten

erfassten SnpLoc Unterelements in die Datenbank geschrieben und mit dem Zusatz `no_map` versehen. Wenn mehrere Einträge mit dem Ausdruck `C57BL` vorhanden sind, werden diese auf Konsistenz der `start` Attribute überprüft und die Daten des ersten SnpLoc Eintrages werden in die Datenbank geschrieben. Sollte eines der `start` Attribute abweichen, wird der Eintrag um den Status `no_map` ergänzt. Bei der Eintragung in die `snplist` wird eine `internal_id` vergeben, auf welche sich alle nachfolgenden Tabelleneinträge für den bearbeiteten SNP beziehen.

Qualitätskontrolle der Genotypdaten durch „encoding“

Beim Start jedes `SsInfo` Unterelements werden die Werte der `ssId` Attribute in die `submissions` Tabelle geschrieben und der Wert des Attributs `ssOrientToRs` (mit der relativen Orientierung zur Referenz-Sequenz) festgehalten.

Wenn das Unterelement `GTypeByInd` auftritt, wird das Unterprogramm `encoding` aufgerufen. In diesem wird der im Attribut enthaltene `gtype` Wert in Großbuchstaben umgewandelt und, je nach dem Wert für `ssOrientToRs`, in seine komplementäre Base übersetzt. Die beiden angegebenen Genotypen werden miteinander verglichen und deren Länge wird analysiert. Sollten die beiden Genotypen nicht identisch sein, wird `status_details` auf `het` gesetzt. Wenn der erste Genotyp länger als 1 ist, wird `status_details` auf `long` gesetzt. Sollte der Genotyp nicht in dem Eintrag in `fasta` vorkommen, wird der Status auf `incon` gesetzt. Eine Ausnahme bildet der `fasta` Eintrag „N“; dieser impliziert das Vorhandensein eines Indel. Wenn dies der Fall ist, wird der Status mit dem Eintrag `indel` ergänzt. Es wird überprüft, ob das Attribut `IndId` in der Tabelle `strains` vorhanden ist, also ob die ID des Mausstammes bereits bekannt ist. Wenn dies nicht der Fall ist, wird die ID des Stammes auf 0 gesetzt. Alle Einträge werden anschließend in die dem Chromosom zugehörige Untertabelle von `snplist_details` geschrieben.

Abschließende Eintragungen in die Tabellen „snplist“ und „snplist_details“

Die Anzahl der `SsIds` pro `internal_id` in der Tabelle `submissions` repräsentiert die Anzahl der Übermittlungen pro SNP. Die Anzahl der `internal_id_strains` repräsentiert die Anzahl der genotypisierten Individuen. Beide Werte werden in die Tabelle `snplist` geschrieben, wobei mehrfach vorkommende Einträge pro `internal_id_strains` nur einfach gezählt werden. Am Ende des Elements `SnpInfo` werden die neuen Einträge in der `snplist_details` Tabelle auf Konsistenz überprüft. Es wird also überprüft, ob zwei gleichlautende Mausstamm-IDs unterschiedliche Genotypen enthalten. Wenn dies der Fall ist, so wird die `status_details`-Spalte um den Eintrag `incon` ergänzt. Der Anteil R_{major} des am häufigsten vorkommenden Allels wird berechnet, indem die Frequenz F_{geno1} bestimmt wird. Wenn das Verhältnis unter 0,5 liegt, ist die Frequenz $F_{major} = 1 - F_{geno1}$, ansonsten ist $F_{major} = F_{geno1}$ (siehe Formel 3.1). Dieser Anteil wird bestimmt, indem der Genotyp jedes Mausstammes nur einmal gezählt wird und inkonsistente sowie heterozygote Einträge nicht berücksichtigt werden. Der erhaltene Anteil

3. ERGEBNISSE

wird in die zum Chromosom gehörende Untertabelle von `snplist` geschrieben. Zuletzt wird die `fasta` Tabelle um den `internal_id` Eintrag ergänzt.

Es bleibt festzuhalten, dass der so berechnete Anteil nicht der Allelfrequenz nach dem Hardy-Weinberg-Gesetz [69, 206] entspricht, da es sich bei den untersuchten Mäusen nicht um Populationen handelt, sondern um ingezüchtete Laborstämme. Der Aufbau der Tabellen aus den Rohdaten ist sehr zeitintensiv und kann ein bis zwei Tage auf einem durchschnittlichen Arbeitsplatzrechner in Anspruch nehmen.

Es sei

- N_{geno1} : Anzahl des ersten Genotyps eines SNP Locus
- N_{geno2} : Anzahl des zweiten Genotyps eines SNP Locus
- F_{geno1} : Anteil des ersten Genotyps

Berechnung des Anteils F_{major} des am häufigsten vorkommenden Allels

$$F_{geno1} = \frac{N_{geno1}}{N_{geno1} + N_{geno2}}$$

$$0 \leq F_{geno1} < 0,5 : F_{major} = 1 - F_{geno1}$$

$$F_{geno1} \geq 0,5 : F_{major} = F_{geno1}$$

(3.1)

3.1.2. Das Frontend `arts.cgi`

Die Eingabeoberfläche von ARTS wurde in Perl geschrieben und regelt Benutzerinteraktionen über den Browser. Es erlaubt dem Benutzer, die geeigneten Marker für ein individuelles Genotypisierungsprojekt zusammenzustellen und die Ergebnisse seiner Suche direkt weiterzuverarbeiten. In Abbildung 3.4 wird ein Screenshot der Benutzeroberfläche dargestellt.

Erstellung der temporären Haupttabelle anhand der Locusdaten und allgemeiner Eigenschaften

Benutzereingabe Der Benutzer kann die chromosomale Region angeben, in der er SNPs suchen möchte. Mit dem Optionsblock „Filter parameters for general SNP entries“ wird angegeben, ob SNPs, bei denen allgemeine Probleme festgestellt wurden, z. B. eine uneindeutige Kartenposition oder mehr als zwei Allele, ausgefiltert werden sollen.

Die allgemeinen Eigenschaften der SNPs in der gewählten Region (z. B. die Mindestanzahl der bekannten Stämme pro SNP) können ebenfalls angegeben werden.

3.1. ADVANCED RETRIEVAL TOOL FOR SNPs (ARTS)

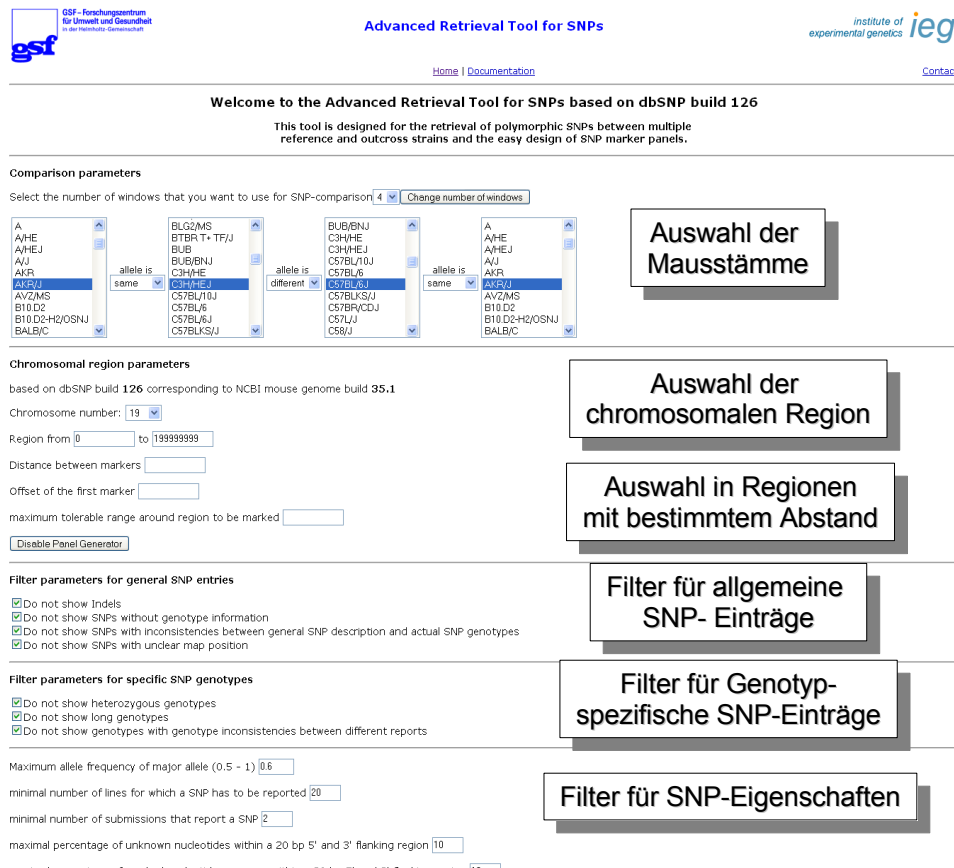


Abbildung 3.4: Die Benutzeroberfläche von ARTS bietet verschiedene Filter- und Eingabeoptionen zur Auswahl von geeigneten SNPs für ein Genotypisierungsprojekt.

Verarbeitung in arts.cgi: 'master' Mit dem `master` Unterprogramm wird eine temporäre `master` Tabelle generiert, in welcher das Ergebnis einer aus den Locusinformationen und den allgemeinen Filterparametern zusammengesetzten SQL Abfrage gespeichert wird. In der `master` Tabelle befinden sich die `internal_ids` der in Frage kommenden SNPs sowie deren Locus.

Reduktion der Locusdaten auf bestimmte Bereiche durch die Option „Panel Generator“

Benutzereingabe (optional) Mit dem Panel Generator können SNPs in frei definierbaren Regionen gefunden werden, die sowohl einen bestimmten Abstand zueinander als auch einen Versatz relativ zum Start der chromosomalen Region haben (vgl. Formel 3.2).

Verarbeitung in arts.cgi: 'panel_generator' Wenn die „Panel Generator“ Option aktiviert wurde, wird die Gültigkeit der für diese Funktion notwendigen Parameter überprüft und die `master` Tabelle auf die Loci der Markerregionen verkürzt. Die Formel 3.2 enthält die zugrunde liegenden Berechnungen. Die errechneten chromosomalen Start- und Zielpositionen werden intern zur späteren Ausgabe vorgehalten.

Es sei

p_{start} : Start der zu untersuchenden chromosomalen Region

p_{end} : Ende der zu untersuchenden Region

o : Versatz der ersten Markerregion relativ zu p_{start}

d : Abstand zwischen den einzelnen Markerregionen

t : Proximale/distale Begrenzung der Markerregion

n : Anzahl der Markerpositionen

Berechnung der Position des ersten Markers

$$m_1 = p_{start} + o$$

Berechnung der Anzahl der Markerpositionen zwischen p_{start} und p_{end}

$$p_{start}, p_{end}, o, d, t, n \in \mathbb{N}$$

$$n = \frac{p_{end} - m_1}{d}$$

Berechnung der chromosomalen Anfangs- und Endposition jeder Markerregion

$$start_{m_1} \cdots start_{m_n} = m_1 + nd - t$$

$$end_{m_1} \cdots end_{m_n} = m_1 + nd + t$$

(3.2)

Erstellung von Mausstamm-spezifischen temporären Tabellen

Benutzereingabe Bis zu sechs Auswahlfenster für Mausstämme können geöffnet und deren Beziehung zueinander (gleiches Alleles, unterschiedliches Allel) kann definiert werden. Es ist möglich, mehrere Mausstämme pro Fenster auszuwählen, um die Trefferwahrscheinlichkeit zu erhöhen. Dies empfiehlt sich insbesondere, wenn für den eigentlich gesuchten Mausstamm nur wenig Information vorhanden ist, aber sehr viele Genotypen für einen nahe verwandten Mausstamm bekannt sind.

Der Genotyp-spezifische Optionsblock weist die Suchmaschine an, nur solche Genotypen in Betracht zu ziehen, für die keine Inkonsistenzen oder Heterozygotien durch `create.pl` gefunden wurden.

Verarbeitung in arts.cgi: 'retrieve_data' Mit dem Unterprogramm `retrieve_data` werden die Genotypen der von der Tabelle `master` vorgegebenen `internal_ids` pro Auswahlfenster gesammelt und in eigene temporäre Tabellen überführt. Hierbei werden Filterparameter aus Optionsblock „Filter parameters for specific SNP genotypes“ in der dazugehörigen SQL Abfrage berücksichtigt.

Vergleich der Genotypdaten der temporären Tabellen

Verarbeitung in arts.cgi : 'comparison' Das Unterprogramm `comparison` übernimmt den Vergleich dieser Tabellen. Zunächst wird die Position des Ungleichheitszeichens bestimmt und die Genotypdaten in den Tabellen links und rechts davon jeweils auf Gleichheit überprüft und in eine neue temporäre Tabelle geschrieben. Letztendlich bleiben also zwei Tabellen übrig, die auf Ungleichheit getestet werden. Das Ergebnis dieses Vergleichs wird in einer temporären Tabelle abgelegt, welche die `internal_ids` und den gemeinsamen Genotyp der Stämme links vom Ungleichheitszeichen sowie den Genotyp der Stämme rechts davon enthält.

Ausgabe der qualifizierten SNPs

Verarbeitung durch arts.cgi: 'output' Die Ausgabe der Daten auf den Browser erfolgt durch das Unterprogramm `output`. Die `rs_id`, die chromosomale Position, allgemeine Informationen sowie die beiden Genotypen und die am 3' und 5' Ende flankierenden Sequenzen werden für jede `internal_id` aus den Tabelle `snplist` und `fasta` gesammelt und ausgegeben. Der maximale Anteil an unbekanntem und maskierten Nukleotiden in den flankierenden Sequenzen wird in diesem Schritt bestimmt und SNPs mit Sequenzen, die die vorgegebenen Kriterien nicht erfüllen, werden nicht ausgegeben.

Es gilt zu beachten, dass die Reihenfolge der beiden Genotypen nicht, wie bei anderen Suchmaschinen üblich, in einer ungeordneten Weise erscheinen, sondern gemäß ihrer Stellung relativ zum Ungleichheitszeichen ausgegeben werden. Dies erleichtert die Verständlichkeit und Weiterverarbeitung der erhaltenen Daten, da ein ansonsten notwendiger zweiter Zuordnungsschritt zwischen Mausstamm und Genotyp entfallen kann.

Ausgabe der endgültigen Tabelle

Benutzereingabe Die am besten geeigneten Marker werden vom Benutzer markiert. Zusätzlich können Angaben zur Länge der zu exportierenden flankierenden Sequenzen und zum Export weiterführender Informationen gemacht werden.

Verarbeitung durch arts.cgi: 'final' Die Verarbeitung der Eingaben und die Ausgabe der Daten erfolgt durch das Unterprogramm `final`. Die Daten werden aus den Tabellen `snplist` und `fasta` ausgelesen und die gewünschte Information im Browser aufgelistet.

Diese kann durch Kopieren und Einfügen weiterverarbeitet werden. Zusätzlich wird durch das Unterprogramm `writereport` eine mit Tabellenkalkulationsprogrammen kompatible Datei erstellt, die heruntergeladen werden kann. Diese Datei kann auch direkt durch das Programm `MassARRAY AssayDesign` (vgl. Tabelle 2.22) weiterverarbeitet werden.

3.2. Eigenschaften des Markersatzes für die hME Methode

Vor dem Umstieg auf die iPlex Methode (vgl. Abschnitt 2.3) wurde ein Markersatz verwendet, der hauptsächlich für die Genotypisierung der im MEP eingesetzten Standardkreuzung zwischen den Mausstämmen C3HeB/FeJ und C57BL/6J geeignet war. Ungefähr die Hälfte der Marker wurde so gewählt, dass auch Kreuzungen zwischen dem Stamm C3HeB/FeJ und BALB/cByJ mit ausreichendem Informationsgehalt genotypisiert werden konnten (vgl. Abschnitt 2.4.1).

Der Großteil der Marker wurde entweder ohne oder mit einer frühen Vorversion der ARTS Suchmaschine (vgl. Abschnitt 3.1) generiert und die Auswertung erfolgte ohne die komplett entwickelten MyGenotype Programme (vgl. Abschnitt 3.4), daher ist die Heterozygotie der Marker relativ niedrig (vgl. Tabelle 3.2) und die chromosomale Verteilung der Marker nicht immer optimal (vgl. Abbildung 3.5). Dieser Markersatz ist jedoch durchaus für eine Kartierung von ENU Mutanten geeignet und wurde zur Kartierung der in den Abschnitten 3.7 und 3.8, sowie für die Kartierung des überwiegenden Teils der in Abschnitt 3.6 aufgelisteten Mauslinien benutzt.

3.2.1. Chromosomale Verteilung und Polymorphie der Marker

Der Markersatz wurde für alle 19 Autosomen erstellt, da ausschließlich Linien mit einem geschlechtsunabhängigen Phänotyp untersucht wurden. Die Genotypen von verschiedenen genetisch recht weit voneinander entfernten Mausstämmen wurden für 126 Marker bestimmt. Die potenzielle Eignung des Markersatzes für Kreuzungen, die nicht bei der Erstellung berücksichtigt wurden, sollte so bestimmt werden. Abbildung 3.5 enthält die chromosomalen Positionen der Marker bezogen auf Maus Genome Assembly 35 sowie den Anteil des seltener auftretenden Allels jedes Markers.

Der durchschnittliche Abstand der Marker zueinander war 17,44 Mb, der durchschnittliche Anteil des seltenen Allels lag bei 0,34. Im weiteren Verlauf der Optimierung des Markersatzes wurde der durchschnittliche Abstand auf 16,4 Mb verkürzt und insbesondere die Markerverteilung auf Chromosom 4, 14 und 17 verbessert. In Anhang A ist der endgültige Markersatz gelistet.

Tabelle 3.1: Genotypisierte Mausstämme

Bezeichnung	Genotypisierungsmethode		
	hME	iPlex	<i>in silico</i>
C3HeB/FeJ	✓	✓	(✓)*
C57BL/6J	✓	✓	✓
BALB/cByJ	✓	✓	✓
AKR/J		✓	✓
102	✓	✓	
129S1/SvImJ			✓
129/Sv	✓	✓	
129X1/SvJ			✓
A/J			✓
BTBR/TfJ			✓
C3H/HeJ		✓	✓*
CAST/EiJ		✓	✓
CBA	✓	✓	
DBA/2J	✓	✓	✓
FVB		✓	✓
GRS	✓	✓	
JF1	✓	✓	
KK/HIJ			✓
MOLF/EiJ			✓
NOD/LtJ			✓
NZW/LacJ			✓
PWD/PhJ			✓
WSB/EiJ			✓
SEC	✓	✓	
SWR	✓	✓	

Zur *in silico* Vorhersage wurden alle Mausstämme mit mehr als einer Million Einträgen in der dbSNP benutzt. *: In der dbSNP sind keine Daten für den Stamm C3HeB/FeJ vorhanden, es wurde daher der Stamm C3H/HeJ benutzt.

3.2. EIGENSCHAFTEN DES MARKERSATZES FÜR DIE HME METHODE

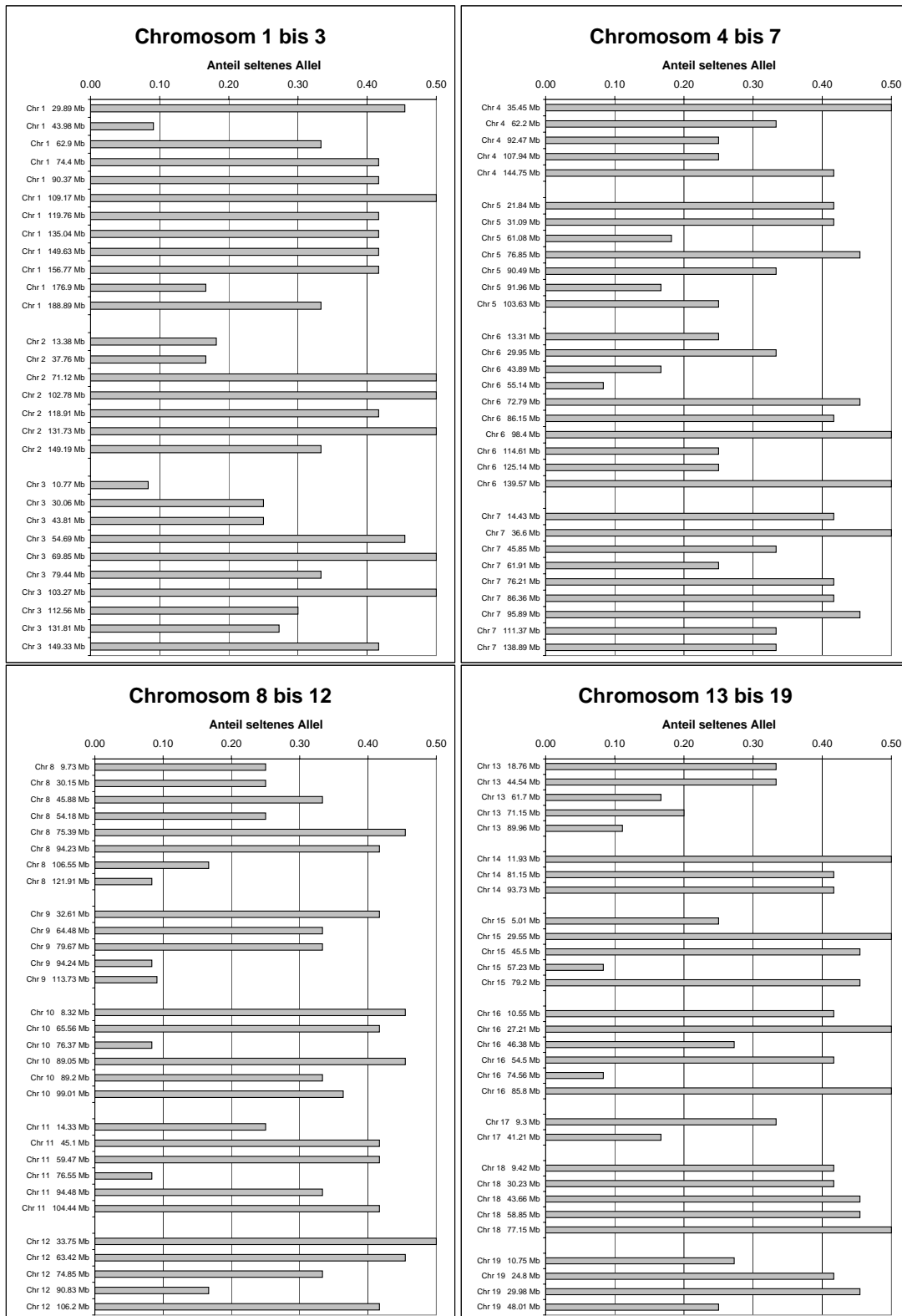


Abbildung 3.5: Chromosomale Verteilung der Marker für die hME Methode. Dargestellt ist der Locus der gewählten Marker sowie der Anteil des seltenen Allels für die in Tabelle 3.1 aufgelisteten Mausstämmen.

3. ERGEBNISSE

3.2.2. Heterozygotie des Markersatzes für Kreuzungen anderer Mausstämme

In Tabelle 3.2 wird die Heterozygotie für Kreuzungen der in Tabelle 3.1 aufgelisteten Mausstämme zusammenfassend dargestellt. Es ergibt sich wie erwartet eine volle Heterozygotie der Marker für die Mausstämme C3HeB/FeJ und C57BL/6J, sowie eine Heterozygotie für 62% aller Marker zwischen C3HeB/FeJ und BALB/cByJ. Desweiteren würden über 60% der Marker eine Heterozygotie bei Kreuzungen zwischen C57BL/6J und den Mausstämmen 102, CBA und DBA/2J zeigen, sowie bei einer Kreuzung zwischen BALB/cByJ und dem Stamm 102.

Tabelle 3.2: Heterozygotie des hME Markersatzes zwischen verschiedenen Mausstämmen

	C3HeB/FeJ	C57BL/6J	BALB/cByJ	102	129/Sv	CBA	DBA/2J	GRS	JF1	SEC	SWR
C3HeB/FeJ	-	100	62	27	45	17	39	44	43	54	51
C57BL/6J	132	-	37	72	54	82	60	55	56	45	48
BALB/cByJ	132	132	-	61	49	53	46	36	53	10	35
102	128	128	128	-	32	33	42	46	47	53	46
129/Sv	121	121	121	121	-	41	42	44	49	46	42
CBA	128	128	128	128	121	-	35	46	40	50	48
DBA/2J	128	128	128	128	121	128	-	43	41	39	44
GRS	128	128	128	128	121	128	128	-	49	35	14
JF1	117	117	117	117	111	117	117	117	-	51	58
SEC	128	128	128	128	121	128	128	128	117	-	32
SWR	128	128	128	128	121	128	128	128	117	128	-

Werte rechts oben: Prozentualer Anteil der Heterozygotie pro Kreuzung. Werte links unten: Anzahl der ausgewerten Marker pro Kreuzung. Grau unterlegt: Kreuzungen mit einer Heterozygotie über 60%.

3.3. Eigenschaften des Markersatzes für die iPlex Methode

Die iPlex Methode hat den Vorteil, dass wesentlich mehr Marker pro Platte genotypisiert werden können. Dadurch wird die Genotypisierung kostengünstiger, der Arbeitsaufwand geringer und die Auswertung erleichtert.

Es wurden Marker gewählt, die nicht nur zwischen den Stämmen C3HeB/FeJ und C57BL/6J Heterozygotie aufweisen sollten, denn zusätzlich sollte die Gesamtheit der gewählten Marker eine Heterozygotie von über 70% bei C3HeB/FeJ×BALB/cByJ und C57BL/6J×AKR/J Kreuzungen bieten (vgl. Abschnitt 2.4.2).

3.3.1. Chromosomale Verteilung und Polymorphie der Marker

Auch dieser Markersatz wurde so erstellt, dass er alle 19 Autosomen abdeckt, da ausschließlich Linien mit einem nachweislich geschlechtsunabhängigen Phänotyp untersucht wurden. Die in Tabelle 3.1 gezeigten Mausstämme wurden mit 156 Markern genotypisiert, um die voraussichtliche Eignung für

Kreuzungen anderer Stämme zu bestimmen. Die chromosomale Verteilung der Marker und den Anteil des seltener vorkommenden Allels werden in Abbildung 3.6 dargestellt.

Der durchschnittliche Abstand der Marker zueinander war 15,07 Mb und der durchschnittliche Anteil des seltenen Allels lag bei 0,39. Zusätzlich wurde auch eine Probe des Mausstamms C3H/HeJ genotypisiert. Es wurde keine Heterozygotie festgestellt, daher kann angenommen werden, dass Marker für den Mausstamm C3H/HeJ mit einer sehr hohen Wahrscheinlichkeit auch für den Stamm C3HeB/FeJ geeignet sind. Die Ergebnisse für den Mausstamm C3H/HeJ wurden nicht bei der Erstellung der Abbildung 3.6 berücksichtigt, weil die Häufigkeit des seltenen Allels dadurch verfälscht worden wäre.

3.3.2. Heterozygotie des Markersatzes für Kreuzungen anderer Mausstämme

Die Heterozygotie für Kreuzungen der in Tabelle 3.1 aufgelisteten Mausstämme wird in Tabelle 3.3 zusammenfassend dargestellt. Die an den Markersatz gestellten Bedingungen für die Heterozygotie wurden erfüllt. Der Markersatz weist eine Heterozygotie von 100% für C3HeB/FeJ×C57BL/6J Kreuzungen auf und besitzt eine Heterozygotie von 82% für Kreuzungen der Stamme C57BL/6J und AKR/J. Kreuzungen zwischen C3HeB/FeJ und BALB/cByJ können mit 73% der Marker ausgewertet werden. Als Nebeneffekt der Auswahl von SNPs mit hoher Heterozygotie ist der Markersatz für eine Reihe anderer Kreuzungen zu über 60% heterozygot. Desweiteren wurde die Heterozygotie zwischen den in Tabelle 3.1 aufgelisteten Mausstämmen vorhergesagt. Diese Mausstämme haben mehr als eine Million Einträge in der dbSNP, wie eine Abfrage der `snplist_details` Tabelle in ARTS ergab. In Tabelle 3.4 sind die vorgesagten Heterozygotien dargestellt. Man beachte, dass es nur geringe Abweichungen zwischen der vorhergesagten und der experimentell ermittelten Heterozygotie gibt.

Da bei der Erstellung dieses Markersatzes auf die neueste Version der dbSNP mit knapp 47 Millionen Einträgen zurückgegriffen werden konnte, wurden hauptsächlich nur solche SNPs verwendet, für die die Genotypen von mindestens 20 Mausstämmen bereits bekannt waren. Daher kann eine gute *in silico* Vorhersage über die Heterozygotie des Markersatzes für Kreuzungen verschiedene Stämme getroffen werden.

3. ERGEBNISSE

Tabelle 3.3: Heterozygotie des iPlex Markersatzes zwischen verschiedenen Mausstämmen

	C3HeB/FeJ	C57BL/6J	BALB/cByJ	AKR/J	102	129/Sv	CAST/EiJ	CBA	DBA/2J	FVB	GRS	JF1	SEC	SWR
C3HeB/FeJ	-	100	73	14	22	38	42	19	32	38	44	48	64	43
C57BL/6J	156	-	26	85	77	61	57	80	67	61	55	51	35	56
BALB/cByJ	155	155	-	61	67	57	58	62	57	54	50	57	15	50
AKR/J	156	156	155	-	32	44	42	25	33	41	45	47	58	43
102	154	154	153	154	-	26	49	35	42	39	41	49	57	42
129/Sv	156	156	155	156	154	-	51	39	55	46	45	48	51	48
CAST/EiJ	152	152	151	152	150	152	-	41	47	53	54	16	53	54
CBA	155	155	154	155	153	155	151	-	36	42	49	50	61	45
DBA/2J	156	156	155	156	154	156	152	155	-	43	49	52	51	44
FVB	155	155	155	155	153	155	151	154	155	-	25	62	47	22
GRS	155	155	155	155	153	155	151	154	155	155	-	60	43	23
JF1	146	146	145	146	144	146	145	145	146	145	145	-	53	62
SEC	153	153	153	153	151	153	149	152	153	153	153	143	-	45
SWR	154	154	154	154	152	154	151	153	154	154	154	145	152	-

Werte links unten: Anzahl der ausgewerten Marker pro Kreuzung. Grau unterlegt: Kreuzungen mit einer Heterozygotie über 60%.

Tabelle 3.4: Vorhergesagte Heterozygotie des iPlex Markersatzes für Mausstämmen mit mehr als einer Million Genotypen in der dbSNP 126

	C3H/HeJ	C57BL/6J	BALB/cByJ	AKR/J	129S1/SvImJ	129X1/SvJ	A/J	BTBRTf/J	CAST/EiJ	DBA/2J	FVB/NJ	KK/HJ	MOLF/EiJ	NOD/Lj	NZW/LacJ	PWD/PhJ	WSB/EiJ
C3H/HeJ	-	100	75	11	40	44	35	47	48	32	41	47	49	44	48	54	43
C57BL/6J	157	-	25	88	61	55	64	54	60	67	60	57	60	57	56	60	56
BALB/cByJ	156	156	-	63	58	52	42	49	61	56	54	58	62	45	59	60	59
AKR/J	156	156	155	-	45	50	34	47	48	33	44	47	48	42	50	52	45
129S1/SvImJ	146	146	146	146	-	8	61	30	58	59	49	51	50	54	46	68	53
129X1/SvJ	144	144	144	144	136	-	59	26	56	60	50	54	53	57	49	70	54
A/J	155	155	154	155	145	143	-	51	51	50	40	55	56	36	53	52	47
BTBRTf/J	144	144	144	144	141	135	143	-	54	52	49	47	53	48	47	76	51
CAST/EiJ	131	131	131	131	127	122	130	128	-	50	62	58	26	56	51	33	45
DBA/2J	158	158	157	157	149	145	156	146	134	-	45	45	54	40	44	70	50
FVB/NJ	144	144	144	144	142	135	143	141	127	147	-	46	67	32	40	74	52
KK/HJ	146	146	146	146	143	137	145	143	129	148	143	-	56	47	45	62	52
MOLF/EiJ	138	138	138	138	135	129	137	137	130	141	135	137	-	64	54	35	47
NOD/Lj	145	145	145	145	142	136	144	143	128	147	142	144	137	-	48	76	55
NZW/LacJ	146	146	146	146	143	137	146	143	130	148	142	144	137	143	-	69	47
PWD/PhJ	51	51	51	51	51	44	51	51	53	54	51	51	53	51	52	-	53
WSB/EiJ	141	141	141	141	139	132	140	139	126	143	139	140	134	139	140	52	-

Werte rechts oben: Prozentualer Anteil der Heterozygotie pro Kreuzung. Werte links unten: Anzahl der ausgewerten Marker pro Kreuzung. Grau unterlegt: Kreuzungen mit einer Heterozygotie über 60%.

3.3. EIGENSCHAFTEN DES MARKERSATZES FÜR DIE IPLEX METHODE

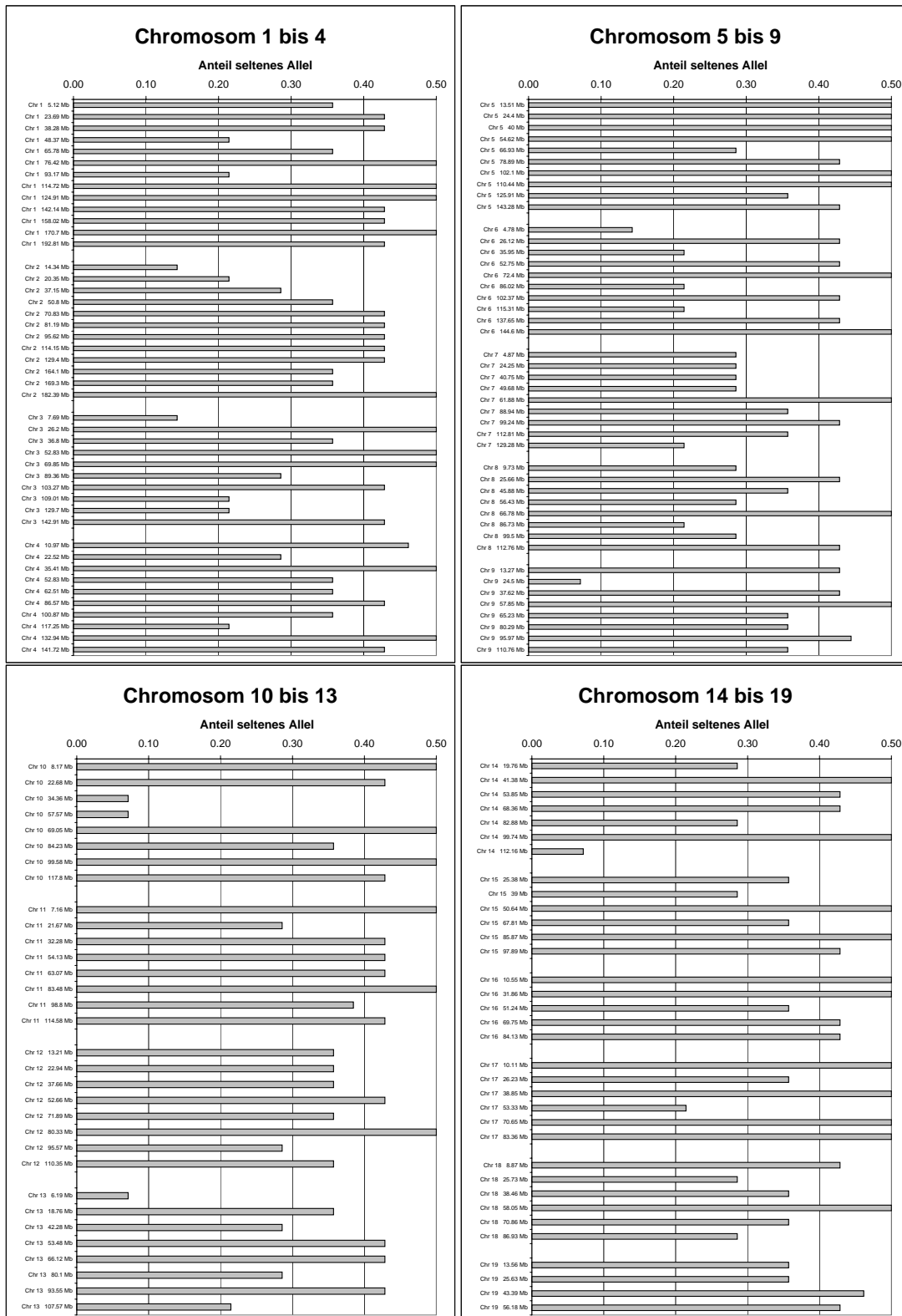


Abbildung 3.6: Chromosomale Verteilung der Marker für die iPlex Methode. Dargestellt ist der Locus der gewählten Marker sowie der Anteil des seltenen Allels für die in Tabelle 3.1 aufgelisteten Mausstämmen.

3. ERGEBNISSE

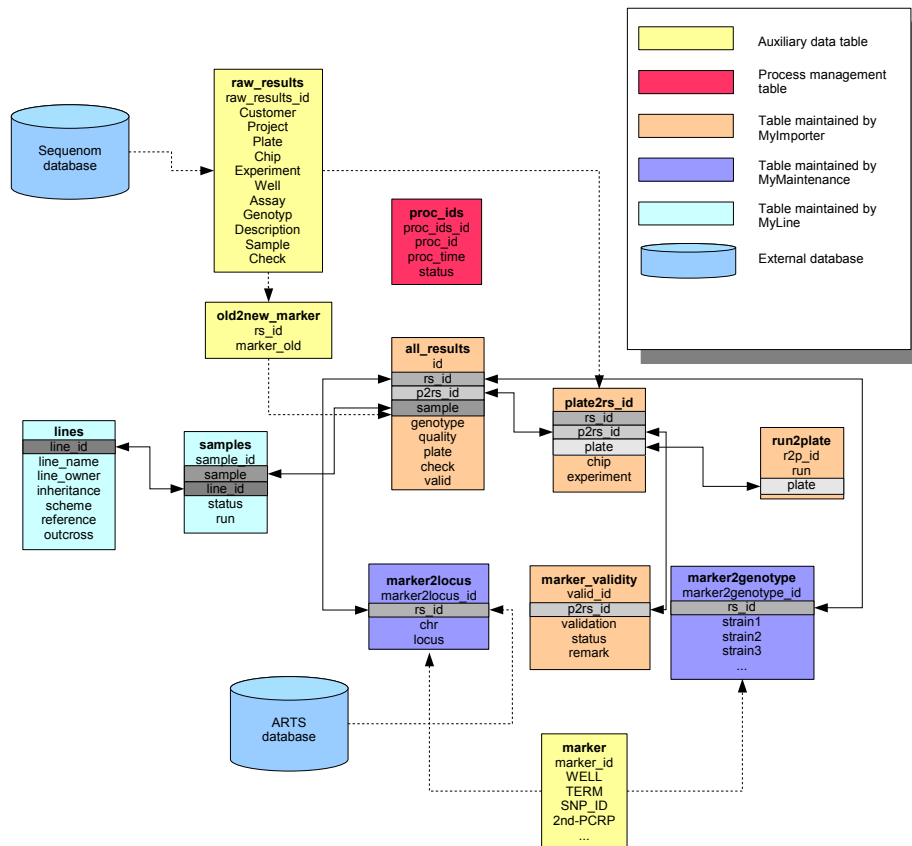


Abbildung 3.7: Schema der MyGenotype Datenbank. Dargestellt ist die Verknüpfung der Tabellen und ihre Interaktion mit den verschiedenen MyGenotype Programmen.

3.4. Die Erstellung der MyGenotype Datenbank

Bei der genomweiten Genotypisierung werden über 60.000 Genotypdaten für 384 Proben ermittelt. Diese Proben müssen den ursprünglichen Mauslinien zugeordnet werden und es muss definiert werden, ob die Probe von einer Mutante oder einem Wildtypstier stammt. Desweiteren muss der ermittelte Genotyp den parentalen Stämmen wieder zugeordnet werden können. Der Import, die Zuordnung und die Auswertung der Daten sollte möglichst automatisiert vonstatten gehen.

Es wurde eine MySQL-Datenbank und eine Reihe von Perl Skripten entwickelt, welche diese Anforderungen erfüllen und eine möglichst komfortable und reproduzierbare Analyse der Genotypisierungsdaten ermöglichen. Die Datenbank hat das in Abbildung 3.7 dargestellte Schema. Alle Perl Skripte greifen mittels des Perl Moduls DBI auf diese Datenbank zu.

3.4.1. Das Programm MyImporter

Importfunktion Die von dem MALDI-TOF MS System erzeugten Daten werden in einer Oracle-Datenbank Version 8i gespeichert. Der Zugriff auf die Daten erfolgt entweder über das Genotyping Frontend von Sequenom, in dem man die Ergebnisse einzelner Platten exportieren kann, oder über ein von Tim Strom (Institut für Humangenetik, GSF Neuherberg) geschriebenes Perl Skript zum Export einer Serie von Genotypisierungsexperimenten (im Folgenden Lauf genannt). MyImporter steuert dieses Skript über die Perl Module LWP::UserAgent und HTML::Form fern und verarbeitet anschließend die für die Mausgenotypisierung relevanten Daten. Der Benutzer kann den zu importierenden Lauf auswählen.

```

matthias@kody:~/mapping - Befehlsfenster 4 - Konsole
Sitzung Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe
M00073_IT00601 3683804 (old ID: W1_MGS_3_30060930)

Genotype distribution
9
A 125
AG 150
G 100
Chi square = 0.30, P = 5.863400E-01

Quality distribution
A 371
B 4
N 1
97.66 % A or B results

Positive H2O Controls: 0.00 %
chi square test looking good
number of automatically called results looking good
H2O controls looking good

More than 90% A or B results and Chi square is okay, automatically validated
  
```

Abbildung 3.8: Bildschirmfoto der MyImporter Validierungsroutine

Es wird überprüft, ob bereits Daten des zu importierenden Laufs vorhanden sind und wenn ja, ob der Benutzer wünscht, dass diese überschrieben werden sollen. In diesem Fall werden bereits vorhandene `sample→line→phenotype` Assoziationen in der temporären Tabelle `samples2lines_temp` zwischengespeichert.

Die anschließend empfangenen Daten werden in ein Format gebracht, welches gut von MySQL eingelesen werden kann. Die Rohdaten werden in der Datenbank abgelegt. Ein aufbereiteter Datensatz mit den wichtigsten Informationen wird in einer temporären Tabelle vorgehalten. Anhand dieser Daten wird die Tabelle `plate2rs_id` (enthält alle Angaben zu Assays und Experimenten und generiert pro Experiment den eindeutigen Bezeichner `p2rs_id`) sowie `run2plate` (enthält alle Plattenbezeichnungen pro Lauf) aktualisiert. Ein Struktogramm ist in Abbildung 3.9 dargestellt.

Validierungsfunktion Aus dieser temporären Tabelle werden alle Assays ausgelesen. Da das erwartete Verhältnis von homozygoten zu heterozygoten Ergebnissen, abhängig von den Kreuzungsschemata der zu untersuchenden Kreuzungen, konstant ist, kann das Verhältnis dieser Ergebnisse zur Qualitätskontrolle genutzt werden. Der Anteil R_{het} der heterozygoten Ergebnisse an der Gesamtzahl aller Ergebnisse eines Laufs wird bestimmt wie in Formel 3.3 dargestellt.

3. ERGEBNISSE

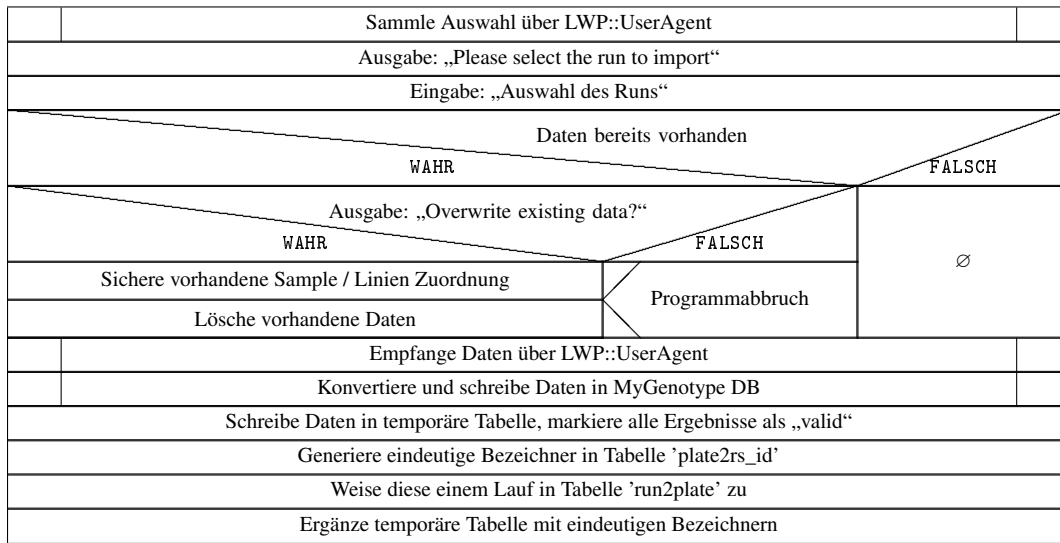


Abbildung 3.9: Die Importfunktion von MyImporter

Es sei

exp_i : Experiment eines Laufs

n : Gesamtzahl der Experimente eines Laufs

het : Anzahl der heterozygoten Ergebnisse pro Experiment

hom : Anzahl der homozygoten Ergebnisse pro Experiment

Berechnung des Anteils R_{het} der heterozygoten Ergebnisse eines Laufs

$$R_{het} = \frac{\sum_{i=1}^n \frac{het}{exp_i}}{\sum_{i=1}^n \frac{het+hom}{exp_i}} \quad (3.3)$$

Die Assays werden der Reihe nach auf Duplikate untersucht und anschließend validiert. Pro Assay kann mehr als ein Experiment vorhanden sein, wenn aus technischen Gründen ein Experiment wiederholt werden musste.

Die Bestimmung erfolgt über den Test, ob mehr als eine p2rs_id mit der zu testenden rs_id des Assays verknüpft ist. Wenn dies der Fall ist, dann existiert mehr als ein Experiment für diesen Assay.

Alle Experimente des als doppelt identifizierten Assays werden in der Tabelle marker_validity auf invalid, duplicate gesetzt. Um inkonsistente Sample → Genotyp Verknüpfungen zu vermeiden, kann der Benutzer die Ergebnisse nur eines Experiments wieder für gültig erklären. Falls kein Experiment vom Benutzer wieder für gültig erklärt wird, wird keine weitere Validitätsprüfung durchgeführt und die Tabelle marker_validity für die betroffenen p2rs_ids nicht mehr geändert.

Wenn ein Experiment ausgewählt wurde, oder keine Duplikate vorhanden sind, wird versucht, die Validität des zugehörigen Experiments nach den unten genannten Gesichtspunkten automatisch zu bestimmen (vgl. Abbildung 3.8). Die Voraussetzungen für eine automatische Validierung sind wie folgt definiert.

1. Die Anzahl der automatisch von der SpectroTYPER RT Software mit einer Vertrauensnote A oder B bestimmten Genotypen ist größer oder gleich 90% der Gesamtergebnisse eines Assays abzüglich H₂O Kontrollen.
2. Das Verhältnis der heterozygoten Genotypen zu allen Ergebnissen eines Assays darf nicht signifikant vom Durchschnitt R_{het} aller Assays abweichen. Dies wird mittels eines χ^2 Tests mit einem Freiheitsgrad bestimmt. Die Hypothese H_0 besagt, dass der Anteil der Heterozygoten des zu prüfenden Assays dem durchschnittlichen Anteil der Heterozygoten aller Assays dieses Laufes folgt (vgl. Formel 3.4). Sie wird abgelehnt, wenn die Wahrscheinlichkeit für ihre Bestätigung unter 0,001 liegt. Die Wahrscheinlichkeit wird mit dem Perl Modul `Statistics::Distributions` berechnet.
3. Die Anzahl der positiven H₂O Kontrollen darf nicht über 25% liegen. Dieser relativ großzügige Wert kann gewählt werden, weil es gelegentlich zu einer unspezifischen Extension eines Extensionsprimers kommen kann, wenn kein weiteres Ziel vorhanden ist. Im Zusammenhang mit einem unauffälligen χ^2 Testergebnis kann jedoch ausgeschlossen werden, dass die Genotypisierung durch eine unspezifische Extension verfälscht worden ist.

Es sei

het_{obs} : Anzahl der heterozygoten Ergebnisse in einem einzelnen Experiment

R_{het} : Anteil der heterozygoten Ergebnisse in einem Lauf

Berechnung des erwarteten Anteils het_{exp} an heterozygoten Ergebnissen

$$het_{exp} = R_{het} \cdot \left(\frac{het_{obs}}{assay} + \frac{het_{obs}}{assay} \right)$$

Annahme: Anzahl der Heterozygoten eines einzelnen Experiments folgt der durchschnittlichen Anzahl der Heterozygoten eines Laufs

Hypothese H_0 : $het_{obs} = het_{exp}$

Test auf Hypothese H_0

$$\chi^2 = 2 \cdot \frac{(het_{obs} - het_{exp})^2}{het_{exp}} \tag{3.4}$$

Sollte eines der Kriterien nicht zutreffen, kann der Benutzer den fraglichen Assay als `valid`, `problems` oder `invalid` kennzeichnen und die Gründe für die Entscheidung angeben, falls `problems` und `invalid`

3. ERGEBNISSE

ausgewählt wurde. Diese Werte werden in die Tabelle `marker_validity` eingetragen.

Nach der Validierung jedes Assays wird die Gültigkeit der Ergebnisse von Proben mit einer ID, die nicht 8-stellig ist, als `control` gekennzeichnet. Proben mit einer 8-stelligen ID werden in die Tabelle `samples` aufgenommen. Falls die Tabelle `samples2lines_temp` existiert, werden die Werte hieraus wieder in die `samples` Tabelle zurückgeschrieben. In Abbildung 3.10 werden die Vorgänge als Struktogramm dargestellt.

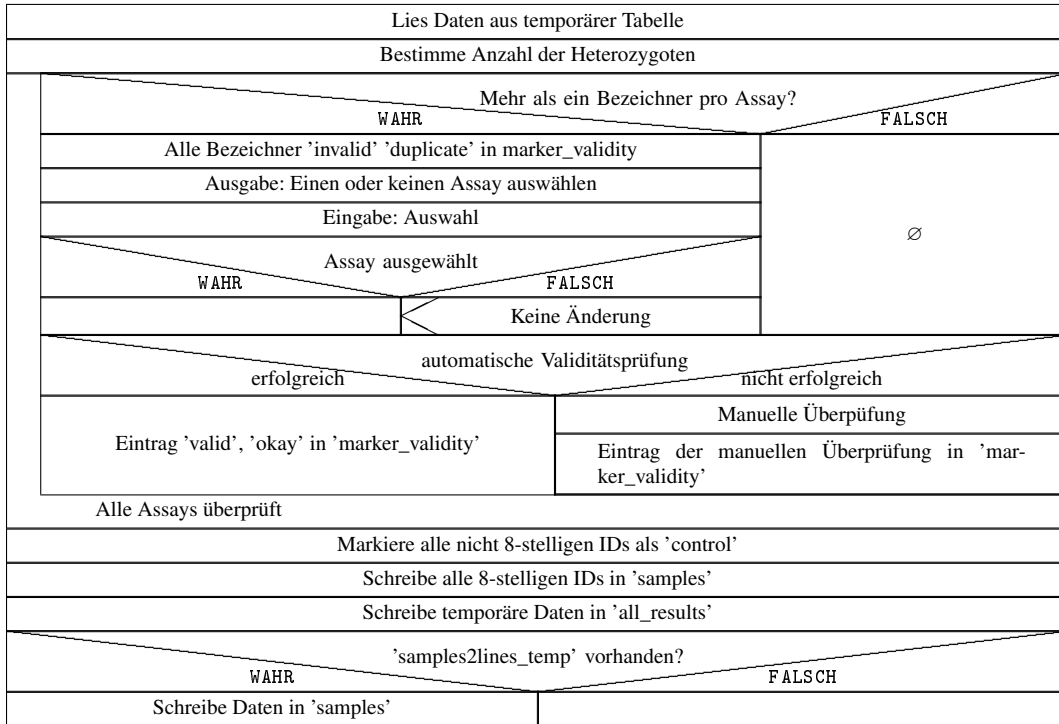


Abbildung 3.10: Die Validierungsfunktion von MyImporter

3.4.2. Das Programm MyMaintenance

Zur Vereinfachung häufig anfallender Abfragen und Aktualisierungen der MyGenotype Datenbank wurde das Perl-Skript `MyMaintenance` erstellt (vgl. Abbildung 3.11). Dieses Programm wird in einer Linux Kommandozeile aufgerufen.

Der Aufbau der Menüs wurde im Quellcode so weit wie möglich modularisiert. Auf diese Weise kann das Skript auf einfache Weise durch neue SQL Abfragen und durch Ein- und Ausgaben von Dateien erweitert werden.

Momentan werden folgende Funktionen unterstützt:

1. Eine Sicherungskopie der kompletten Datenbank kann erstellt werden. Hierbei wird auf das Programm `mysqlhotcopy` zurückgegriffen.

2. Assaylisten können importiert und exportiert werden. Die Listen entsprechen dem MassARRAY Assay Design 3.0 Format.
3. Marker IDs können mittels der ARTS Datenbank einem Locus zugeordnet werden.
4. Genotypen von neuen Mausstämmen können importiert werden. Dieser Vorgang ist komplex und wird deshalb in einem eigenen Unterpunkt und in Abbildung 3.12 behandelt.
5. Die Tabelle `marker_validity` kann exportiert werden.

Die Mausstamm / Genotyp Zuordnung

Ein Genotypisierungsexperiment kann die ID einer Probe einen neuen Mausstamm repräsentieren. Dieser kann als Referenzstamm für zukünftige Untersuchungen dienen. Nach Eingabe der ID, der Nomenklatur des Mausstamms und der Nummer des Experiments wird eine Plausibilitätsprüfung durchgeführt, die Genotypen aller Assays temporär in `temp_geno` geschrieben, und auf Genotyp-Inkonsistenzen innerhalb der Assays untersucht. Inkonsistente Genotypen werden temporär nach `temp_own_incon` geschrieben und aus `temp_geno` entfernt.

Dann wird überprüft, ob es sich um einen neuen Mausstamm oder einen existierenden Mausstamm handelt. Sollte es sich um einen neuen Mausstamm handeln, wird eine neue Tabellenspalte mit dem Namen des Stammes in `marker2genotype` angelegt.

Falls bereits Daten für den angegebenen Mausstamm existieren, werden die neuen Genotypen mit den bereits vorhandenen Genotypen verglichen. Inkonsistente Ergebnisse werden temporär in `temp_incon` abgelegt. Alle alten Ergebnisse werden aus `marker2genotype` entfernt. Der Benutzer wird aufgefordert, für alle Einträge in `temp_incon` den richtigen Genotyp anzugeben oder die Ergebnisse zu verwerfen. Nach erfolgter Prüfung wird entweder die Tabellenspalte eines existierenden Mausstammes mit den Werten aus `temp_geno` aktualisiert oder eine neue Tabellenspalte angelegt. Das Struktogramm in Abbildung 3.12 zeigt den schematischen Ablauf.

3.4.3. Das Programm MyLine

Der Administrator kann über dieses in Perl geschriebene Browserinterface die 8-stelligen Proben-ID eines Genotypisierungsexperiments einer Mauslinie zuordnen. Durch das Einfügen der IDs in die entsprechenden Felder kann er angeben, ob die Proben Mutanten, Wildtypen oder Tiere mit unbekanntem Phänotyp sind. In Abbildung 3.13 ist das allgemeine Vorgehen dargestellt.

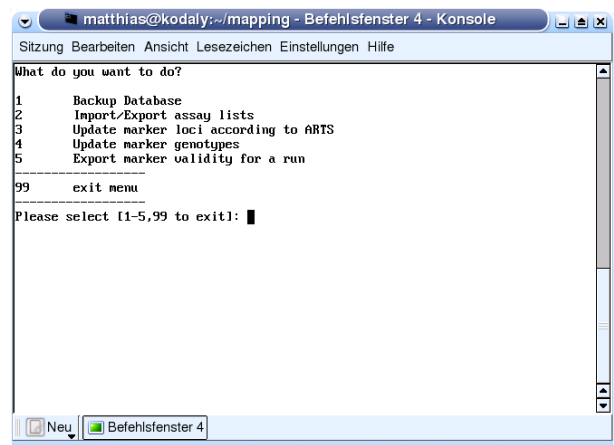


Abbildung 3.11: Bildschirmfoto des My-Maintenance Hauptmenüs

3. ERGEBNISSE

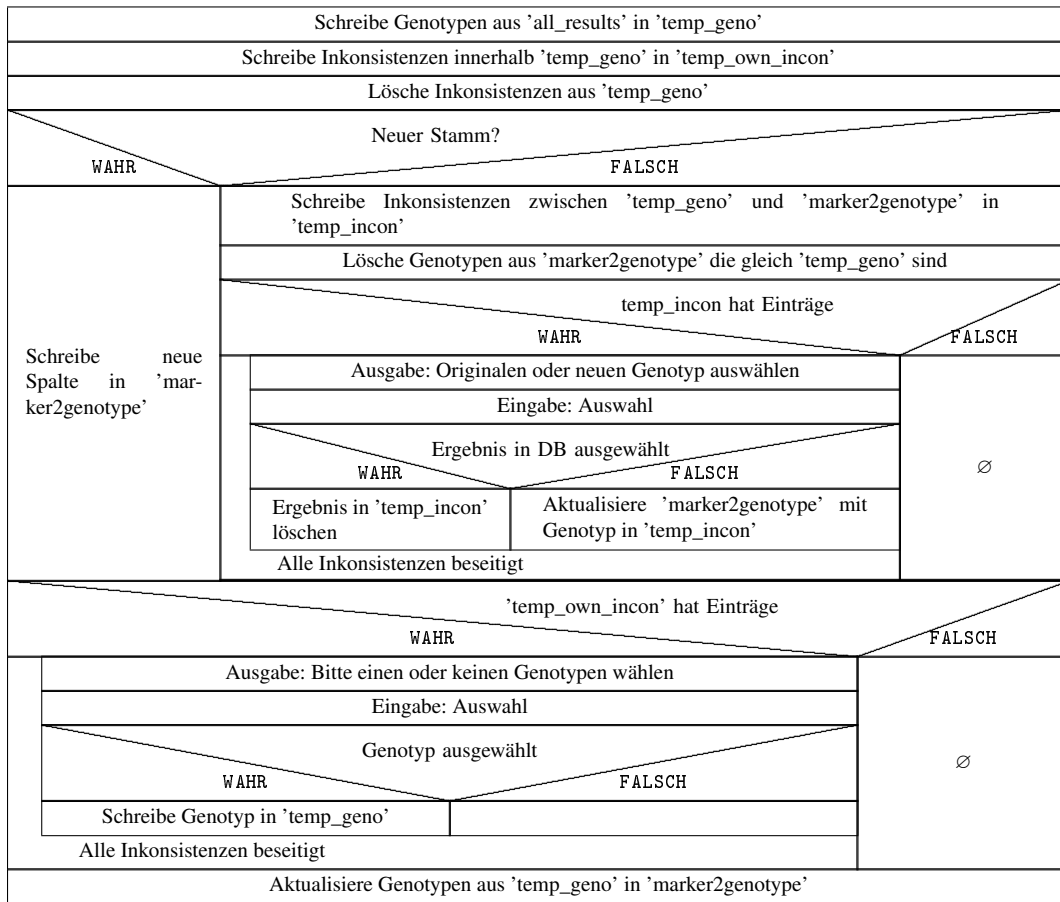


Abbildung 3.12: Die Mausstamm→Genotyp Zuordnung in MyMaintenance

Das CGI-Frontend `myline_frontend.pl` startet das Skript `myline.pl`, welches anhand der übermittelten Daten überprüft, ob es sich um die Anlegung einer neuen oder die Überschreibung einer existierenden Mauslinie handelt.

Bei der Anlegung einer neuen Mauslinie wird ein Eintrag in der Tabelle `lines` erstellt und die IDs in der Tabelle `samples` mit der entsprechenden Nummer der Mauslinie und dem passenden Phänotypen versehen. Falls eine Mauslinie überschrieben werden soll, wird das Skript `myline_overwrite.pl` gestartet. Nach Bestätigung des Benutzers wird überprüft, ob die eingegebenen IDs bereits einer anderen Linie zugeordnet sind. In diesem Falle werden die problematischen IDs aufgelistet und das Programm abgebrochen, um bereits getätigte Zuordnungen nicht zu überschreiben.

Andernfalls werden alle zu der Mauslinie gehörenden Daten in `lines` aktualisiert, in `samples` gelöscht und anhand der Eingaben neu zugeordnet.

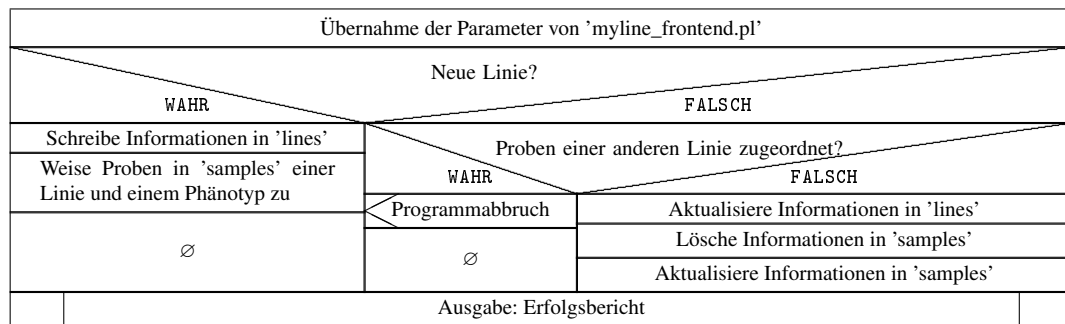


Abbildung 3.13: Das Struktogramm von MyLine

3.4.4. Das Programm MyGenotype

Benutzer können die Genotypisierungsergebnisse über ein CGI-Interface abrufen. Hierzu muss die Mauslinie, gegebenenfalls die Nummer des Genotypisierungsexperiments, sowie das Ausgabeformat gewählt werden. Zusätzlich kann angegeben werden, ob problematische Assays (Eintrag `problems` in der Tabelle `marker_validity`) angezeigt werden sollen. Ebenso kann die Grenze angegeben werden, unterhalb der keine Statistik mehr für einen Assay erhoben werden soll.

Das CGI Frontend `mygenotype_frontend.pl` startet das Skript `mygenotype.pl`, welches zunächst überprüft, ob die Mauslinie in mehr als einem Lauf analysiert wurde. Falls dies der Fall ist, wird das Programm abgebrochen und der Benutzer aufgefordert, nur einen Lauf auszuwählen. Die Genotypen aller Samples der ausgewählten Linie und gegebenenfalls des ausgewählten Laufs werden in eine neue Tabelle geschrieben.

Anschließend werden alle Assays nach verschiedenen Kriterien gefiltert. Genotypen werden als `bad` markiert, in denen ein Eintrag `invalid` für den entsprechenden Assay in der Tabelle `marker_validity` vorliegt. Falls der Filter für Assays mit dem Eintrag `problems` in der Tabelle `marker_validity` aktiviert wurde, werden diese Werte als `problem` gekennzeichnet. Genotypen bei denen in der `marker2genotype` Tabelle keine Polymorphie zwischen den parental Stämmen verzeichnet ist, werden mit dem Eintrag `mono` markiert. Daraufhin wird jedem Ergebnis ein Allel zugeordnet.

Genotypen, die dem Allel des ersten parental Stammes entsprechen, werden diesem zugeordnet und als `ref` (Referenz) bezeichnet. Homozygote Genotypen, die nicht dem ersten parental Stamm entsprechen, werden dem zweiten parental Stamm zugeordnet und als `ocr` (Outcross) bezeichnet. Eine Besonderheit bei dieser Art der Zuordnung ist, dass nur jeweils eines der beiden parental Allele bekannt sein muss, da das zweite Allel über die Definition „nicht das erste Allel und Länge 1“ bestimmt wird. Alle Genotypen mit zwei Nukleotiden werden als `het` (Heterozygot) gekennzeichnet und Genotyp-Einträge ohne Nukleotid werden mit `-` (nicht vorhanden) markiert. Nachdem dies erfolgt ist, wird entsprechend der Benutzereingabe eines der in Tabelle 3.5 aufgelisteten Ausgabeformate gestartet. Ein Ablaufschema ist in Abbildung 3.14 dargestellt.

3. ERGEBNISSE

Tabelle 3.5: Ausgabeskripte für den Aufruf aus mygenotype.pl

Skriptname	Funktion
mygenotype_html.pl	Stellt die Ergebnisse der Genotypisierungen im Browser dar
mygenotype_rqt1.pl	Ein tabellarisches Format, welches in R/QTL [24] eingelesen werden kann
mygenotype_dchip.pl	Auflistungen des Genotypisierungsergebnisse, der Markerpositionen und der Samples, welche in dChip [106] eingelesen werden können
mygenotype_mm.pl	Ein tabellarisches Format, welches in MapManager QTX [114] eingelesen werden kann

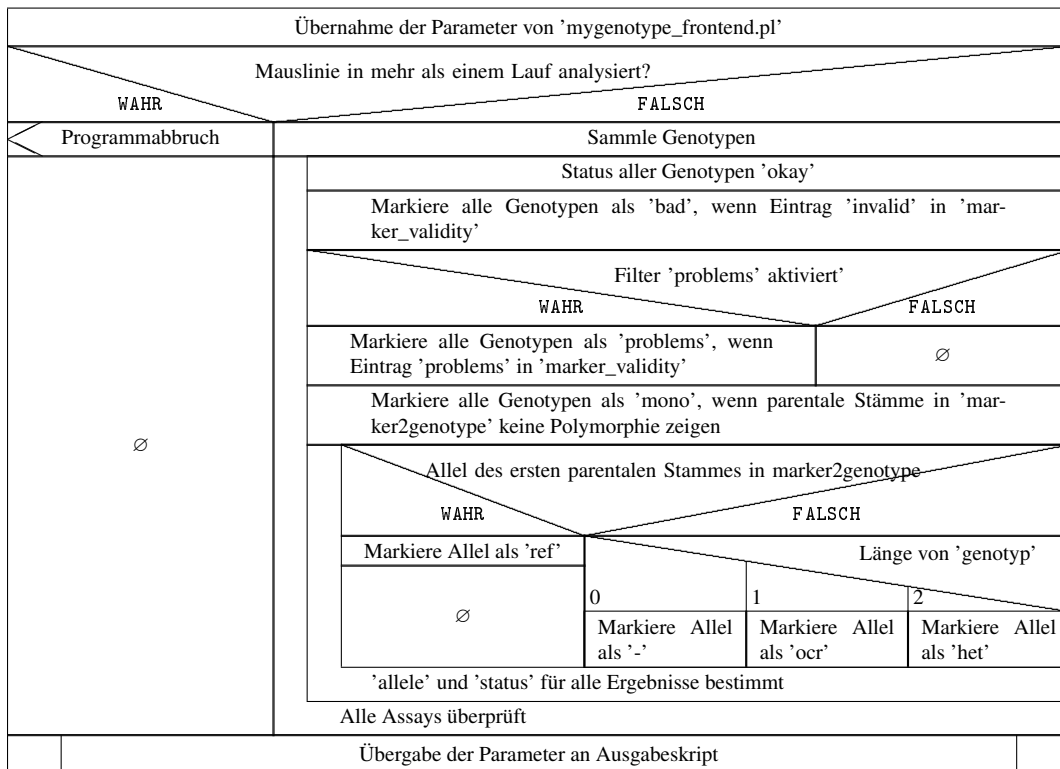


Abbildung 3.14: Ablaufschema des Programms MyGenotype

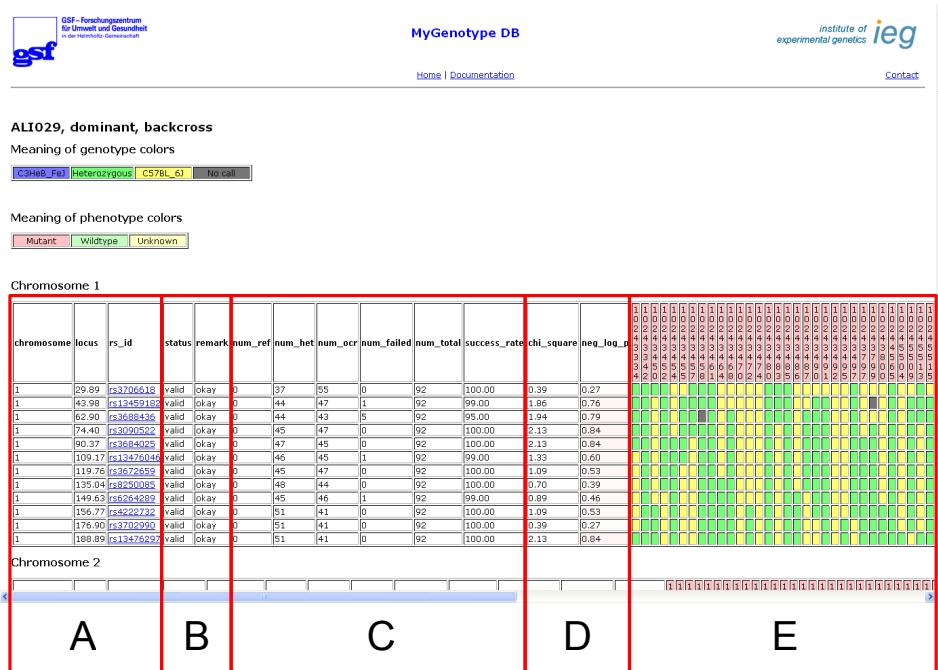


Abbildung 3.15: Bildschirmfoto des mygenotype_html.pl Ausgabemoduls. Genotypen und Phänotypen werden in unterschiedlichen Farben dargestellt. Die Ergebnisse für die einzelnen Marker und Proben werden zusammen mit einer statistischen Auswertung aufgelistet. A: Locusinformationen und Markerbezeichnung; B: Ergebnisse der Validierung; C: Anzahl der einzelnen Allele und Erfolgsrate; D: Statistische Angaben; E: Auflistung der Allele.

mygenotype_html.pl Die Ausgabe im Browser ermöglicht es dem Benutzer, einen schnellen Überblick über die Genotypisierungsdaten einer Mauslinie zu bekommen (vgl. Abbildung 3.15). Es wird ein χ^2 Test angewandt, welcher sich aus den Ausführungen in [175, Kapitel 9.1] ergibt. Für eine Rückkreuzung (N2) gilt die Hypothese H_0 , dass im Falle von Nicht-Kopplung des zu kartierenden Locus und des untersuchten Markers der Anteil von parentalen und rekombinanten Allelen bei 0,5 liegt. Die Wahrscheinlichkeit P dieser Aussage wird mit dem Perl Modul `Statistics::Distributions` berechnet. Wenn $P > 0,001$ ist, dann ist H_0 bestätigt. Wenn $P \leq 0,001$ ist, dann wird der zu kartierende Locus mit einer Wahrscheinlichkeit von mindestens 99% mit dem untersuchten Marker gekoppelt vererbt. In einem Rückkreuzungsexperiment können sowohl Mutanten als auch deren wildtypische Geschwister genotypisiert werden. Das Ergebnis der Kopplungsanalyse sollte im Falle von Kopplung ein gegensätzliches Verhältnis von Mutanten und Wildtypen beispielsweise mit parentalem Allel vorliegen.

Um eine statistische Analyse durchführen zu können, wird also die Anzahl der Mutanten mit rekombinanten Allel und die Anzahl der Wildtypen mit parentalem Allel addiert und mit der Summe aus Mutanten mit parentalem Allel und Wildtypen mit rekombinantem Allel verglichen. In Formel 3.5 ist dieser Kopplungstest dargestellt.

3. ERGEBNISSE

Es sei

a_{obs} : Anzahl Mutanten mit rekombinantem Allel + Anzahl Wildtypen mit parentalem Allel
 b_{obs} : Anzahl Mutanten mit parentalem Allel + Anzahl Wildtypen mit rekombinantem Allel
 t : Gesamtzahl erfolgreich genotypisierter Individuen

Es gilt

$$a_{obs} + b_{obs} = t$$

Annahme: Nicht-Kopplung von Marker und Locus

$$\text{Hypothese } H_0 : a_{obs} = b_{obs} = \frac{1}{2}t$$

Test auf Hypothese H_0

$$\chi^2 = \frac{(a_{obs} - a_{exp})^2}{a_{exp}} + \frac{(b_{obs} - b_{exp})^2}{b_{exp}} \quad (3.5)$$

Der in [175, Kapitel 9.1] beschriebene χ^2 Test mit drei Freiheitsgraden für die Auswertung von F2 Daten setzt ein Segregationsverhältnis von 1/8 nicht rekombinanten, 1/2 einfach rekombinanten, 1/8 doppelt rekombinanten und 1/4 nicht oder doppelt rekombinanten Allelen voraus. Durch die ausschließliche Analyse von Mutanten in der Genotypisierung ändert sich das zu erwartende Segregationsverhältnis auf 1/4 nicht rekombinante Allele, 1/2 einfach rekombinante Allele und 1/4 doppelt rekombinante Allele. Dies reduziert die Anzahl der Freiheitsgrade eines χ^2 Tests auf zwei.

Es sei

a_{obs} : Anzahl Mutanten mit nicht rekombinantem Allel
 b_{obs} : Anzahl Mutanten mit einfach rekombinantem Allel
 c_{obs} : Anzahl Mutanten mit doppelt rekombinantem Allel
 t : Gesamtzahl erfolgreich genotypisierter Individuen

Es gilt

$$a_{obs} + b_{obs} + c_{obs} = t$$

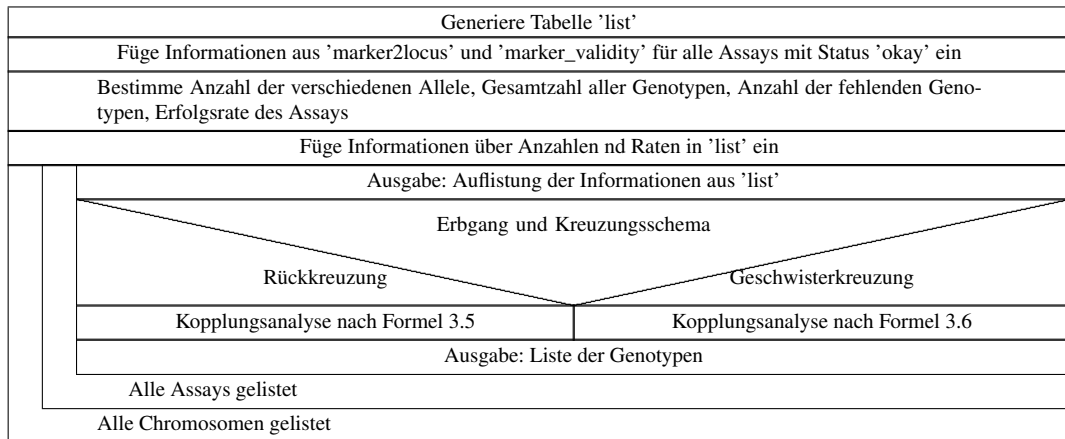


Abbildung 3.16: Das Ablaufschema von 'mygenotype_html.pl'

Annahme: Nicht-Kopplung von Marker und Locus

Hypothese $H_0 : a_{obs} = c_{obs} = 1/4t, b_{obs} = 1/2t$

Test auf Hypothese H_0

$$\chi^2 = \frac{(a_{obs} - 1/4t)^2}{1/4t} + \frac{(b_{obs} - 1/2t)^2}{1/2t} + \frac{(c_{obs} - 1/4t)^2}{1/4t} \quad (3.6)$$

Zur Darstellung der Ergebnisse wird zunächst die temporäre Tabelle `list` für allgemeine Informationen erstellt, in die die chromosomale Position aller Marker aus Tabelle `marker2locus` sowie die Validität aus `marker_validity` eingetragen werden. In diese Tabelle werden nur die `p2rs_ids` übernommen, deren Genotypen den Status `okay` haben. Daraufhin wird eine Reihe von temporären Tabellen angelegt, in welche die Anzahl der verschiedenen Allele, die Gesamtzahl aller Genotypen, sowie die Anzahl der fehlenden Genotypen abgelegt wird. Desweiteren wird die Erfolgsrate aller Assays bestimmt. Die Tabelle `list` wird anschließend mit diesen Werten ergänzt.

Es startet die chromosomenweise Auflistung der Ergebnisse. Nach Ausgabe der Kopfzeile erfolgt für jeden Assay die Auflistung der allgemeinen Parameter (Locus und Gültigkeit des Markers, Anzahl der verschiedenen Allele, Erfolgsrate des Assays), die Bestimmung und Ausgabe der in 3.5 und 3.6 dargestellten Kopplungstests, sowie die Auflistung aller Genotypen geordnet nach Phänotyp und ID (vgl. Abbildung 3.16). Zur besseren Übersicht werden die ermittelten P Werte als negativer dekadischer Logarithmus dargestellt [198]. Man kann die Hypothese H_0 also ablehnen, wenn $-\log_{10}(P) \geq 3$ ist.

3. ERGEBNISSE

mygenotype_rqtl.pl Das R/QTL [24] Format ist ein tabellarisches Format, welches mit Trennzeichen arbeitet. Es listet, im Gegensatz zu `mygenotype_html.pl`, die Ergebnisse der einzelnen Assays spaltenweise und die Ergebnisse der Proben zeilenweise auf. Die erste Zeile besitzt die Bezeichnungen der einzelnen Phänotyp-Eigenschaften sowie die Bezeichnungen der Marker (hier erfolgt eine Auflistung der `rs_ids` nach Chromosom und Locus). Die zweite Zeile hat keinen Eintrag bei den Phänotyp-Eigenschaften und enthält eine Zuordnung der Marker zu den Chromosomen.

Es folgt die Ausgabe des Phänotyps sowie der erhaltenen Genotypen. Die Liste wird im Browser ausgegeben und kann vom Benutzer gespeichert werden.

mygenotype_dchip.pl Bevor dChip [122] benutzt werden kann, ist es erforderlich, die Ausgabetafel mit den Allelen des Referenz- und des Auskreuzungsstammes, sowie mit denen des parentalen Tieres zu ergänzen. Diese Angaben dienen dem Programm dazu, nicht-informative Marker auszufiltern. Da dies bereits im `mygenotype.pl` Hauptprogramm geschehen ist, erhält der Referenzstamm für jeden Marker ein A, der Auskreuzungsstamm für jeden Marker ein B und ein virtuelles parentales Tier ein AB. Im Browser werden drei Listen ausgegeben, welche separat gespeichert werden.

Die erste Liste enthält die zeilenweise Auflistung aller Proben und die spaltenweise Auflistung der Marker. In Zeile 1 werden die Bezeichnungen „Mutant“, „Background“ und „Parental“ als erstes gelistet, danach folgen die aufsteigend geordneten IDs der Proben. In den folgenden Zeilen werden die Genotypen für jede `rs_id` gelistet.

Die zweite Liste enthält Informationen über die benutzten Marker. Das Programm verlangt nach einer Kopfzeile mit den Elementen `SNP_ID Chromosome Chromosomal Location Genetic Map Strand dbSNP RS ID`. Die Werte für die nachfolgende Liste sind an den Stellen `Chromosomal Location` und `Genetic Map` sowie `SNP_ID` und `dbSNP RS ID` identisch. Die Elemente `Genetic Map` und `dbSNP RS ID` sind für die korrekte Funktion des Programmes unerheblich und können im Grunde beliebige Werte annehmen. `Genetic Map` wird jedoch gleich `Chromosomal Location` und `dbSNP RS ID` wird gleich `SNP_ID` gesetzt.

In der dritten Liste schließlich wird der Phänotyp jeder Probe aufgeführt. Im Falle einer dominanten Vererbung gelten der Referenzstamm und das parentale Tier als „Affected“, bei rezessiver Vererbung ist das parentale Tier „Unaffected“.

mygenotype_mm.pl Das Modul `mygenotype_mm.pl` stellt die Genotypisierungsdaten in einem Format dar, welches von MapManager QTX [114] importiert werden kann.

Zunächst werden die letzten fünf Zeichen der IDs der Mauslinie getrennt durch ein Tabulatorzeichen aufgelistet. Anschließend wird der Name der Mauslinie, das Chromosom des Markers und die `rs_id` des Markers gelistet, ebenfalls jeweils durch ein Tabulatorzeichen getrennt. Die zugehörigen Genotypen werden danach ohne Trennzeichen aufgelistet. A steht hierbei für das maternale Allel, H für ein heterozygoten Allel und B für das paternale Allel. Ausfälle werden als - (minus) dargestellt.

3.5. Automatisierte DNA Extraktion

3.5.1. Anwendung

Ergebnisse der Überprüfung auf DNA Fragmentierung

Ziel der DNA-Extraktion war es, möglichst undegrierte DNA zu erhalten, um die Möglichkeit zu haben, auch Protokolle durchzuführen, bei denen hochmolekulare DNA Fragmente benötigt werden (z. B. Pulsfeld-Gelelektrophorese, Southern Blot). Bei Kontrollen der DNA auf Fragmentation mittels 0,8%-iger Agarosegele war stets eine klare Bande im Bereich >20 kbp sichtbar. Es konnte keine physikalische Scherung oder Degradation beobachtet werden (vgl. Abbildung 3.17). Die DNA ist damit von ihrer Qualität her sowohl für PCR Anwendungen als auch für Protokolle für hochmolekulare DNA geeignet.

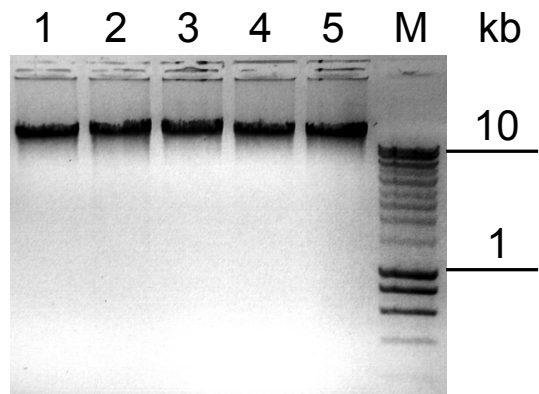


Abbildung 3.17: Die extrahierte DNA zeigt keinerlei Degradation oder physikalische Scherung. Die Ansicht wurde zur besseren Darstellung invertiert.

Ergebnisse Reinheit

Die Messung der Absorption im UV-Spektrum einiger Eluate aus einer routinemäßigen Extraktion ergab ein Verhältnis der Absorption bei 260 nm und 280 nm von durchschnittlich bei 1,9 mit einer Bandbreite zwischen 1,79 und 2,02 ($n=22$). Reine DNA weist ein Verhältnis $\geq 1,8$ auf.

Es wurde jedoch ein Absorptionsverhältnis von 260 nm zu 230 nm festgestellt, welches unabhängig von der erhaltenen Ausbeute zwischen 0,61 und 1,69 schwankte. Bei reiner DNA wäre ein Absorptionsverhältnis $\approx 2,0$ zu erwarten. Die Unabhängigkeit der Werte von der Ausbeute bei prinzipiell gleich behandelten Proben deutet auf die Anwesenheit einer Waschpufferkomponente im Eluat hin. Laut des Herstellers des eingesetzten Kits und nach eigener Erfahrung wirkt sich diese Pufferkomponente jedoch nicht auf nachfolgende Experimente aus. Speziell bei der Multiplex-PCR der Genotypisierung (vgl. Abschnitt 2.3) wurden keine signifikanten Unterschiede zu manuell extrahierter DNA festgestellt.

3.6. Ergebnisse Mauslinien

In Tabelle 3.6 werden die Ergebnisse der routinemäßigen Kartierung vorgestellt. Die Linienbezeichnung, der Erbgang, das Kreuzungsschema, das Chromosom und der negative dekadische Logarithmus der Wahrscheinlichkeit P des gekoppelten Markers [198] werden dargestellt. Wenn die ursächliche Mutation gefunden wurde, wird auch dies kurz erwähnt.

Von insgesamt 55 genotypisierten Mauslinien konnte die ENU-induzierte Mutation bislang in zwölf Linien gefunden werden. Die Analyse von fünf Mauslinien ergab eine Kopplung mit einem Locus, in dem sich ein Gen befindet, für das bereits viele ENU-induzierte Allele mit einem sehr ähnlichen Phänotyp beschrieben werden. Diese Linien wurden bei der Sequenzierung wegen der vermutlich niedrigen Chance auf die Entdeckung neuer Allele zurückgestellt. Die restlichen Linien sind entweder offen für neue Kollaborationspartner oder werden momentan weiter charakterisiert.

Der durchschnittliche Abstand der gefundenen Mutation zum Marker mit der statistisch signifikantesten Kopplung betrug 5,9 Mb (ca. 3 cM).

Kartierung *Dll1*-abhängiger ENU Mausmutanten Es wurden verschiedene Mausmutanten aus dem *Dll1*-abhängigen Sensitized Screen des MEP genotypisiert. Hierbei erwies sich der hME Markersatz als nur bedingt geeignet für die Untersuchung von Linien, die nicht Bestandteil der Grundstämme waren. Die geringe Anzahl an untersuchten Tieren pro Linie und die niedrigen Anzahl an informativen Markern sowie der nicht ingezüchtete Hintergrund der auf 129X1 basierten *Dll1*-Knockoutlinie erschwerten die Auswertung jedoch.

Es wurden die Linien DFE001, DFE002 und DTP001 untersucht. Nur bei Linie DFE001 konnte reproduzierbar eine Kopplung auf Chromosom 2 festgestellt werden [155].

Tabelle 3.6: Kartierungsergebnisse der untersuchten Mauslinien. **Erbgang rez:** Rezessiv, **dom:** dominant; **Kreuzungsschema B:** Rückkreuzung, **I:** Geschwisterverpaarung; **Abstand Marker** gibt den Abstand der gefundenen Mutation zum am stärksten gekoppelten Marker an.

Linienbezeichnung	Erbgang	Kreuzungsschema	Markersatz	Chromosom	$-\log_{10}(P)$ von χ^2	Mutation gefunden	Abstand Marker	Bemerkungen
ABE016	dom	B	hME	8	4,76	nein		
AEA001	dom	B	iPlex	8	12,92	nein		
ALI014	dom	B	hME	8	3,29	ja	9,9 Mb	Manuskript in Vorbereitung
ALI015	dom	B	hME	8	4,4	nein		
ALI027	dom	B	hME	6	11,46	nein		
ALI029	dom	B	hME	13	13,92	ja	4,5 Mb	siehe Abschn. 3.8.
ALI030	rez	I	iPlex	3	31,27	ja	1,1 Mb	
ALI031	dom	B	hME	11	3,19	nein		
ALI032	rez	I	hME	6	7,17	nein		

Tabelle 3.6: Kartierungsergebnisse Fortsetzung

Linienbezeichnung	Erbgang	Kreuzungs- schema	Marker- satz	Chromosom	$-\log_{10}(P)$ von χ^2	Mutation gefunden	Abstand Marker	Bemerkungen
ALI034	dom	B	iPlex	6	5,11	nein		
ALI035	rez	I	hME	10	17,8	nein		
ALI036	dom	B	hME	15	3,76	nein		
ALI037	rez	I	hME	9	19,11	ja	15,6 Mb	
ALI038	dom	B	hME	13	16,91	Verdacht		
ALI042	dom	B	iPlex	3	15,48	nein		
ATE001	dom	B	iPlex	5	13,36	ja	9,6 Mb	Manuskript angenommen
ATE004	rez	I	iPlex	9	24,11	nein		
BCC005	rez	I	hME	15	25,96	nein		
CCH001	rez	I	hME	6	24,36	nein		
CHO004	dom	B	iPlex	1	12,16	nein		
CHO028	rez	I	iPlex	5	8,27	nein		
CTA002	dom	B	hME	5	5,8	nein		
DSK016	dom	B	iPlex	7	18,51	nein		
FER002	rez	I	hME	19	3,82	nein		
FER003	rez	I	hME	9	13,92	nein		
GLS001	dom	B	hME	11	15,05	nein		
GLS004	dom	B	hME	7	18,07	ja	0,3 Mb	siehe [72]
HST001	dom	B	hME	7	5,41	nein		
HST002	rez	I	iPlex	6	10,43	nein		
HST009 ¹	rez	I	hME	2	27,25	ja	7,1 Mb	siehe [3]
HST012	rez	I	hME	7	9,51	ja	4,1 Mb	siehe [3]
KTA047	dom	B	hME	6	8,83	nein		
KTA050	dom	B	hME	2	16,13	nein		
MVG114	rez	I	iPlex	8	16,72	nein		
MYO005	rez	I	hME	5	6,86	nein		
PAN017	dom	B	hME	2	3,4	nein		
RCO009	dom	B	hME	11		nein		Ohne statistische Auswertung
RCO011	rez	I	hME	1	5,11	nein		
RCO012	dom	B	hME	15	7,23	ja	3,7 Mb	siehe [157]
RCO013	dom	B	hME	15	9,99	ja	3,7 Mb	siehe [157]
RCO014	rez	I	hME	19	28,91	nein		
RCO015	dom	B	hME	7	10,5	nein		
RCO016	rez	I	hME	1	22,83	nein		
RCO017	dom	B	hME	11	17,26	nein		
RCO019	dom	B	iPlex	11	18,29	nein		
SCO009	dom	B	hME	5	19,32	nein		
SCO010	dom	B	iPlex	10	10,83	Verdacht		
SCO011	dom	B	iPlex	10	12,25	Verdacht		
SMA004	rez	I	hME	4	11,07	ja	2,9 Mb	siehe Abschn. 3.7

¹C3HeB/FeJ×BALB/cByJ

Tabelle 3.6: Kartierungsergebnisse Fortsetzung

Linienbezeichnung	Erbgang	Kreuzungsschema	Marker-satz	Chromosom	$-\log_{10}(P)$ von χ^2	Mutation gefunden	Abstand Marker	Bemerkungen
SMA005 ¹	dom	B	iPlex	1	4,36	nein		
TRE001	rez	B	hME	6	25	nein		
TUB010	rez	I	hME	1	22,55	ja	8,4 Mb	
WBS012	dom	B	hME	9	9,84	nein		
WBS014	dom	B	hME	1	16,47	Verdacht		
WBS015	dom	B	iPlex	1	15,23	Verdacht		

3.7. Charakterisierung der Mausmutante SMA004

Die Mauslinie SMA004 wurde im Rahmen des Recessive Screens des MEP [77] isoliert. Zunächst fiel der kleine Wuchs des Gründertieres auf, daher wurde nach der internen Nomenklatur der Name SMA (=small) und die laufende Nummer 4 vergeben.

Der makroskopische Phänotyp der Tiere zeigte sich neben dem kleinen Wuchs am deutlichsten durch das reduzierte Gewicht im Vergleich zu nicht betroffenen Geschwistern, durch eine reduzierte Mobilität, sowie den fortschreitenden Verlust des Greifvermögens. Weiterhin wurde eine Fellreduktion am gesamten Körper und die Ergrauung des Fells um die Nasenspitze, sowie eine Verkrümmung des Rückgrades (Skoliose bzw. Kyphose) festgestellt. Tiere mit dem $Zmpste24^{SMA004/SMA004}$ Genotyp zeigen mit fortschreitendem Alter eine zunehmende Hinfälligkeit und sterben nach maximal neun Monaten. Die Tiere wurden einer weiteren Phänotypisierung unterzogen, sowie die genetische Ursache und mögliche molekulare Auswirkungen auf den Phänotyp studiert.



Abbildung 3.18: Tiere aus der SMA004 Stammzucht im Alter von drei Monaten. Links $Zmpste24^{+/+}$, rechts $Zmpste24^{SMA004/SMA004}$.

3.7.1. Verpaarungsmodus

Die Zucht der Tiere erwies sich als schwierig, da $Zmpste24^{SMA004/SMA004}$ Tiere sowohl männlichen als auch weiblichen Geschlechts eine reduzierte Fertilität haben, was vermutlich auf die reduzierte Mobilität

¹Spontanmutation auf C57BL/6J

der Tiere zurückzuführen ist. Im Verlauf der Arbeit wurde nur einmal die Geburt dreier lebender Jungtiere durch ein *Zmpste24*^{SMA004/SMA004} Weibchen beobachtet.

Die Linie ist rezessiv ohne embryonale Sterblichkeit auf C3HeB/FeJ Hintergrund, was sich anhand der Verteilung der Genotypen von Nachkommen aus *Zmpste24*^{SMA004/+} × *Zmpste24*^{SMA004/+} Verpaarungen zeigen lässt (vgl. Tabelle 3.7). Für die Erhaltungszucht bedeutete dies, dass ohne Genotypisierungsmöglichkeit nur sechs bis acht Wochen alte Männchen und Weibchen ohne Phänotyp verpaart werden konnten. Bei den Tieren handelte es sich also zu 2/3 um *Zmpste24*^{SMA004/+} Tiere und zu 1/3 um *Zmpste24*^{+/+} Tiere.

Die Wahrscheinlichkeit, durch zufällige Verpaarung zweier Tiere ohne Phänotyp eine Verpaarung zweier *Zmpste24*^{SMA004/+} Tiere zu erhalten, lag also bei $2/3 \cdot 2/3 = 4/9$. Diese Zuchtpärchen konnten als *Zmpste24*^{SMA004/+} × *Zmpste24*^{SMA004/+} identifiziert werden, wenn in der Nachkommenschaft mindestens ein *Zmpste24*^{SMA004/SMA004} Tier gefunden wurde. Nach Identifikation der Mutation konnten die Tiere genotypisiert werden und es wurden in der Regel Verpaarungen zwischen einem *Zmpste24*^{SMA004/+} -Männchen und zwei *Zmpste24*^{SMA004/+} -Weibchen angesetzt.

Tabelle 3.7: Genotypverteilung der Nachkommen aus *Zmpste24*^{SMA004/+} × *Zmpste24*^{SMA004/+} Verpaarungen

	<i>Zmpste24</i> ^{SMA004/SMA004}	<i>Zmpste24</i> ^{SMA004/+}	<i>Zmpste24</i> ^{+/+}	Σ
Anzahl	44	83	40	167
Anteil	26%	50%	24%	100%

3.7.2. Kartierung der Mutation

Eine genomweite Kopplungsanalyse wurde mit der ersten Version des in Abschnitt 2.4.1 vorgestellten Panels durchgeführt. Eine Auswertung der Genotypdaten von 24 Mutanten mit dem Programm MyGenotype (vgl. Abschnitt 3.4.4) ergab eine starke Kopplung ($\chi^2 = 51$; $-\log_{10}(P) = 11,07$) auf Chromosom 4 an Marker rs3682306 (123,45 Mb¹). Eine zweite Analyse des Markers rs3682306 mit anderen Proben, sowie die Analyse der Mikrosatelliten D4Mit187 (100,25 Mb¹), D4Mit334 (119,21 Mb¹) und D4Mit12 (123,87 Mb¹) bestätigten die Gültigkeit des Ergebnisses (vgl. Abbildung 3.19).

Durch die Filterung aller infrage kommenden Gene in der Region zwischen 115 und 125 Mb, die in der Literatur mit den Gewebesuchworten „bone“ und „cartilage“ verknüpft waren (Tabelle 3.8) war es möglich, die Anzahl der Kandidatengene mithilfe von LitMiner (siehe Tabelle 2.23 und [112]) zu reduzieren. Insbesondere das Gen *Zmpste24* galt bei weiterführender Literaturrecherche aufgrund von Ähnlichkeiten des beobachteten Phänotyps mit bereits vorhandenen Knock-Out Linien (vgl. [13, 146]) als mögliches Kandidatengen.

¹Locus-Angaben sind bezogen auf Mouse Genome Assembly 36.1

3. ERGEBNISSE

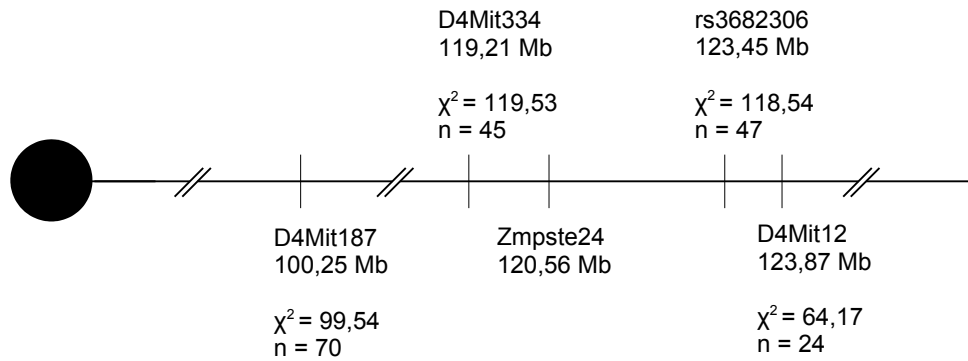


Abbildung 3.19: Kartierung der Mutation von SMA004 auf Chromosom 4

Tabelle 3.8: LitMiner-Kandidatengene für die Suchworte „bone“ und „cartilage“

Suchwort „bone“			Suchwort „cartilage“		
Genname	Locus [Mb]	Score	Genname	Locus [Mb]	Score
Bmp8b	121.12	33.57	Col9a2	119.18	151.32
Ermap	117.31	11.19	Bmp8a	121.33	7.76
Zmpste24	119.2	11.19	Slc2a1	117.24	1.53
Csf3r	124.04	7.95			
Atp6v0b	115.88	7.46			
Bmp8a	121.33	6.03			
Mpl	116.44	4.38			
4930538K18Rik	117.34	4.31			
Col9a2	119.18	3.36			
Artn	115.93	2.8			
Tie1	116.47	1.97			
Inpp5b	122.76	1.97			
Ptch2	115.1	1.24			
Cap1	120.88	1.08			

Score: Der „Overrepresentation score“ (vgl. [112]) ist ein Maß für die Häufigkeit einer bestimmten Schlüsselwort-Schlüsselwort-Kombination (hier: Genname und Suchwort) in der Literatur.

3.7.3. Untersuchung der genomischen DNA-Sequenz von *Zmpste24*

PCR Primer wurden von ExonPrimer (vgl. Tabelle 2.22) mit der cDNA Sequenz AK083566 (vorläufige RefSeq ID: NM_172700) und der korrespondierenden gDNA Sequenz aus den UCSC Genome Browser inclusive 1000 bp der 3'- und 5'-Sequenz mit den Standardparametern und der „rodent“ Mispriming Bibliothek berechnet. Eine PCR-Amplifizierung der Exons von *Zmpste24* wurde mit Pools von genomischer DNA (je fünf Mutanten und fünf Tiere der C3HeB/FeJ Referenzlinie) durchgeführt.

Zur Sequenzierung wurden die selben Primer wie bei der Amplifizierung eingesetzt. Zur Kompensation von nicht auswertbaren Sequenzspuren zu Beginn und am Ende einer Sequenzierreaktion wurde die Sequenzierung eines PCR Produkts jeweils mit dessen Sense Primer und dem Antisense Primer angesetzt. Die Auswertung erfolgte mit der Phred/Phrap/PolyPhred/Consed Pipeline (vgl. Tabelle 2.22). Abbildung 3.20 zeigt das Ergebnis der mit Hilfe von PolyPhred identifizierten Mutation. Eine G→A Mutation unmittelbar nach Exon 6 konnte sowohl bei der Sequenzierung mit dem Sense Primer und dem Antisense Primer gefunden werden. Ein Vergleich mit den Sequenzen anderer Mauslinien und anderer Species mit ClustalX (vgl. Tabelle 2.22 und [193]) bestätigte, dass sich die beobachtete Mutation an einer hochkonservierten Splice-Donor-Stelle im Genom befindet und auf die ENU-Mutagenese zurückzuführen ist (vgl. Abbildung 3.21).

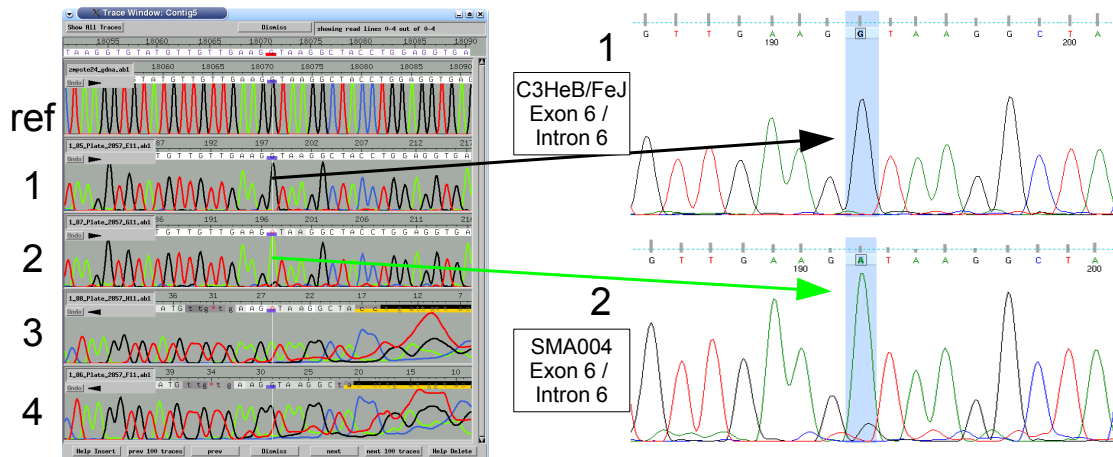


Abbildung 3.20: G→A Mutation nach Exon 6 in *Zmpste24*. ref: simulierte Referenzsequenz, 1: Wildtyp-Sequenz Sense, 2: SMA004-Sequenz Sense, 3: SMA004-Sequenz Antisense, 4: Wildtyp-Sequenz Antisense.

3. ERGEBNISSE

```

          **** ***** ** **** ***** * * * *
SMA004   TTTGACTAAGGTGTATGTTGTTGAAGTAAGGCTACCTGGAGGTGAG
C3HeJ/FeJ TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
C57Bl/6J TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
AKR      TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
Balb/cByJ TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
Cast/Ei  TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
CBA      TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
CFW      TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
FVB      TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
SEC      TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACCTGGAGGTGAG
Rattus_norvegicus CCTGACTAAGGTGTATGTCGTTGAAGGTAAGGCTACCTGGAGGTGAC
Canis_familiaris  TTTGACTAAGGTGTATGTTGTTGAAGGTAAGGCTACTGAGAGGTGAG
Homo_sapiens_sapiens TTTGACGAAGGTGTATGTTGTGGAAGGTAAGGCTACCTGGGGATAAG
Pan_troglodytes  TTTGACGAAGGTGTATGTTGTGGAAGGTAAGGCTACCTGGGGATAAG

```

Abbildung 3.21: Die gefundene G→A Mutation befindet sich in einer hochkonservierten Region, wie das Alignment mit verschiedenen Mausstämmen und Spezies zeigt.

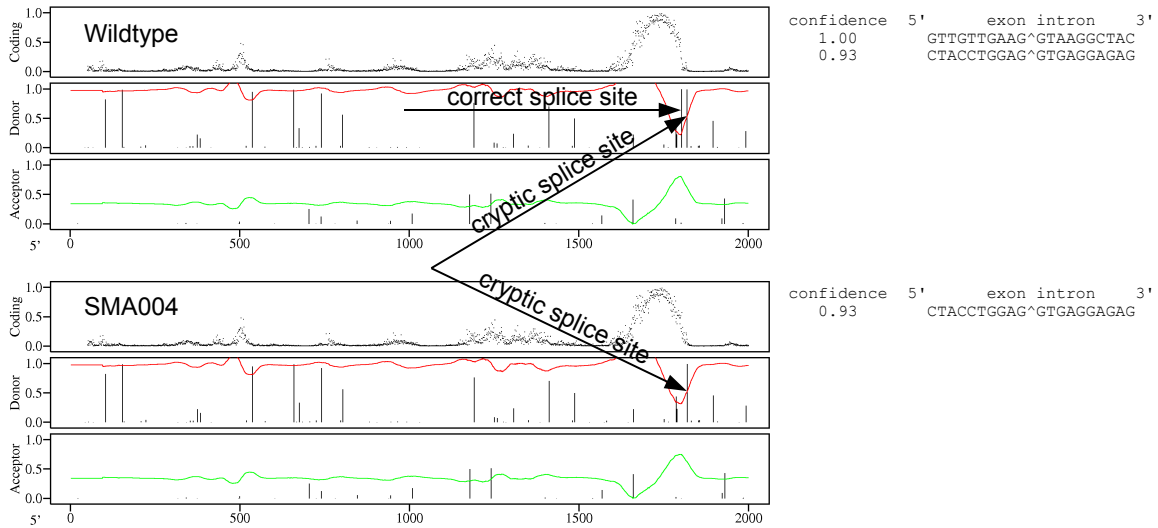
3.7.4. Untersuchung der mRNA von SMA004

Die Mutation einer Splice-Donor-Stelle kann verschiedene Konsequenzen für die Prozessierung der nascenten RNA zur reifen mRNA haben. In der Literatur werden Fälle der Transkription des kompletten Introns beschrieben, der Übergehung des der Mutation vorhergehenden Exons und des alternativen Splicings an einer kryptischen Splice-Donor-Stelle. All diese Fälle können einzeln aber auch gemeinsam auftreten, wie zum Beispiel von Maruyama *et al.* beschrieben [117].

Eine Analyse der genomischen Sequenz von 3'-Exon 5 bis 5'-Exon 7 inclusive Introns mit NetGene2 (vgl. Tabelle 2.21 und [70]) ergab, dass eine kryptische Splice-Donor-Stelle 16 bp nach der eigentlichen Splice-Donor-Stelle in Intron 6 existiert, welche mit hoher Wahrscheinlichkeit aufgrund der Mutation aktiviert wird.

Es wurden vier Primer generiert (vgl. Anhang A), um diese Hypothese zu beweisen oder die tatsächlichen Auswirkungen der Mutation mittels RT-PCR zu studieren (Abbildung 3.23). Die erwarteten Längen der PCR Produkte werden in Tabelle 3.9 aufgelistet. Die Gesamt-RNA von Mutanten und Wildtypen wurde aus Lebergewebe extrahiert und in komplementäre Einzelstrang-DNA umgeschrieben.

Die Untersuchung ergab, dass die kryptische Splice-Donor-Stelle aktiviert wird. Ein Durchlesen von Intron 6 und ein Übergehen von Exon 6 findet nicht statt (vgl. Abbildung 3.24). Der Einbau von 16 zusätzlichen Nukleotiden und die daraus resultierende Verschiebung des Leserahmens wurde per Sequenzierung bewiesen. Diese Verschiebung führt zur Transkription eines Stopp-Codons und zu einer Verkürzung von Zmpste24 von 465 Aminosäuren im Wildtyp auf 262 Aminosäuren in Zmpste24^{SMA004/SMA004} Tieren (Gly257fsX263).



Upper panel: coding potential, middle: donor site prediction and variable 90% threshold, lower: acceptor prediction and 90% threshold.

Abbildung 3.22: Vorhersage von Splice-Donor-Stellen in Intron 6 für die genomische Wildtyp-Sequenz und die SMA004 Sequenz mit NetGene2 [70]. In der genomischen Sequenz des Wildtyps wird sowohl die korrekte als auch eine mögliche kryptische Splice-Donor-Stelle vorhergesagt. In der Sequenz von SMA004 wird nur die kryptische Splice-Donor-Stelle vorhergesagt.

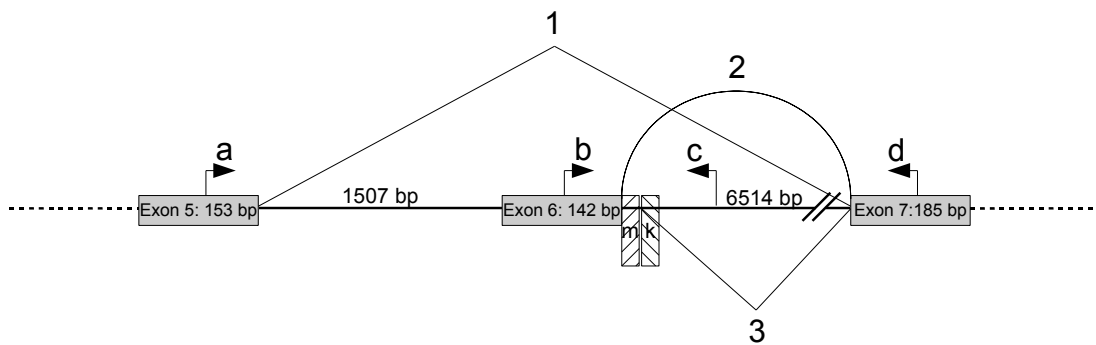


Abbildung 3.23: Splicevarianten von SMA004. a,b,c,d: Positionen der benötigten Primer; m: mutierte Splice-Donor-Stelle in Intron 6; k: vorhergesagte kryptische Splice-Donor-Stelle in Intron 6 16 bp nach der natürlichen Splice-Donor-Stelle; 1: Übergehen von Exon 6; 2: Durchlesen von Intron 6; 3: Splicing an kryptischer Splice-Donor-Stelle in Intron 6.

3. ERGEBNISSE

Tabelle 3.9: Größen der PCR-Produkte von Kombinationen der Primer aus Abbildung 3.23

Primer-kombination	Übergehen von Exon 6			Durchlesen von Intron 6			Splicing an kryptischer Splice-Stelle		
	SMA004	wt	gDNA	SMA004	wt	gDNA	SMA004	wt	gDNA
a,c	k.P.	k.P.	(2020 bp)	513 bp	k.P.	(2020 bp)	k.P.	k.P.	(2020 bp)
a,d	197 bp	<i>339 bp</i>	(8360 bp)	(6853 bp)	339 bp	(8360 bp)	355 bp	<i>339 bp</i>	(8360 bp)
b,c	k.P.	k.P.	419 bp	419 bp	k.P.	<i>419 bp</i>	k.P.	k.P.	419 bp
b,d	k.P.	245 bp	(6759 bp)	(6759 bp)	245 bp	(6759 bp)	261 bp	<i>245 bp</i>	(6759 bp)

K.P.: Kein Produkt zu erwarten; in Klammern: hypothetische Produktlängen, die wegen der PCR Konditionen nicht erreicht werden; **fett**: Erwartetes PCR-Produkt, falls der Fall zutrifft; **kursiv**: Positivkontrolle

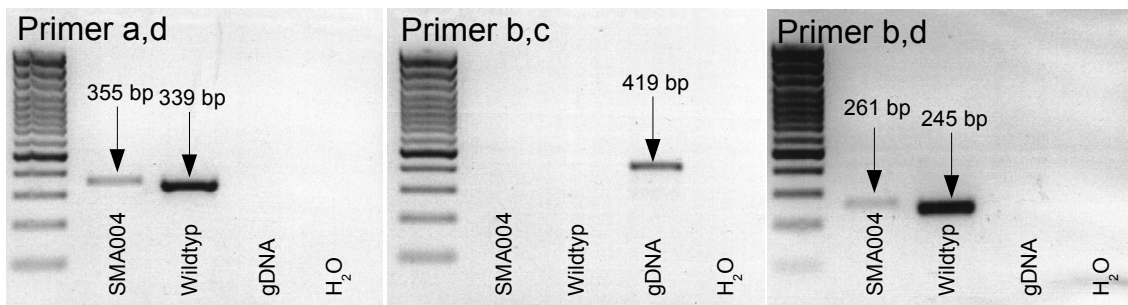


Abbildung 3.24: Ergebnisse der PCR-Analyse von *Zmpste24*^{SMA004/SMA004} cDNA, *Zmpste24*^{+/+} cDNA und genomischer DNA mit den in Abbildung 3.23 gezeigten Primern. Die Aktivierung der kryptischen Splice-Donor-Stelle kann anhand der Ergebnisse der Primerkombinationen a,d und b,d gezeigt werden. Die Ansicht wurde zur besseren Darstellung invertiert.

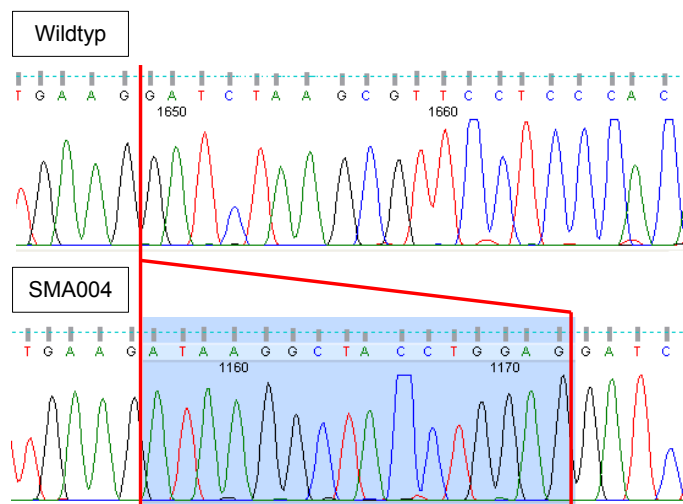


Abbildung 3.25: Die Aktivierung der kryptischen Splice-Stelle in *Zmpste24*^{SMA004/SMA004} in Intron 6 führt zur Anwesenheit von 16 zusätzlichen Nukleotiden in der mRNA.

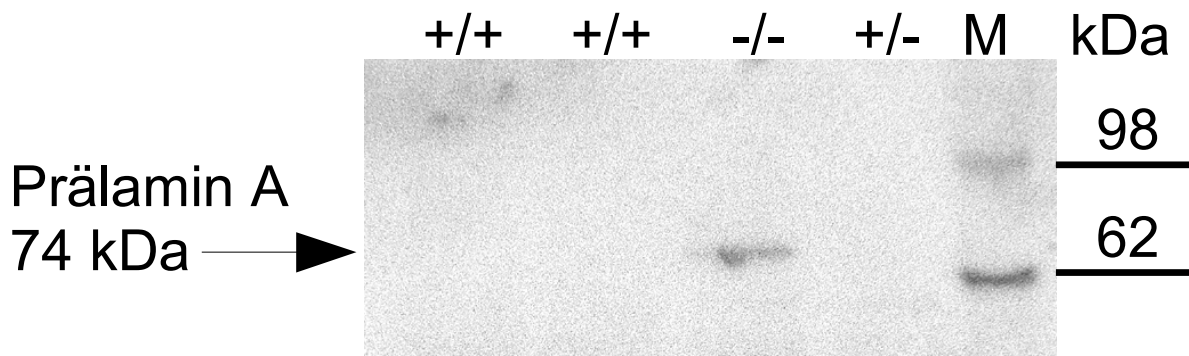


Abbildung 3.26: Western Blot eines Proteinextrakts aus der Niere mit einem Antikörper gegen das C-terminale Ende von Prälamin A, welches bei vollständiger Prozessierung zu reifem Lamin A mehr vorhanden ist.

3.7.5. Western Blot Lamin A

Auf Exon 8 wird das Motiv HELGH codiert. Im translatierten Protein wird dieses Motiv von den Aminosäuren 335 bis 339 gebildet. Dieses konservierte Zink-Metalloprotease-Motiv (allgemeine Form HEXXH), ist notwendig für die proteolytische Prozessierung von Proteinen (vgl. [189, 21]). Durch die Verschiebung des Leserahmens und dem vorzeitigen Stopp der Transkription fehlt dieses Motiv in *Zmpste24*^{SMA004/SMA004} Tieren, daher ist ein Verlust der Proteaseaktivität von *Zmpste24* zu erwarten.

Dies wirkt sich insbesondere auf die Prozessierung von farnesyliertem Prälamin A, dem bislang einzig bekannten Substrat von *Zmpste24*, zu reifem Lamin A aus. Dieses Protein wird in zwei Schritten von *Zmpste24* in seine reife Form gebracht (vgl. Abbildung 1.11) Es wurde daher untersucht, ob das Produkt des fehlerhaften *Zmpste24* Gens in *Zmpste24*^{SMA004/SMA004} Mäusen vollständig die Fähigkeit zur Prozessierung von Prälamin A verliert und ob die Prozessierung von Prelamin A abhängig von der Dosis von *Zmpste24* ist. Zu diesem Zwecke wurde ein Western Blot mit einem Antikörper durchgeführt, der spezifisch an das im reifen Lamin A nicht vorhandene C-terminale Ende von Prälamin A bindet.

Prälamin A konnte in den untersuchten Wildtypen und dem *Zmpste24*^{SMA004/+} Tier nicht nachgewiesen werden, in der untersuchten *Zmpste24*^{SMA004/SMA004} Mutante zeigte sich jedoch eine deutliche Bande (siehe Abbildung 3.26). Dies beweist, dass die Insertion von 16 zusätzlichen Nukleotiden zu einem nicht-funktionellen *Zmpste24* Protein führt, und dass eine um die Hälfte reduzierte Dosis von *Zmpste24* in *Zmpste24*^{SMA004/+} Tieren ausreichend für die vollständige Umsetzung von Prälamin A zu reifem Lamin A ist.

3. ERGEBNISSE

Tabelle 3.10: Bestimmung der Knochendichte und des Körpergewichts von weiblichen *Zmpste24^{SMA004/SMA004}*, *Zmpste24^{SMA004/+}*, *Zmpste24^{+/+}* Tieren mittels pDEXA.

Parameter	+/+ (n=10)	+/- (n=8)	-/ (n=10)	+/+ vs. +/- P value	+/+ vs. -/ P value	+/- vs. -/ P value
BMD [mg/cm ²]	68 ± 4	72 ± 1	64 ± 2	n.s.	n.s.	< 0.05
pBMD [mg/ρ ²]	57 ± 4	62 ± 1	50 ± 4	n.s.	n.s.	< 0.05
BMC [mg]	920 ± 134	1124 ± 133	359 ± 92	n.s.	< 0.01	< 0.01
Bone Content [%]	2.97 ± 0.21	3.33 ± 0.23	2.58 ± 0.83	n.s.	n.s.	n.s.
Body Length [cm]	9.80 ± 0.20	9.90 ± 0.10	8.40 ± 0.10	n.s.	< 0.001	< 0.0001
Body Weight [g]	30.46 ± 3.02	33.36 ± 1.95	14.67 ± 0.65	n.s.	< 0.001	< 0.0001
Lean mass [units]	15.66 ± 1.19	12.51 ± 1.56	12.42 ± 2.36	n.s.	n.s.	n.s.
Lean Content [units×100/g]	53.53 ± 6.12	38.62 ± 6.01	82.30 ± 14.40	n.s.	n.s.	< 0.05
Fat mass [units]	10.96 ± 2.73	16.98 ± 3.15	1.32 ± 1.32	n.s.	< 0.05	< 0.01
Fat Content [units×100/g]	33.51 ± 6.78	49.61 ± 6.34	10.80 ± 10.80	n.s.	n.s.	< 0.05

3.7.6. Röntgenologische Untersuchung von SMA004

Röntgenabsorptiometrie Es wurde eine Varianzanalyse zur statistischen Auswertung der erhaltenen pDEXA Daten durchgeführt. *Zmpste24^{SMA004/+}* Tiere zeigen keine signifikanten Änderungen im Vergleich zu *Zmpste24^{+/+}* Tieren. Die Körperlänge und das Körpergewicht sind signifikant (P<0,001) reduziert bei *Zmpste24^{SMA004/SMA004}* Tieren im Vergleich zu *Zmpste24^{+/+}*- und *Zmpste24^{SMA004/+}* Tieren. Auch der Knochenmineralgehalt (Bone Mineral Content, BMC) ist im Vergleich zu *Zmpste24^{+/+}*- und *Zmpste24^{SMA004/+}* Tieren signifikant (p<0,01) erniedrigt.

Die Fettmasse ist im Vergleich zu *Zmpste24^{+/+}*- und *Zmpste24^{SMA004/+}* Tieren ebenfalls signifikant erniedrigt (P<0,05 und P<0,01). Nur im Vergleich zu *Zmpste24^{SMA004/+}* Tieren ist sowohl die gesamte als auch die periphere Knochenmineraldichte (periphäre Bone Mineral Density, BMD/pBMD), der Fett- und der Muskelgehalt jeweils mit einer Signifikanz von P<0,05 reduziert.

Computertomographie Jeweils zwei *Zmpste24^{SMA004/SMA004}*, *Zmpste24^{SMA004/+}* und *Zmpste24^{+/+}* Tiere wurden mittels CT untersucht. Die Skelette der *Zmpste24^{SMA004/+}* Tiere zeigten keine Auffälligkeiten im Vergleich zu den Skeletten der *Zmpste24^{+/+}* Tiere.

Eine Reihe von Auffälligkeiten wurde jedoch bei *Zmpste24^{SMA004/SMA004}* Tieren festgestellt (vgl. Abbildung 3.27). Bereits makroskopisch sichtbar war die starke Kyphose der *Zmpste24^{SMA004/SMA004}* Tiere (Abbildung 3.27 A). Bei der weiteren Analyse wurden kallusartige Kalzifikationen an den Rippenansätzen (Abbildung 3.27 B) festgestellt. Desweiteren wurden Läsionen des Os zygomaticum gefunden (Abbildung 3.27 C). Außerdem wurde ein starker Abbau des Os Ilium beobachtet (Abbildung 3.27 C und D).

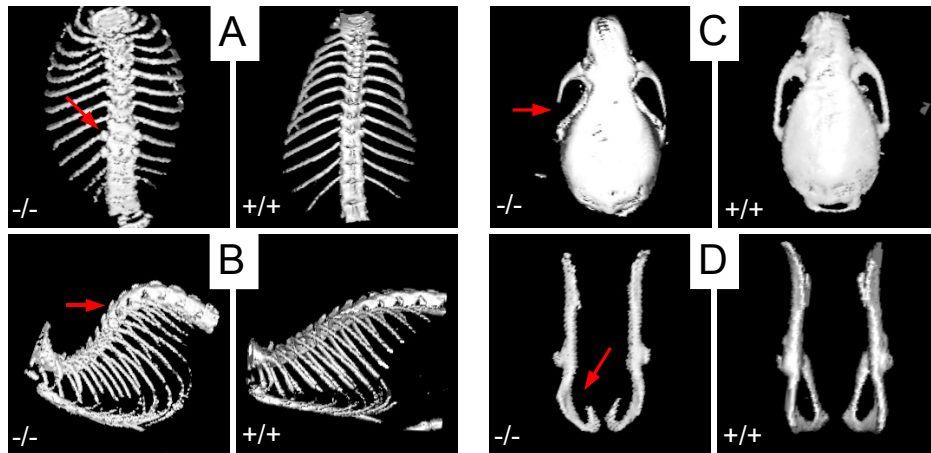


Abbildung 3.27: CT Scans von $Zmpste24^{SMA004/SMA004}$ (-/-) und $Zmpste24^{+/+}$ (+/+). A: Kallusartige Kalzifikationen an den Rippenansätzen; B: Starke Kyphose; C: Läsion des Os zygomaticum; D: Abbau des Os Ilium

3.7.7. Gewichtskurve

Die Gewichte von $Zmpste24^{SMA004/SMA004}$ (n=11), $Zmpste24^{SMA004/+}$ (n=19) und $Zmpste24^{+/+}$ (n=14) Geschwistern wurden nach dem Absetzen bis zu einem Alter von 26 Wochen ermittelt. Aus Abbildung 3.28 wird ersichtlich, dass $Zmpste24^{SMA004/SMA004}$ Tiere deutlich langsamer an Gewicht zunehmen und ein wesentlich reduziertes Endgewicht haben als $Zmpste24^{+/+}$ und $Zmpste24^{SMA004/+}$ Tiere. Eine mögliche Ursache kann die weithin in der Literatur beschriebene Lipodystrophie $Zmpste24$ - oder $Lmna$ -defizienter Mäuse sein. Eine alternative Erklärung wäre, dass $Zmpste24^{SMA004/SMA004}$ Tiere aufgrund ihrer eingeschränkten Mobilität nicht ausreichend Nahrung und Wasser zu sich nehmen können.

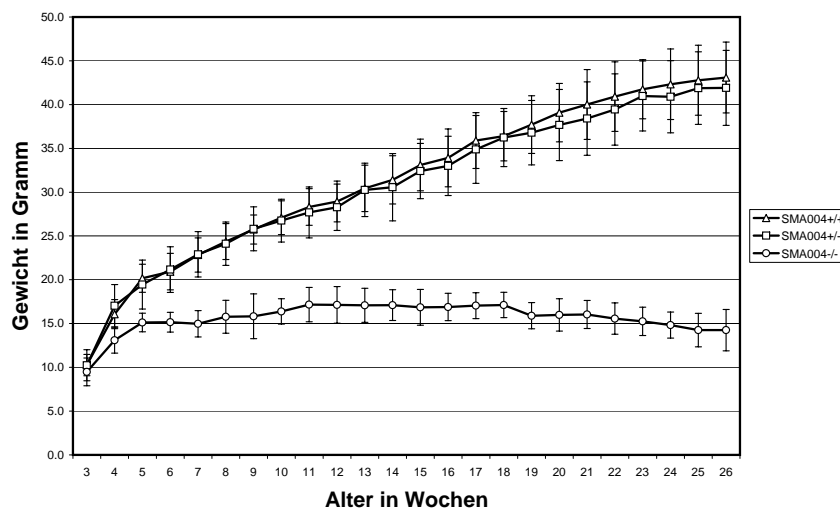


Abbildung 3.28: Die Gewichtszunahme der $Zmpste24^{SMA004/SMA004}$ (n=11) Tiere unterscheidet sich deutlich von der von $Zmpste24^{SMA004/+}$ (n=19) und $Zmpste24^{+/+}$ Tieren (n=14).

Tabelle 3.11: Zielgene von Transkriptionsfaktoren, die mit Lamin A interagieren

Grundkomponente	Transkriptionsfaktor	reguliertes Gen
Lamin A	Srebf1 [110]	<i>Vegf</i> [123]
		<i>ApoA5</i> [84]
		<i>Pcsk9</i> [35]
<i>Casp2</i> [111]		
	Mok2 [42]	<i>Rbp3</i> [5]
		<i>Pax3</i> [5]
	pRB [141]	<i>Rbp1</i> [14]

3.7.8. Quantifizierung von Zielgenen von Lamin A Interaktionspartnern

Zur Entstehung der progerieartigen Symptome von *Zmpste24*-defizienten Mäusen trägt vermutlich maßgeblich die Destabilisierung des Zellkerns durch die Anwesenheit von farnesyliertem Prälamina A bei. Alternativ hierzu ist weitläufig bekannt, dass Lamin A und Lamin C mit einer Anzahl von verschiedenen Zellkomponenten interagiert, z. B. Komponenten der Zell- und Zellkernarchitektur (z. B. Actin [166], LAP2 α [37]), Chromatin (z. B. DNA [184], Histone [191]) und Transkriptionsfaktoren (Srebf1a/c [110], Mok2 [42], pRB [141]), wie von Zastrow *et al.* in einer Übersichtsarbeit zusammengestellt wurde [212]. Es wurde daher die Hypothese aufgestellt, dass durch die Abwesenheit von reifem Lamin A die Interaktion zwischen Laminfilamenten und den Transkriptionsfaktoren (TF) Srebf1a/c, Mok2 und pRB die Transkription von Zielgenen dieser TF in *Zmpste24*^{SMA004/SMA004} Tieren vom Normalzustand abweichen würde. Zielgene dieser mit Lamin A interagierenden Transkriptionsfaktoren wurden mittels des Programmes Bibliosphere (vgl. Tabelle 2.21) ermittelt. Es wurden insgesamt sieben Gene als Ziel der Untersuchungen ausgewählt (Tabelle 3.11).

Mit Primer3 (vgl. Tabelle 2.21 und [154]) wurden Primer an den Exon-Exon-Grenzen der Zielgene erstellt, um eine hochspezifische relative mRNA-Quantifizierung durchzuführen.

Die Effizienz der generierten Primersätze wurde zusammen mit Primersätzen für die Housekeeping-Gene β -Actin, *Gapdh* und *Hprt1* bestimmt. Alle Primerpaare zeigten eine Effizienz über 1,7 und waren daher für eine Quantifizierung geeignet.

Expression verschiedener Gene im Muskel Die Expression der sieben in Tabelle 3.11 aufgelisteten Gene wurde in einem RNA-Pool von fünf *Zmpste24*^{+/+} Proben und in fünf *Zmpste24*^{SMA004/SMA004} RNA-Proben mittels RT-PCR untersucht. Alle Proben stammten von sechs Wochen alten Tieren.

Das Gen für Proprotein convertase subtilisin/kexin type 9 (*Pcsk9*) zeigte eine Überexpression in den fünf *Zmpste24*^{SMA004/SMA004} Proben im Vergleich zum RNA-Pool der *Zmpste24*^{+/+} Tiere (35,2 Zyklen \pm 0,49 in *Zmpste24*^{SMA004/SMA004}; 37,26 Zyklen \pm 0,321 in *Zmpste24*^{+/+}). Tabelle 3.12 zeigt die Ergebnisse des Vergleichs der Expression zwischen dem Wildtyp-Pool und den *Zmpste24*^{SMA004/SMA004} Proben.

Tabelle 3.12: Ergebnisse des Expressionsanalyse von Muskelproben

	Mittelwert <i>Zmpste24</i> ^{+/+} Pool	Standard- abweichung <i>Zmpste24</i> ^{+/+} Pool	Mittelwert <i>Zmpste24</i> ^{SMA004/SMA004} Einzelproben	Standard- abweichung <i>Zmpste24</i> ^{SMA004/SMA004} Einzelproben	Differenz der Mittelwerte
ApoA5	>41	>41	>41	>41	n.v.
Casp2	27,69	0,04	27,41	0,51	0,28
Pax3	28,35	0,22	29,10	0,61	-0,75
Pcsk9	37,26	0,32	35,20	0,49	2,06
Rbp1	25,13	0,05	24,96	0,52	0,17
Rbp3	>41	>41	>41	>41	n.v.
Vegf	23,96	0,03	23,80	0,46	0,16
<i>β</i> -Actin	17,03	0,01	17,36	0,44	-0,33
Gapdh	13,26	0,04	13,53	0,33	-0,27
Hprt1	26,06	0,04	25,87	0,40	0,19

Die Werte geben den Crossingpoint an, also den Punkt des signifikant detektierbaren Amplifikationsstarts. Grau hinterlegt: Housekeeping-Gene.

Expression von *Pcsk9* in der Leber Zur Bestätigung der Hypothese, dass die Expression von *Pcsk9* von der Interaktion zwischen Lamin A und *Srebfl* abhängt, wurde RNA aus der Leber von je fünf *Zmpste24*^{+/+}- und *Zmpste24*^{SMA004/SMA004} Tieren extrahiert. Die Proben stammten von denselben Tieren, bei denen die Expression im Muskel analysiert wurde.

Die Untersuchung der Expression von *Pcsk9* erfolgte über die Bestimmung der Genexpression im Verhältnis zu *Hprt1*. Die Ergebnisse der quantitativen PCR wurden anhand einer Kalibrator-cDNA normalisiert. Als Kalibrator-cDNA wurde *Zmpste24*^{+/+} Probe Nummer 1 eingesetzt, um einen realen Bezug zu einer Einzelprobe und damit eine Validierungsmöglichkeit zu haben.

Eine signifikante Expressionsänderung von *Pcsk9* konnte in den Leberproben jedoch nicht festgestellt werden (vgl. Tabelle 3.13).

Tabelle 3.13: Ergebnisse der Expressionsanalyse von *Pcsk9* in je 5 *Zmpste24*^{SMA004/SMA004} und *Zmpste24*^{+/+} Tieren

	1	2	3	4	5	Mittelwert	Standardabweichung
<i>Zmpste24</i> ^{SMA004/SMA004}	0,85	1,77	1,1	1,64	0,8	1,23	0,4
<i>Zmpste24</i> ^{+/+}	1,02	0,76	0,94	1,06	0,78	0,91	0,12

Es wird die Expressionsdifferenz zwischen den Genen *Hprt1* und *Pcsk9* relativ zur Kalibrator-cDNA (*Zmpste24*^{+/+} Probe 1) angegeben.

3. ERGEBNISSE

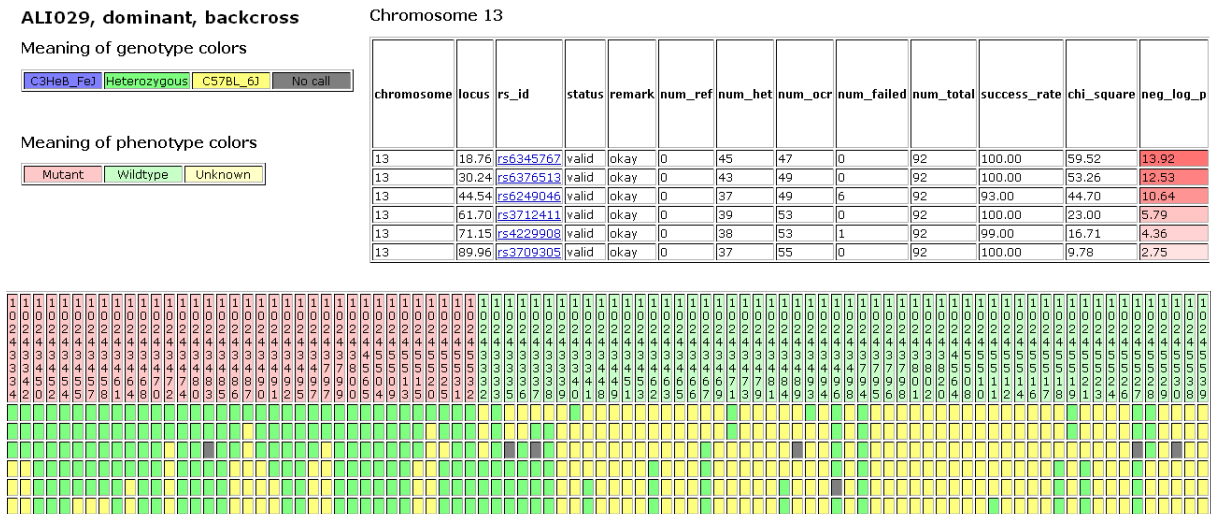


Abbildung 3.30: Kopplung des ALI029 Locus mit Marker rs6345767 auf Chromosom 13. Oben rechts: Statistikausgabe von MyGenotype, $-\log_{10}(P)$ von χ^2 . Unten: Auflistung der Genotypisierungsergebnisse für die oben rechts aufgelisteten Marker.

3.8. Charakterisierung der Mausmutante ALI029

Diese Mauslinie wurde im Rahmen des Dominant Screens des MEP [77] isoliert. Sie fiel wegen präaxialer Polydaktylie auf. Sie wurde zur Kartierung der Mutation auf den Stamm C57BL/6J ausgekreuzt und die Nachkommen wiederum auf C57BL/6J rückgekreuzt.



Abbildung 3.29: Die Mauslinie ALI029 zeigt eine präaxiale Polydaktylie.

3.8.1. Kartierung der Mutation

Im Rahmen der allgemeinen Kartierung mit dem SNP Markersatz wurden DNA-Proben von 36 Mutanten und 56 Wildtypen der Mauslinie ALI029 mit dem SNP Mapping Panel kartiert. Die Auswertung mit dem Programm MyGenotype ergab eine starke Kopplung ($\chi^2 = 59,52$; $-\log_{10}(P) = 13,92$) auf Chromosom 13 mit Marker rs6345767 (19,4 Mb¹). Aufgrund dieser Kartierungsergebnisse und des Polydaktylie-Phänotyps der ALI029 Mutanten wurde das Gen für GLI-Kruppel family member GLI3 (*Gli3*) (15,25 Mb¹) als möglichen Kandidatengen in Betracht gezogen.

¹Locus-Angaben beziehen sich auf Mouse Genome Assembly 36.1

3.8.2. Untersuchung der genomischen DNA Sequenz von *Gli3*

Die Exons 10-13 wurden gezielt sequenziert, da von diesen das C2H2 Zinkfinger-Motiv von Gli3 (Aminosäure 480-632) codiert wird. Die Auswertung der Ergebnisse erfolgte analog zu Abschnitt 3.7.3 mit der Phred/Phrap/PolyPhred/Consed Pipeline. Es wurde eine heterozygote G→A Mutation im 5' Bereich vor Exon 11 gefunden (vgl. Abbildung 3.31). Es wurde vermutet, dass die Mutation zur Entstehung einer kryptischen Splice-Stelle führt und dadurch die Funktion des Zinkfinger-Motivs gestört wird.

Die Analyse dieses Bereiches mit Netgene2 ergab, dass in der genomischen DNA Sequenz von ALI029 eine neue Splice-Stelle vorhergesagt wird, die möglicherweise in Konkurrenz mit der natürlichen Splice-Stelle steht.

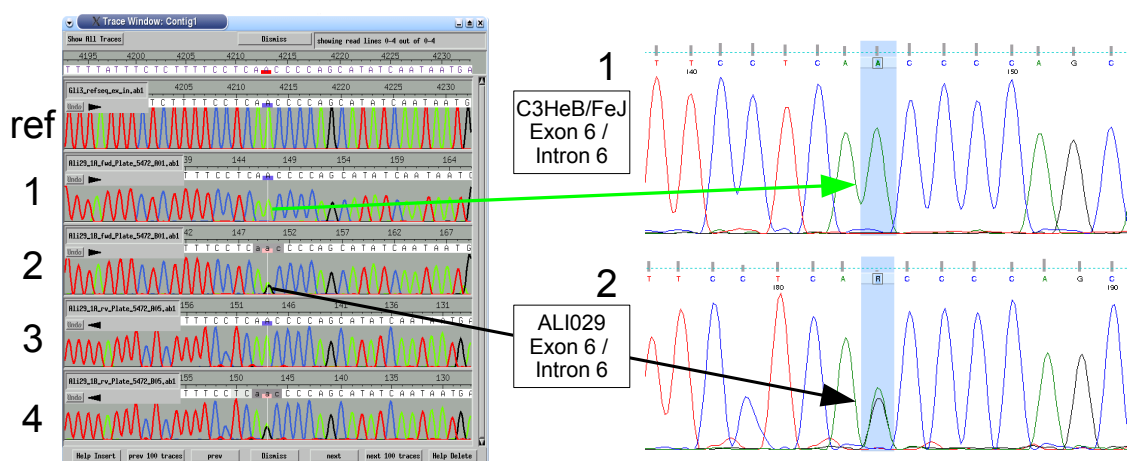


Abbildung 3.31: Eine heterozygote A→G Mutation vor Exon 11 in *gli3*. ref: simulierte Referenzsequenz, 1: Wildtyp-Sequenz Sense, 2: ALI029-Sequenz Sense, 3: ALI029-Sequenz Antisense, 4: Wildtyp-Sequenz Antisense

3. ERGEBNISSE

```

** * ** * * * ***** ***** *****
ALI029 TTATTTCTCTTTTCCTCAGCCCCAGCATATCAATAATGACCACATTCAT
C3HeB/FeJ TTATTTCTCTTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
C57Bl/6J TTATTTCTCTTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
Balb/cByJ TTATTTCTCTTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
Cast/Ei TTATTTCTCTTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
FVB TTATTTCTCTTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
Rattus_norvegicus TTAATTTCTCTTTCCTCAACCCCAGCATATCAATAATGACCACATTCAT
Canis_familiaris TTCCTTTCTTCTTCCAACCCCAGCATATAAATAATGACCATATTCAC
Homo_sapiens_sapiens TTCTCTCTCTTTTTCCAAACCCCAGCATATAAATAACGACCATATTCAT
Pan_troglodytes TTCTCTCTCTTTTTCCAAACCCCAGCATATAAATAATGACCATATTCAT

```

Abbildung 3.32: Die gefundene G→A Mutation befindet sich am Ende einer hochkonservierten Region, wie das Alignment mit verschiedenen Mausstämmen und Spezies zeigt.

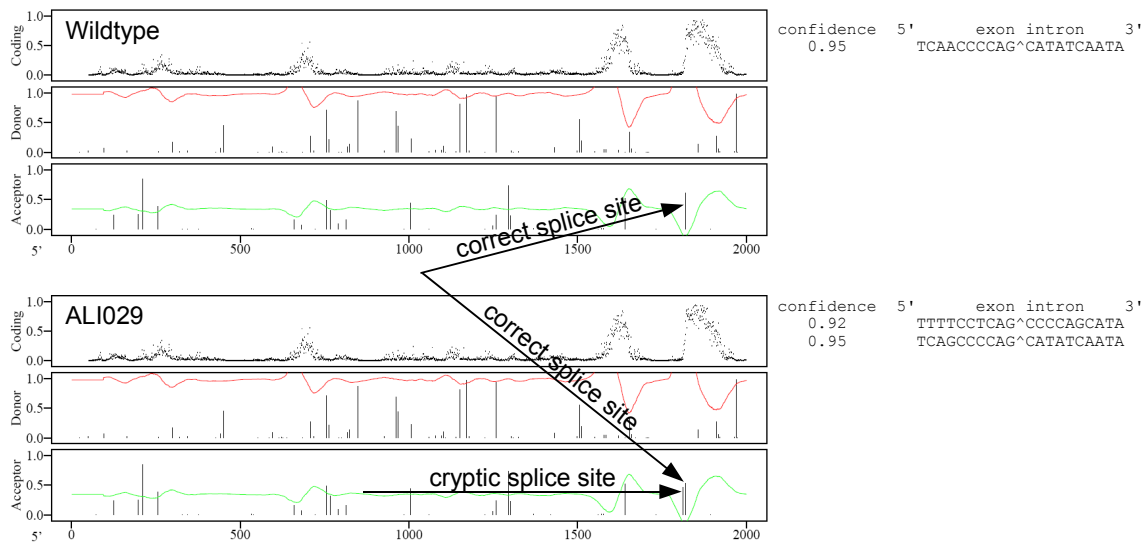


Abbildung 3.33: Vorhersage von Splice-Donor-Stellen in Intron 10 für die genomische Wildtyp-Sequenz und die ALI029 Sequenz mit NetGene2 [70]. In der genomischen Sequenz des Wildtyps wird nur die korrekte Splice-Akzeptor-Stelle vorhergesagt. In der Sequenz von ALI029 wird zusätzlich eine kryptische Splice-Akzeptor-Stelle vorhergesagt.

4. Diskussion

4.1. Design von Markersätzen für die Kopplungsanalyse

4.1.1. Design einfacher Markersätze

Um einen einfachen Markersatz zusammenzustellen reicht es aus, SNPs zu benutzen, die Polymorphismen zwischen zwei Mausstämmen aufzeigen. Um SNP Marker auszuwählen, die diese Bedingung erfüllen, gibt es in Internet diverse Tabellen und Suchmaschinen, z. B. die dbSNP (www.ncbi.nlm.nih.gov/SNP/MouseSNP.cgi) und die SNP Suchmaschinen von Perlegen (mouse.perlegen.com/mouse/browser.html) und Roche (mousesnp.roche.com/cgi-bin/msnp.pl). Die Suche gestaltet sich jedoch mühsam, da es erforderlich ist, jede gewünschte Region einzeln nach den gewünschten Markern zu durchsuchen. Desweiteren ist in der Regel nicht gewährleistet, dass der eingetragene Polymorphismus zwischen den zu vergleichenden Mauslinien real ist oder auf ein Sequenzartefakt zurückzuführen ist. Ein Markersatz, der auf diese Weise generiert wurde, wird nur Marker enthalten, die für die zu untersuchende Kreuzung informativ sind. Die Genotypisierungsergebnisse von Kreuzungen anderer, nicht nah verwandter Stämme lassen sich nicht vorhersagen.

Der erste in dieser Arbeit vorgestellte Markersatz (vgl. Abschnitt 2.4.1) wurde in dieser Form erstellt. Der zweite Markersatz sollte bereits für die Genotypisierung von Auskreuzungen des C3HeB/FeJ Stammes mit zwei verschiedenen Stämmen geeignet sein. Hierbei ergaben sich Probleme bei der Auffindung geeigneter Marker, da die zu dieser Zeit verfügbaren Suchmaschinen nicht in der Lage waren, Polymorphismen zwischen mehr als zwei Mausstämmen zu finden.

4.1.2. Design von Markersätzen mit hoher Heterozygotie

Zu Beginn dieser Arbeit existierte noch kein auf SNPs basierender Markersatz, der ohne weitere Recherche und Optimierung im Rahmen des MEP einsetzbar gewesen wäre. Petkov *et al.* präsentierten erstmals 2004 einen 1638 SNPs umfassenden Markersatz, der für die phylogenetische Analyse von 102 Mausstämmen eingesetzt wurde und auf eine hohe Heterozygotie zwischen den Stämmen ausgelegt war [147]. Seit kurzem werden Markersätze zur Genotypisierung von Mäusen auch von kommerziellen Anbietern wie Illumina oder Affymetrix angeboten. Sie sind als Kits erhältlich, die aus mehreren Hundert bis Tausend Markerassays bestehen und auf die Genotypisierung der am häufigsten in der Wissenschaft eingesetzten Mausstämme angepasst sind.

Der Einsatz solcher Markersätze ist sinnvoll bei der Erstellung von Stammbäumen oder bei der Analyse von komplexen Erkrankungen und QTLs. Diese Ansätze machen die Analyse vieler verschiedener

4. DISKUSSION

Mausstämme erforderlich, um über den Vergleich von Haplotypen Rückschlüsse auf Verwandtschaften oder gemeinsame Allele ziehen zu können. Eine weitere Anwendung findet sich bei der Erstellung vieler verschiedener rekombinanter Inzuchtstämme, die im Rahmen des Collaborative Cross [30] entstehen sollen. Hier dient ein Markersatz mit einer hoher Heterozygotie der genetischen Kontrolle und der Auswahl passender Zuchtpaare, um in wenigen Schritten eine hohe Anzahl an RI-Stämmen zu erzeugen.

Der Nachteil des Einsatzes eines Markersatzes mit hoher Heterozygotie ist, dass die verwendeten Marker nie vollständig heterozygot zwischen zwei Mausstämmen sein werden. Daraus resultiert, dass der Informationsgehalt der erhobenen Daten immer niedriger als 100% ist und der physikalische Abstand zwischen zwei informativen Markern sehr groß sein kann. Dies wird kompensiert durch den erhöhten Einsatz von Markern, mit dem versucht wird, eine Auflösung und Markerverteilung zu erreichen, wie sie ein Markersatz zur Genotypisierung von nur zwei Mausstämmen bieten würde.

Ein solcher Markersatz kann zur routinemäßigen Kopplungsanalyse von Kreuzungen weniger Mausstämme eingesetzt werden, jedoch wird bei jeder Analyse eine hohe Zahl an Daten produziert, die keine Information enthalten und unter Umständen ist die Markerverteilung lückenhaft. Unter Einsatz des Markersatzes von Petkov *et al.* hätten 132 SNPs die erforderliche Heterozygotie zwischen den im MEP eingesetzten Mausstämmen gezeigt [147]. Die Verteilung dieser Marker hätte deutliche Lücken aufgewiesen. 981 Marker (59,9%) wären informativ für eine C3HeB/FeJ×C57BL/6J Kreuzung, 786 Marker (48%) wären informativ für die Analyse von C57BL/6J×AKR/J Kreuzungen. Heterozygotie zwischen C3HeB/FeJ und BALB/cByJ zeigen 519 Marker (31,3%).

Aus dem kürzlich von Moran *et al.* vorgestellten Satz von 394 Markern wären maximal 66 Marker für einen Einsatz im MEP geeignet gewesen, wobei keine Daten für den Stamm AKR/J vorlagen, was die Anzahl der qualifizierten Marker vermutlich weiter reduzieren würde [122]. 241 Marker (61,1%) sind informativ für eine C3HeB/FeJ×C57BL/6J Kreuzung und 125 Marker (31,7%) sind informativ für C3HeB/FeJ×BALB/cByJ Kreuzungen.

Das „Low Density Linkage Panel“ von Illumina benutzt 377 Marker, von denen 271 (71,9%) informativ zwischen den Stämmen C3HeB/FeJ und C57BL/6J sind. 212 Marker (56,2%) wären bei der Kopplungsanalyse einer Kreuzung zwischen den Stämmen C57BL/6J und AKR/J informativ und 111 Marker (29,4%) wären für die Auswertung einer Kreuzung zwischen C3HeB/FeJ und BALB/cByJ geeignet (vgl. www.illumina.com). Die letztere Kreuzung wäre mit diesem Markersatz nicht auswertbar, da sehr viele Chromosomen nur lückenhaft abgedeckt sind und sich auf Chromosom 9 kein heterozygoter Marker befindet. 25 Marker zeigen die passende Heterozygotie für alle im MEP relevanten Kreuzungen.

4.1.3. Design einer Mischform

Für das MEP [77] hat sich insbesondere die Mauslinie C3HeB/FeJ als geeignet für die Mutagenese durch ENU erwiesen [180]. Für die Kartierung werden Träger einer ENU-induzierten Mutation auf den Stamm C57BL/6J ausgekreuzt [78]. Da in den Kreuzungsversuchen gelegentlich die Penetranz des untersuchten Phänotyps aufgrund von Modifiereffekten des C57BL/6J Genoms reduziert ist, oder der C57BL/6J Hintergrund die Ergebnisse eines Genotypisierungsexperiments verändern kann, wird auf den alternativen

Auskreuzungsstamm BALB/cByJ zurückgegriffen. Desweiteren wird im Rahmen des *DIII* Sensitized-Screens die ENU Mutagenese mit Tieren des Stammes C57BL/6J durchgeführt. Um auch hier einen alternativen Rückkreuzungsstamm zu C3HeB/FeJ zu haben, können Tiere des Stammes AKR/J eingesetzt werden. Die Anzahl der möglichen Kreuzungsschemata bewegt sich also in engen Grenzen mit einem deutlichen Übergewicht der Kreuzung C3HeB/FeJ×C57BL/6J.

Zusätzlich zur Analyse monogener Erkrankungen wird ein Sensitized-Screen mit heterozygoten Trägern eines loss-of-function Allels des Delta1 Gens, *DIII^{LacZ}* [79], auf dem 129X1 (nach früherer Nomenklatur 129/SvJ) Hintergrund durchgeführt. Hierzu werden Männchen des Stammes C57BL/6J mit ENU behandelt und mit *DIII^{LacZ}* Weibchen verpaart und die Nachkommen auf *DIII*-abhängige Phänotypen untersucht. Da es sich bei diesem Stamm nicht um einen reinen Inzuchtstamm handelt [176, 194], sind anderswo erhobene Genotyp-Daten nicht unbedingt auf den für den Sensitized Screen eingesetzten 129X1 Stamm übertragbar. Es bot sich daher an, einen Markersatz zu generieren, der eine Mischform der in Abschnitt 4.1.1 und 4.1.2 dargelegten Kriterien der SNP-Selektion bildet.

Mit diesem sollte erreicht werden, sowohl eine möglichst hohe Heterozygotie für Kreuzungen zwischen den vier „Kern“-Stämmen (C3HeB/FeJ, AKR/J, C57BL/6J, BALB/cByJ) zu erreichen, als auch durch die gezielte Auswahl von hochpolymorphen Markern Kreuzungen zwischen dem Stamm C3HeB/FeJ und einem „unbekannten“ Stamm (in diesem Falle dem 129X1 Stamm aus dem Sensitized Screen) genügend Information für eine statistisch signifikante Kopplungsanalyse zu erhalten.

Der dritte Markersatz (vgl. Abschnitt 2.4.2) stellt eine Lösung für dieses Problem dar. Aus Tabelle 3.3 wird ersichtlich, dass diese Ziele erreicht werden. 100% der Marker sind informativ für Kreuzungen der Stämme C3HeB/FeJ und C57BL/6J und 85% der Marker sind informativ für Kreuzungen zwischen C57BL/6J und AKR/J.

73% der Marker sind heterozygot zwischen dem Stamm C3HeB/FeJ und BALB/cByJ, was im Vergleich zu anderen Markersätzen sehr hoch ist. Die gewünschte Heterozygotie von annähernd 100% wird in diesem Fall jedoch nicht erreicht. Angesichts der Tatsache, dass diese beiden Mausstämme relativ nah verwandt sind (vgl. hierzu die Stammbäume von Beck und von Petkov [11, 147]), und dass sich auf dem Mausgenom große Regionen von niedriger Diversität befinden [203] ist dieser Wert jedoch nicht erstaunlich. So können mit ARTS beispielsweise auf Chromosom 10 in der Region zwischen 28,4 Mb und 98,3 Mb, ca. 53% des gesamten Chromosoms, nur 17 Polymorphismen zwischen diesen Linien gefunden werden. Auf dem gesamten Chromosom 10 befinden sich nur zwei SNPs, die die in Abschnitt 2.4.2 definierten Kriterien voll erfüllen. Daher mussten bei der Auswahl geeigneter SNPs einige Kompromisse zugunsten der Auswertbarkeit der am häufigsten eingesetzten Kreuzung zwischen C3HeB/FeJ und C57BL/6J eingegangen werden.

Die Existenz von Regionen mit einer niedrigen Anzahl an Polymorphismen wurde bereits von Wade *et al.* dokumentiert und die Existenz von 23 nichtpolymorphen Regionen über 20 Mb und die niedrige Diversität auf Chromosom 10 von Wiltshire *et al.* nachgewiesen [203, 207]. Die Struktur des Mausgenoms ist daher verschieden von der des menschlichen Genoms, in dem von Haplotyp-Blöcken von einigen 10 kb ausgegangen wird [151].

4.1.4. Schwierigkeiten bei der Auffindung geeigneter Marker

Zur Zeit der Generierung des zweiten Markersatzes für die hME Methode (vgl. Abschnitt 2.4.1), der nicht nur für die Analyse einer C3HeB/FeJ×C57BL/6J Kreuzung, sondern auch einer C3HeB/FeJ×BALB/cByJ Kreuzung geeignet sein sollte, gab es nur die Möglichkeit, mit den verfügbaren Suchmaschinen SNPs zu finden, die eine Polymorphie zwischen zwei Stämmen zeigten. Die Erstellung eines Satzes von ca. 80 Markern hätte es unter Zuhilfenahme dieser Tools notwendig gemacht, mindestens 80 Mal die zugrunde liegenden Datenbanken zuerst nach geeigneten Markern für eine C3HeB/FeJ×C57BL/6J Kreuzung, dann mindestens 80 Mal für eine C3HeB/FeJ×BALB/cByJ Kreuzung abzufragen. Die Ergebnisse hätten jeweils gespeichert und in ein Format gebracht werden müssen, das einen halbautomatischen Vergleich, beispielsweise in einem Tabellenkalkulationsprogramm, erlaubt.

Regionen, in denen auf diese Weise keine passenden Marker gefunden werden konnten, hätten neu evaluiert oder verschoben werden müssen. Insgesamt wäre diese Arbeit sehr zeitaufwendig und fehleranfällig gewesen. Die Nachvollziehbarkeit des Vorgehens bei der Fehlersuche wäre ebenfalls nicht gewährleistet gewesen.

Mittlerweile bieten die meisten Suchmaschinen die Option an, polymorphe SNPs zwischen einem Referenzstamm und mehreren Auskreuzungsstämmen zu finden. Nach wie vor wird allerdings nicht die Option angeboten, einen Vergleich zwischen mehreren Referenz- und Auskreuzungsstämmen durchzuführen, wie es in dem in Abschnitt 4.1.3 geschilderten Fall jedoch sinnvoll wäre.

Ein weiteres Manko speziell der Maus-dbSNP ist, dass die erhobenen Daten wie die Daten humaner SNPs ausgewertet werden. So ist die *in silico* Validierung durch die mindestens zweimalige Beobachtung des seltener auftretenden Allels pro RefSNP Cluster (vgl. www.ncbi.nlm.nih.gov/SNP/snp_legend.cgi?legend=validation) keine gültige Methode zur Validierung von Polymorphismen zwischen Inzuchtstämmen. Bei Kopplungsanalysen bzw. Assoziationsstudien im humanen Bereich wird die Häufigkeit des Auftretens eines bestimmten Allels in Verbindung mit einer bestimmten erblichen Eigenschaft untersucht. Hierfür ist die Validierung eines Polymorphismus durch das bloße Vorhandensein eines selten auftretenden Allels legitim.

Bei der Validierung von SNPs zwischen Inzuchtstämmen müssen andere Kriterien angelegt werden. Polymorphismen kommen innerhalb einer ingezüchteten Kolonie *per definitionem* nicht vor, es existiert also nur ein Allel. Die Frage nach der Möglichkeit des Vorhandenseins eines Allels in einer ausgezüchteten Population verschiebt sich also hin zu der Frage nach dem tatsächlichen Vorhandensein unterschiedlicher Allele in zwei oder mehr Inzuchtkolonien. In anderen Worten: Die für ausgezüchtete Populationen valide Aussage „In einer bestimmten Population kommt ein Allel mit einer Frequenz $0 < f < 1$ vor“ kann bei Kreuzungen zwischen zwei ingezüchteten Stämmen nicht getroffen werden, da die Allelfrequenz entweder 0,5 (Polymorphismus vorhanden) und 1 (Polymorphismus nicht vorhanden) beträgt. Die zweimalige Beobachtung eines Allels innerhalb einer Population von verschiedenen Kolonien von Inzuchtstämmen ist daher für die Kopplungsanalyse unerheblich und kann nicht zur Validierung benutzt werden.

Mit ARTS wird ein Konzept eingeführt, das die Chancen erheblich erhöht, SNPs zu identifizieren, die eine *in vitro* Validierung bestehen. So bietet zum Beispiel die Zahl der Institute, die einen SNP übermit-

teln, sowie die Anzahl der untersuchten Mausstämme eine gute Einschätzung über die Verlässlichkeit von Genotypdaten. Diese Daten sind in der dbSNP nicht ohne weitere Recherche verfügbar. In anderen Suchmaschinen sind solche Daten gar nicht vorhanden, weil nur Daten aus einer Quelle benutzt werden. Die berechneten Allelhäufigkeiten werden ebenfalls nicht von der dbSNP Suchmaschine angezeigt. In den Rohdaten (vgl. Abschnitt 3.3) befinden sich zwar Angaben zur Allelhäufigkeit pro Bericht und pro SNP, die Berechnung der Allelfrequenz pro SNP erfolgt allerdings nur aus dem Anteil der ermittelten Genotypen an der Gesamtsumme aller Daten ohne Rücksicht auf doppelt erhobene oder inkonsistente Genotypen. Eine exakte Berechnung der Allelhäufigkeit findet nicht statt.

Die gezielte Auswahl von SNPs, die mit einer Häufigkeit gegen 0,5 in einer gewissen Mindestanzahl von Inzuchtstämmen vorkommen, erhöht die Chancen, einen Markersatz zu erstellen, der zur Untersuchung von Kreuzungen verschiedener Mausstämme benutzt werden kann. Dies kann durch den Vergleich der Heterozygotie des nicht dahingehend optimierten hME Markersatzes (vgl. Abschnitt 3.2) und des iPlex Markersatzes (vgl. Abschnitt 3.2) bewiesen werden.

4.2. SNP Genotypisierung im Hochdurchsatz

4.2.1. Verwaltung und Auswertung der generierten Daten

Die praktische Arbeit für die Genotypisierung und die spätere Auswertung der Daten erfordert eine Vielzahl von Einzelschritten, die eine umfangreiche Dokumentation unbedingt erforderlich machen. In dieser Arbeit wurde Wert darauf gelegt, einen strukturierten Prozess zu generieren, der sowohl bei der praktischen Arbeit im Labor als auch bei der Datenverarbeitung an jedem Punkt Einfachheit und Nachvollziehbarkeit gewährleistet.

Zur Speicherung und Auswertung der Ergebnisse wurde das MyGenotype System entwickelt. Es besteht aus einer Datenbank, in welche die generierten Rohdaten importiert werden können, sowie aus einer Reihe von Programmen, die eine komfortable Kopplungsanalyse ermöglichen.

Bei der Planung des Projekts wurde schnell deutlich, dass für die Anzahl der Daten, die bei der DNA-Extraktion und der Genotypisierung erhoben werden, ein hohes Maß an Dokumentation notwendig ist. Diese Dokumentation sollte möglichst standardisiert ablaufen und mit möglichst wenig manuellem Arbeitsaufwand verbunden sein. Neben der Dokumentation der DNA Extraktion, deren Formulare später bei der Erstellung der Dateien für die Genotypisierung hilfreich sind, musste ein System entwickelt werden, das die Auswertung der erhobenen Genotypdaten einfach und nachvollziehbar macht.

Zu Anfang des Projekts existierte keine publizierte Lösung zur Verwaltung und Aufbereitung speziell von Genotypdaten von Maus-Inzuchtstämmen. Es gab jedoch einige Programme, z. B. MapManager QTX [114] und R/QTL [24], mit denen fertige Genotypdaten ausgewertet werden konnten. Eine Konvertierung der erhobenen Genotypdaten in ein Format für die Auswertung mit diesen Programmen lag daher nahe.

Für die Verwaltung der Proben und Genotypdaten musste eine individuelle Lösung gefunden werden. Die Tatsache, dass mehrere Mauslinien in einem Lauf analysiert werden, macht es notwendig, die Verteilung

4. DISKUSSION

der einzelnen Proben auf einer Platte sehr genau zu planen. Nach Durchführung der Genotypisierung müssen alle Proben wieder zweifelsfrei einer Mauslinie zugeordnet werden. Gleichzeitig ist es erforderlich, gegebenenfalls einer Probe den richtigen Phänotyp zuzuordnen. Die Genotypen einer Probe müssen zudem den korrespondierenden parentalen Allelen zugewiesen und die benutzten Marker einem Locus zugeordnet werden. Es hat sich desweiteren als nützlich erwiesen, eine einfache Darstellung der Daten für einen schnellen Überblick ohne detaillierte Kenntnisse der oben genannten Programme anzubieten. Zu Beginn des Projekts wurden die erhaltenen Genotypdaten heruntergeladen und halbautomatisch mit Hilfe von Tabellenkalkulationsprogrammen und Perl-Skripten ausgewertet. Es stellte sich jedoch heraus, dass diese Form der Auswertung sehr aufwendig ist und aus vielen Schritten besteht, die automatisiert werden können. Nach und nach erfolgte daher die Umstellung auf ein Datenbanksystem und die Erstellung diverser Programme zur Verwaltung und Auswertung der erhaltenen Genotypdaten (vgl. Abschnitt 3.4).

Im Rahmen dieser Arbeit wurden mehr als 830.000 Genotypdaten von mehr als 4200 DNA Proben generiert, die in die MyGenotype Datenbank eingelesen wurden. Mit wenigen Eingaben können diese Daten analysiert und manipuliert werden. In Kürze wird das MyGenotype-System als Teil der Genotypisierungsplattform veröffentlicht und ist damit das bislang einzige öffentlich verfügbare Datenbanksystem, das speziell auf die genomweite Kopplungsanalyse von Mäusen ausgelegt ist. Die Infrastruktur vergleichbarer Genotypisierungsplattformen wurde im Gegensatz dazu nicht offengelegt [122].

Das MyGenotype-System hat eine offene Struktur und kann einfach erweitert und in bestehende Datenbanksysteme eingebettet werden.

4.2.2. Konzeptionsbeweise

Mit Hilfe der beiden hier vorgestellten Markersätze war es möglich, innerhalb des Zeitraums der Arbeit eine genomweite Kopplungsanalyse von insgesamt 55 Mauslinien mit einem ENU-induzierten monogenischen Defekt durchzuführen. Bei zwölf dieser Linien konnte die zugrunde liegende Mutation entweder im Rahmen dieses Projekts oder von Kollaborationspartnern gefunden werden. Der Abstand der gefundenen Mutation zum Marker mit der stärksten Kopplung lag bei durchschnittlich 5,9 Mb (ca. 3 cM), was den theoretischen Erwartungen entspricht (vgl. Abschnitt 1.5).

Zwei dieser Mauslinien werden in dieser Arbeit als Konzeptionsbeweis für eine erfolgreiche Kartierung verschiedener Erbgänge und Kreuzungsschemata vorgestellt. Eine davon ist die rezessive Mauslinie SMA004, ein Modell für HGPS und MAD. Diese Mauslinie wurde in dieser Arbeit näher untersucht. Die zugrunde liegende Mutation verursacht den Wegfall einer Splice-Donorstelle. Das dadurch hervorgerufene alternative Splicing an einer kryptischen Splice-Stelle 16 bp nach der ursprünglichen Splice-Stelle führt zu einem vorzeitigen Stopp der Translation von *Zmpste24*. Die verkürzte Variante des Proteins verliert damit das konservierte Zink-Metalloprotease-Motiv HEXXH der Aminosäuren 335 bis 339 und damit die Fähigkeit, Prälamin A zu reifem Lamin A endoproteolytisch zu prozessieren, wie anhand eines Western Blots bewiesen werden konnte.

Eine dominante Mauslinie ist die ALI029, ein Modell für präaxiale Polydaktylie. Durch die gefundene

Mutation wird vermutlich eine zusätzliche kryptische Splice-Stelle aktiviert, die in Konkurrenz mit der natürlichen Splice-Stelle steht.

Von den restlichen gefundenen Mutationen werden fünf dominant und fünf rezessiv vererbt. Bei fünf weiteren Mauslinien wurde eine Kopplung mit einem Locus festgestellt, in dem sich ein Gen befindet, für das bereits viele ENU-induzierte Allele mit einem sehr ähnlichen Phänotyp beschrieben werden. Zu den bereits publizierten oder sich in der Publikation befindlichen Linien gehören die RCO012 und RCO013 Linien, in denen Missense Mutationen in der α -helikalen rod-Domäne von Krt71 gefunden wurden, einem Gen, das an der Bildung des inneren Haarwurzelschafts maßgeblich beteiligt ist [157]. Die Kandidatengenanalyse erfolgte direkt nach der genomweiten Kartierung mit dem hME Markersatz (vgl. Abschnitt 3.2).

Die Region der Kandidatengene von ATE1 wurde bereits 2004 auf einen Bereich von 9 cM auf Chromosom 5 eingegrenzt [171]. Bei der Sequenzierung von einigen Kandidatengen in dieser Region wurde eine Mutation in dem Gen für Enamelin (Enam), einem Bestandteil des Zahnschmelzes, identifiziert.

Es zeigte sich jedoch, dass die damals eingesetzten Marker D5Mit18 und D5Mit10 zwar auf der genetischen Karte die Region von *Enam* umgeben, auf der physikalischen Karte jedoch außerhalb dieses Bereichs liegen. Um Fehler bei der Korrelation der genetischen und der physikalischen Karte auszuschließen, wurde die Kopplungsanalyse mit dem iPlex Markersatz wiederholt. Mit diesem konnte die Region des *Enam* Gens zweifelsfrei bestätigt und weiter eingegrenzt werden.

Aufgrund der Kopplungsanalyse der Mauslinie GLS004 mit dem hME Markersatz konnte eine Mutation in Insulin 2 (*Ins2*) gefunden werden. Diese Mauslinie ist damit ein weiteres Modell für Diabetes mellitus Typ I [72].

Die Korrelation zwischen erhöhten Harnstoffkonzentrationen im Urin und ENU-induzierten genetischen Ursachen gelang in mehreren Fällen [3]. Die Loci der betroffenen Gene konnten ebenfalls mit dem hME Markersatz identifiziert werden. Die Mauslinie HST009 wurde auf BALB/cByJ ausgekreuzt und erfolgreich kartiert, was beweist, dass die Heterozygotie des hME Markersatzes ausreichend ist, um eine Kopplungsanalyse durchzuführen. Probleme bestanden bei dieser Kartierung hinsichtlich der Markerverteilung, die nicht auf allen Chromosomen optimal war. So konnte auf Chromosom 9 nur ein Marker und auf den Chromosomen 4, 13, 14 und 18 nur zwei Marker analysiert werden. Die Kopplung der Mutation mit einem Marker auf Chromosom 2 ist jedoch so deutlich, dass unter der Annahme einer monogenischen Erkrankung eine weitere Kopplung auf diesen Chromosomen ausgeschlossen werden kann.

Ebenso konnten die Markersätze teilweise erfolgreich zur Kartierung von Modifiern des *Dll1* Gens auf einem unbekanntem genetischen Hintergrund eingesetzt werden [155].

Allgemein kann festgehalten werden, dass die Identifikation der ursächlichen Mutation durch die genomweite Kopplungsanalyse im Hochdurchsatz stark beschleunigt wurde. Die teilweise sehr starke Kopplung zwischen einem Marker und der ENU-induzierten Mutation legt desweiteren den Schluß nahe, dass die Variablen bezüglich der Anzahl und der Verteilung der Marker korrekt gewählt wurden.

4.3. Gesamtkonzeption

4.3.1. Schlussfolgerungen

Die genomweite Kopplungsanalyse von Mausmutanten durch die Genotypisierung von Längenpolymorphismen (RFLP, Minisatelliten, Mikrosatelliten) wird nach und nach von der SNP-Genotypisierung abgelöst. In den letzten Jahren gab es große Fortschritte, was die Kosten und den Durchsatz einer SNP-basierten Genotypisierung betrifft.

Im Bereich der Humangenetik wurden schon vor längerer Zeit die theoretischen Grundlagen für eine genomweite Genotypisierung von SNPs geschaffen [97] und entsprechende Datenbanken und bioinformatische Techniken standen zur Verfügung.

Die Entwicklung geeigneter Datenbanken für die Genotypisierung von Mäusen, einem der wichtigsten Modelltiere der Säugetiergenetik, rückte dabei, auch aufgrund der relativ spärlichen Datenlage, zunächst in den Hintergrund. Erst in jüngster Zeit, mit Abschluß der großen Mausgenotypisierungs- und Resequenzierungsprojekte, sind diese Werkzeuge verfügbar geworden [107, 207].

Dank dieser aus der öffentlichen Hand stammenden Informationsquellen und den kürzlich von Firmen wie Celera und Perlegen zur Verfügung gestellten Daten ist es mittlerweile möglich, auf ein Reservoir von 47 Millionen Genotypdaten an knapp 5 Millionen Loci (dbSNP 126) bei der Erstellung neuer Markersätze zurückzugreifen. ARTS und MyGenotype sind momentan die einzigen öffentlich verfügbaren Programme, mit denen ein Gesamtprojekt für die Genotypisierung von nicht-humanen Proben geplant, durchgeführt und ausgewertet werden kann.

Gegenüber anderen kommerziell erhältlichen und öffentlich verfügbaren Markersätzen und Programmen bietet das Konzept von ARTS den Vorteil, individuelle Markersätze auf einfachem Wege direkt aus den Daten der dbSNP zu generieren. Diese Vorgehensweise hat sich im MEP bewährt und kann als Konzept für andere Institute eingesetzt werden, die ebenfalls einen genomweiten Ansatz bei der ENU-Mutagenese verfolgen, beispielsweise das ENU Mutagenesis Programme am MRC in Harwell, Großbritannien [135]. Mit MyGenotype ist es möglich, die mit diesem Markersatz erhaltenen experimentellen Genotypdaten zu verwalten und auszuwerten. Es wurde hierbei auf eine möglichst einfache Bedienbarkeit und die Benutzung mit anderen Auswertungsprogrammen Wert gelegt.

Die hier vorgestellte Genotypisierungspipeline bestehend aus automatisierter DNA-Extraktion, ARTS, der SNP-Genotypisierung mittels MALDI-TOF MS und der Auswertung durch MyGenotype hat sich als integraler Bestandteil der Charakterisierung von ENU-induzierten Mausmutanten bewährt, wie aus den zahlreichen bereits publizierten oder sich in der Publikation befindlichen wissenschaftlichen Artikeln hervorgeht (vgl. Abschnitt 5).

4.3.2. Ausblick

Im Hinblick auf die neuen Fragestellungen wie zum Beispiel Gen-Umwelt-Interaktionen und die Analyse von QTLs bietet die ENU-Mutagenese in Verbindung mit dem hier vorgestellten Markersatz interessante Optionen.

So zeigen zum Beispiel F₁ Hybriden aus Verpaarungen zwischen ENU-induzierten Mutanten auf dem C3HeB/FeJ-Hintergrund und dem C57BL/6J Stamm nicht immer den auf dem C3HeB/FeJ Hintergrund dominanten Phänotyp. Es muss also mindestens ein genomisches Element auf dem C57BL/6J Hintergrund existieren, das den Phänotyp dominant modifiziert (Modifier).

Eine Rückkreuzung auf das mutante Elterntier oder eine Geschwisterverpaarung wird den Effekt mit einer von der Zahl der Modifier abhängigen Wahrscheinlichkeit wieder zum Vorschein bringen. Bei einem Kartierungsexperiment werden im Falle nur eines Modifiers allerdings zwei Loci Kopplung zeigen. Es existiert also ein Locus, auf dem sich die ENU-induzierte Mutation befindet und ein weiterer Locus, auf dem der von C57BL/6J stammende Modifier abwesend ist.

Eine Unterscheidung dieser Loci ist zunächst nicht möglich. Im einfachsten Fall können anhand von Literaturrecherche oder der Analyse von Genexpression Kandidatengene in beiden Loci gefunden werden, von denen eines eine ENU-induzierte Mutation enthält. Alternativ dazu muss zunächst über eine Verpaarung mit konsomen oder rekombinanten Inzuchtstämmen der Locus der Mutation von dem des Modifiers unterschieden werden. Eine Eingrenzung der ENU-induzierten Mutation erfolgt dann über weitere Kopplungsanalysen und einen Kandidatengenansatz. Eine Eingrenzung des Locus durch weitere Rückkreuzungen auf den C3HeB/FeJ Hintergrund ist denkbar.

Die Identifikation des Modifiers kann entweder über weitere Kreuzungsversuche mit rekombinanten Inzuchtstämmen erfolgen oder über die Auskreuzung der ursprünglichen Mutante auf C3HeB/FeJ Hintergrund auf wenige Inzuchtstämme, die verschiedene Allelmuster in der Region des Modifiers haben. Bereits durch den Phänotyp der F₁ Generation sollte es möglich sein, eine Aussage über die Allelstruktur des Modifiers zu treffen. Durch eine *in silico* Kartierung [66, 105] kann der Locus des Modifiers weiter eingegrenzt werden. Zur Stützung einer Kandidatengenhypothese kann auch ein Zugriff auf die GNF Expressionsdatenbank nützlich sein [187].

Für diese Schritte ist eine genomweite Kopplungsanalyse mit dem hier vorgestellten Markersatz möglich. Geeignete SNPs für die Feinkartierung können schnell über das ARTS Frontend gefunden werden. Ebenso kann die ARTS Datenbank mit einem modifizierten Frontend zur Darstellung von Allelstrukturen genutzt werden.

Durch eine weitere Modularisierung und Zusammenlegung von ARTS und dem MyGenotype-System kann ein umfangreiches und flexibles System erstellt werden, das alle Belange von groß angelegten Genotypisierungsprojekten erfüllt. Hier gibt es einige Berührungspunkte zu den in der ENU-Mutagenese [143] und der Deutschen Mauslinik [59] für Zucht- und Phänotypisierungsaufgaben eingesetzten Systemen. Als erster Schritt kann die Verwaltung von Genotyp- und Phänotypdaten zur Reduktion von Fehlerquellen vereinheitlicht werden. Durch eine Möglichkeit, neue Marker und die damit gewonnenen Genotypdaten in die Datenbank einzutragen, könnte der Überblick über die bisherige Grob- und

Feinkartierungsarbeit gewährleistet werden. Ein direktes Nebeneinander von Genotyp-, Phänotyp- und Haplotypdaten würde desweiteren die Planung von QTL Experimenten erleichtern. Ebenfalls wäre es denkbar, vor dem Hintergrund der allgemeinen SNP-Daten aus ARTS und den gewonnenen Genotypdaten aus der MyGenotype-Datenbank neuartige, haplotypbasierte Analyseverfahren zu entwickeln, die einen Einstieg in die komplexe Genetik darstellen.

4.4. SMA004 - Ein Modell für *Zmpste24* Defizienz

4.4.1. Auffindung und Charakterisierung der Mutation

Der Locus der rezessiven ENU-induzierten Mausmutante SMA004 wurde mit der ersten Version des Markersatzes mit einer Auflösung von 30 Mb gefunden und die Kopplung mit einem Marker auf Chromosom 4 durch die Genotypisierung von Mikrosatelliten bestätigt. Durch eine datenbankbasierte Literaturrecherche konnten einige Kandidatengene identifiziert werden, von denen *Zmpste24* als der vielversprechendste Kandidat ausgewählt wurde.

Die gefundene G→A Mutation an der Splice-Donor-Stelle nach Exon 6 in *Zmpste24* führt zur Aktivierung einer kryptischen Splice-Donor-Stelle in Intron 6. Der dadurch verursachte Einschub von 16 bp in die mRNA führt zu einer Verschiebung des Leserahmens und zum Einbau eines vorzeitigen Stoppcodons. Dies konnte anhand einer exonspezifischen PCR mit anschließender Sequenzierung bewiesen werden. Durch dieses vorzeitige Translationsende fehlt das für Zink-Metalloproteasen typische HEXXH Motiv und das verkürzte *Zmpste24* Protein verliert dadurch die Fähigkeit, farnesyliertes Prälamina A posttranslational zu prozessieren, wie anhand eines Prälamina A-spezifischen Western Blots bewiesen werden konnte.

Dieser Mechanismus wurde bereits anhand von zwei Knockout-Mauslinien in der Literatur beschrieben [146, 13]. Zwei unterschiedliche Strategien wurden bei der Erstellung dieser Mauslinien verfolgt.

Die von Pendás *et al.* vorgestellte Mutante ist ein Knockout, der durch den Austausch von Exon 2 und 3 durch eine IRES β -*geo*-Kassette erhalten wurde [146]. Das Transkript dieses Gens wurde nach Codon 39 unterbrochen. Die Lokalisation und Stabilität des verkürzten Transkripts wurde nicht untersucht.

Das in der SMA004 Mauslinie gefundene *Zmpste24* Allel überschneidet sich funktionell insbesondere mit der von Leung *et al.* generierten Mausmutante, da die fünfte Transmembrandomäne von dem 3'-Ende von Exon 5 und dem 5'-Ende von Exon 6 codiert wird. Durch die Verschiebung des Leserahmens nach Exon 6 bleiben vermutlich auch bei dieser Variante fünf der sieben Transmembrandomänen intakt, jedoch wird das HEXXH Motiv nicht mehr codiert.

Prinzipiell konnten alle von Bergo *et al.* in *Zmpste24*^{-/-} Mäusen beobachteten Phänotypen ebenfalls in *Zmpste24*^{SMA004/SMA004} Mäusen gefunden werden. Das allgemeine Erscheinungsbild, die langsame Gewichtszunahme, die kallusartigen Kalzifikationen an den Rippenansätzen, sowie die multiplen Knochenbrüche und der damit verbundene Abbau des Os zygomaticum im Bereich des Schädels konnten in *Zmpste24*^{SMA004/SMA004} Mäusen reproduziert werden.

Der Phänotyp der SMA004 Mauslinie scheint auch mit der von Pendás *et al.* beschriebenen Knockout

Mutante vergleichbar zu sein. Vor allem die gezeigte Gewichtskurve stimmt mit dem in dieser Studie erhaltenen Ergebnis überein. Die von Bergo *et al.* beschriebene Alopecia, die im Alter von ca. 15 Monaten bei *Zmpste24*^{+/-} Mäusen auftritt, konnte über den Zeitraum der Studie bei *Zmpste24*^{SMA004/+} Tieren nicht beobachtet werden. Dieser Effekt ist eventuell eine Folge des gemischten genetischen Hintergrundes der *Zmpste24* Knockout Mäuse ($\approx 50\%$ C57BL/6 und $\approx 50\%$ 129/SvJae).

4.4.2. Einfluß der Mutation auf den Phänotyp

Ein wesentlicher Aspekt der unvollständigen Prozessierung von Prälamina A zu reifem Lamin A ist der Verlust einer geordneten Zellkernstruktur [188, 45, 32, 124]. Damit einhergehend sind Störungen der Expression von Genen zu erwarten, die direkt oder indirekt von einer korrekten Anordnung der Laminfilamente abhängen [199]. Zwei alternative Modelle zur Beschreibung der Ursache dieser Störungen sind hierbei denkbar.

Erstens ist zu erwarten, dass durch die Deformation des Zellkerns und die inkorrekte Verteilung von Hetero- und Euchromatin [38] die Transkription einer zufälligen Zahl von Genen in einer betroffenen Zelle in unregelmäßiger Weise verläuft und dadurch letztendlich die Apoptose als Kontrollantwort ausgelöst wird. Hierfür spricht eine Studie, in der beschrieben wird, dass das Tumorsuppressorprotein p53 in Organen von *Lmna*- oder *Zmpste24*-defizienten Mäusen hyperaktiviert wird [199]. Alternativ zu den in dieser Studie von Varela *et al.* getroffenen Schlußfolgerungen kann diese Hyperaktivierung und die damit verbundene Aktivierung von Zielgenen von p53 nicht nur als eine Antwort auf einen durch die Akkumulation von Prälamina A instabilisierten Zellkern verstanden werden, sondern auch als Antwort auf eine zunehmend deregulierte Expression verschiedener Gene. Gegen diese Hypothese spricht jedoch, dass die Symptome der Erkrankung wenig Varianz zeigen. Auch eine gelegentliche Bildung von Tumoren in *Zmpste24*^{-/-} Mäusen, wie sie beispielsweise durch eine zufällige Überexpression von Onkogenen zu erwarten wäre, wurde weder in dieser Studie beobachtet noch in anderen Publikationen erwähnt.

Zweitens kann die Hypothese aufgestellt werden, dass nur die Expression der Gene gestört ist, die direkt oder indirekt von der Interaktion mit Laminfilamenten abhängen. In der Literatur finden sich viele Studien, in denen Interaktionen von Lamin A mit verschiedenen Komponenten des Zellkerns, des Chromatins und mit Transkriptionsfaktoren beschrieben werden. Als Hinweis auf die Bedeutung von Lamin A Interaktionen kann die Steuerung des Zellzyklus durch die Bindung von Faktor pRB durch die nukleäre Lamina gesehen werden [86]. Eine Folge des Verlusts dieser Interaktionen wären vermutlich Expressionsänderungen, die sich in einer geordneten Weise vollziehen. Gegen diese Hypothese sprechen die Ergebnisse von Fong *et al.* [53, 52]. In diesen zeigte es sich, dass durch die Reduktion von Prälamina A in der Zellkernmembran oder durch die vollständige Abwesenheit von Prälamina A beziehungsweise reifem Lamin A, und damit eines Großteils Laminfilamente, die Symptome der Defizienz von *Zmpste24* gemildert oder eliminiert werden können. Eine Folge davon wäre, dass die Abwesenheit von intakten Laminfilamenten keinen entscheidenden Einfluß auf die Funktion einer Zelle hat.

Im Rahmen dieser Arbeit wurde die Expression verschiedener Zielgene untersucht, um beide Hypothe-

4. DISKUSSION

sen zu testen. Es wurden speziell solche Gene ausgewählt, von denen bekannt ist, dass sie unter der Kontrolle von Transkriptionsfaktoren stehen, die nachgewiesenermaßen mit Lamin A interagieren. Falls ein konstanter Expressionsunterschied dieser Gene im Gewebe von *Zmpste24*^{SMA004/SMA004} Mäusen gefunden wird, ist davon auszugehen, dass dieser auf einen Verlust der Interaktion zwischen einem Transkriptionsfaktor und Lamin A zurückzuführen ist. Unterschiede in der Expression sowohl im Vergleich zu den Proben vom *Zmpste24*^{+/+} Mäusen als auch zwischen verschiedenen Proben von *Zmpste24*^{SMA004/SMA004} Mäusen würden auf einen Verlust der Expressionskontrolle hindeuten.

Das Muskelgewebe ist stark von Mutationen in *Lmna* und *Zmpste24* betroffen [99], daher wurden zunächst Studien mit Proben des Muskelgewebes durchgeführt. Es zeigte sich, dass von sieben untersuchten Genen nur *Pcsk9* in den Muskeln von *Zmpste24*^{SMA004/SMA004} Mäusen im Vergleich zu Wildtypkontrollen des selben Alters differenziell exprimiert war. Die Expression der restlichen Gene war nicht signifikant verändert. Zur Überprüfung der Annahme, dass die Überexpression von *Pcsk9* auf eine fehlende Interaktion mit Lamin A zurückzuführen ist, wurde die Expression dieses Gens in Leberproben bestimmt. Hier wurde jedoch kein signifikanter Expressionsunterschied gefunden.

Der allgemeine Nachweis einer Kopplung der Genexpression an eine Interaktion von Transkriptionsfaktoren und Lamin A konnte somit nicht erbracht werden. Auch läßt sich anhand der Ergebnisse kein Hinweis darauf finden, dass die Regulation der Expression verschiedener Gene durch die Abwesenheit von intakten Laminfilamenten zufällig gestört ist, da mit einer Ausnahme alle untersuchten Gene, inklusive dreier Housekeeping-Gene, keine Expressionsänderung im Vergleich zu *Zmpste24*^{+/+} Proben zeigen. Die Interaktion von Lamin A mit verschiedenen Transkriptionsfaktoren scheint eine geringere Rolle zu spielen als angenommen, da keine signifikante Veränderung der Genexpression von Zielgenen dieser Transkriptionsfaktoren gefunden oder bestätigt werden konnte.

Allgemein bekräftigen diese Ergebnisse kürzlich publizierte Daten, nach denen Lamin A nicht essentiell für die korrekte Funktion des Zellkerns zu sein scheint [53, 52] und der Phänotyp durch den ganzen oder teilweisen Entzug von unreifem oder reifem Lamin A abgemildert wird.

Dennoch bleibt die Frage offen, warum *Pcsk9* im Muskelgewebe von *Zmpste24*^{SMA004/SMA004} überexprimiert ist. Dieses Gen spielt eine wichtige Rolle bei der Aufnahme von Low-Density-Lipoprotein (LDL), die Funktionsweise des Genprodukts ist jedoch nicht endgültig verstanden [7]. Eine Überexpression von *Pcsk9* induziert die Degradation des LDL-Rezeptors (LDLR) [118]. Die Verminderung der Aufnahme von LDL im Muskel sollte zu einer Erhöhung der LDL-Konzentration im Blutplasma und letztendlich zu Hypercholesterinämie führen. Die zu erwartende Konsequenz einer erhöhten LDL Konzentration wäre die Drosselung der Expression von *Pcsk9* und damit ein Anstieg von LDLR in der Leber, um diese Differenz wieder auszugleichen. Diese Expressionsänderung kann jedoch nicht beobachtet werden.

Unter der Annahme, dass die Expression von *Pcsk9* von der Interaktion von *Srebf1* und Lamin A abhängt, wäre zu erwarten, dass *Pcsk9* in allen Geweben überexprimiert ist. Umgekehrt bedeutet dies, dass die Expression von *Pcsk9* nicht effizient gedrosselt werden kann. Es kann daher angenommen werden, dass die Expression in der Leber nur scheinbar normal ist, da die Expression von *Pcsk9* vermutlich von anderen, weniger wirksamen Faktoren als *Srebf1* gedrosselt wird und nicht auf ein von der normalen Expression unterscheidbares Niveau fallen kann.

Alternativ dazu wäre denkbar, dass die Expression von *Pcsk9* nicht von einer Interaktion von *Srebf1* und Lamin A abhängt und es sich bei der Überexpression von *Pcsk9* im Muskel um einen durch die Störung der Zellkernarchitektur verursachten Sekundäreffekt handelt.

Weitere Studien der Interaktion von Lamin A und *Srebf1* sowie die Aufklärung der Rolle von *Pcsk9* in *Zmpste24*-defizienten Muskelzellen können Aufschluss über den zugrunde liegenden Mechanismus liefern und eventuell zu einem weiteren Verständnis des Cholesterin-Haushalts und zur Funktionsweise von *PSCK9* beitragen.

4.4.3. Ausblick

Der Phänotyp von *Zmpste24*^{SMA004/SMA004} Mäusen wird in einem Alter von 4-6 Wochen sichtbar. Dieser Phänotyp muss eine Folge von Änderungen des Expressionsmusters von einer Vielzahl von Genen sein, auch wenn die primäre Ursache hierfür nicht endgültig geklärt ist. Entweder beruht die Ursache auf einem Effekt, der durch Störungen der Bindung zwischen Filamenten aus reifem Lamin A und dessen Interaktionspartnern hervorgerufen wird, oder auf der Desorganisation der Zellkernmembran, in die farnesyliertes Prälamina A akkumuliert ist.

Eine Beobachtung der Genexpression zu verschiedenen Zeitpunkten bis zur offensichtlichen Manifestation des Phänotyps könnte einen tieferen Einblick in die eigentliche Ursache der progerieartigen Symptome von *Zmpste24*^{-/-} Mutanten geben.

Die Analyse dieser Expressionsdaten könnte Aufschluss über mögliche neue Interaktionspartner von Lamin A geben. Über Sequenzvergleiche der Promotorregionen von differentiell exprimierten Genen können eventuell Bindungsstellen für Transkriptionsfaktoren gefunden werden, die mit Lamin A interagieren. Eine Herausforderung wird es hierbei sein, die primär von einer Interaktion mit Lamin A abhängigen Gene von sekundär regulierten Genen zu unterscheiden. Die bisher bekannten Interaktionspartner von Lamin A und die durchgeführten Analysen können hierbei zum Training eines Modells zur Unterscheidung solcher Gene eingesetzt werden.

Ein weiterer Ansatz zum besseren Verständnis der Rolle von *Zmpste24* im Organismus wäre eine detaillierte Analyse der Entstehung der Lipodystrophie in *Zmpste24*^{SMA004/SMA004} Mäusen. Der Transkriptionsfaktor *Srebf1* und dessen Interaktion mit Lamin A könnte hierbei eine Schlüsselrolle spielen, da bekannt ist, dass *Srebf1* sowohl eine wichtige Rolle in der Differenzierung von Adipozyten als auch in der Cholesterinbiosynthese spielt [93].

5. Liste eigener Publikationen

Artikel in wissenschaftlichen Zeitschriften

M. Klaften and M. Hrabé de Angelis. ARTS: a web-based tool for the set-up of high throughput genome-wide mapping panels for the SNP genotyping of mouse mutants. *Nucleic Acids Res*, 33(Web Server issue): W496-W500, Jul 2005.

M. Klaften, K. Abe, S. Wagner, D. Soewarto, H. Fuchs, M. Hrabé de Angelis. A novel concept for high throughput genetic mapping of mouse mutants by SNP genotyping. (in Vorbereitung)

F. Runkel, **M. Klaften**, K. Koch, V. Böhnert, H. Büssow, H. Fuchs, T. Franz, and M. Hrabé de Angelis. Morphological and molecular characterisation of two novel Krt2-6g mutations: Krt2-6grco12 and Krt2-6grco13. *Mamm Genome*, 17(12): 1172-1182, Dec 2006.

H. Seedorf, **M. Klaften**, F. Issmail, H. Fuchs, U. Seedorf, and M. Hrabé de Angelis. Identification of a nonsense mutation in the enamel gene in a mouse model for codominant Amelogenesis imperfecta. *J Dent Res* (akzeptiert)

I. Rubio-Aliaga, D. Soewarto, S. Wagner, **M. Klaften**, H. Fuchs, S. Kalaydjiev, D. H. Busch, M. Klempt, B. Rathkolb, E. Wolf, K. Abe, S. Zeiser, G. Przemeczek, J. Beckers, and M. Hrabé de Angelis. A genetic screen for modifiers of Delta1 dependent signaling function in the mouse. *Genetics*, 175(3): 1451-1463, Mar 2007

N. Herbach, B. Rathkolb, E. Kemter, L. Pichl, **M. Klaften**, M. Hrabé de Angelis, E. Wolf, B. Aigner, and R. Wanke. Early onset diabetes mellitus and severe loss of beta cells in heterozygous Ins2 mutant mice lacking the insulin 2 intrachain disulfide bond. *Diabetes*, 56(5): 1268-1276, May 2007

B. Aigner, B. Rathkolb, N. Herbach, E. Kemter, C. Schessl, **M. Klaften**, M. Klempt, M. Hrabé de Angelis, R. Wanke, and E. Wolf. Screening for increased plasma urea levels in a large-scale ENU mouse mutagenesis project reveals kidney disease models. *Am J Physiol Renal Physiol*, 292(5): F1560-F1567, May 2007

T. S. Lisse, C. Cohrs, F. Thiele, K. Abe, W. Hans, **M. Klaften**, M. Schulz, H. Fuchs, S. H. Ralston, and M. Hrabé de Angelis. Characterization and mapping of ALI034: A new ENU-derived murine model for osteoarthritis. *Calcif Tissue Int* (akzeptiert)

Patente

M. Klaften, T. Hoefler, and M. Hrabé de Angelis. Rapid Screening of Clones for the Generation of Genetically Modified Organisms and DNA Containing Systems Based on Genetic Markers. European Patent Office (in Vorbereitung)

6. Abkürzungsverzeichnis

3-HPA	3-Hydroxypicolinic acid
AIL	Advanced Intercross Line
ARMS	Amplification Refractory Mutation System
ARTS	Advanced Retrieval Tool for SNPs
BMC	Bone Mineral Content
BMD	Bone Mineral Density
cDNA	Complementary DNA
CT	Computertomographie
ddNTP	Dideoxyribonukleosidtriphosphat
DNA	Deoxyribonucleic Acid
dNTP	Desoxyribonukleosidtriphosphat
DWP	Deep Well Plate
EDTA	Ethylene-diamine-tetraacetic acid
EMS	Ethylmethansulfonat
ENU	<i>N</i> -Ethyl- <i>N</i> -nitrosourea
ES Zelle	Embryonale Stammzelle
Face-1	Farnesylated protein-converting enzyme 1
Fdft1	Farnesyldiphosphat Farnesyltransferase I
Gli3	GLI-Kruppel family member GLI3
HGPS	Hutchinson-Gilford-Progerie-Syndrom
hME	Homologe Massenextension
Icmt	Isoprenylcysteincarboxylmethyltransferase
LDL	Low-Density-Lipoprotein
LDLR	Low-Density-Lipoprotein Receptor
Lmna	Lamin A
Lmnb1	Lamin B1
Lmnb2	Lamin B2
MAD	Mandibuloakrale Dysplasie
MALDI	Matrix assisted laser desorption/ionization
MEF	Murine embryonale Fibroblasten
MEP	Münchener ENU Mausmutagenese-Projekt
MHC	Major Histocompatibility Complex

6. ABKÜRZUNGSVERZEICHNIS

mRNA	Messenger RNA
OMIM	Online Mendelian Inheritance in Man
PAGE	Polyacrylamide Gel Electrophoresis
PAK	Primärer Antikörper
pBMD	peripheral Bone Mineral Density
PCR	Polymerase Chain Reaction
Pcsk9	Proprotein convertase subtilisin/kexin type 9
pDEXA	peripheral Dual Energy X-Ray Absorptiometry
pRB	Retinoblastomprotein
QTL	Quantitative Trait
RD	Restriktive Dermatopathie
RFLP	Restriktionsfragmentlängenpolymorphismus
RI Stamm	Rekombinanter Inzuchtstamm
RNA	Ribonucleic Acid
RT	Raumtemperatur
RT-PCR	Reverse Transkriptase PCR
SAK	Sekundärer Antikörper
SAP	Shrimp Alkaline Phosphatase
SDS	Sodium Dodecylsulfate
SNP	Single Nucleotide Polymorphism
Srebf1	Sterol regulatory element binding transcription factor 1
SSCP	Single Strand Length Polymorphism
STR	Short Tandem Repeat
TF	Transkriptionsfaktor
TOF-MS	Time-of-flight mass spectrometry
TZ	Tischzentrifuge
VNTR	Variable Number Tandem Repeat
Zmpste24	Zink-Metalloprotease STE24 homolog

7. Literaturverzeichnis

- [1] U. Aebi, J. Cohn, L. Buhle, and L. Gerace. The nuclear lamina is a meshwork of intermediate-type filaments. *Nature*, 323(6088):560–564, 1986.
- [2] A. K. Agarwal, J.-P. Fryns, R. J. Auchus, and A. Garg. Zinc metalloproteinase, ZMPSTE24, is mutated in mandibuloacral dysplasia. *Hum Mol Genet*, 12(16):1995–2001, Aug 2003.
- [3] B. Aigner, B. Rathkolb, N. Herbach, E. Kemter, C. Schessl, M. Klaften, M. Klempt, M. H. de Angelis, R. Wanke, and E. Wolf. Screening for increased plasma urea levels in a large-scale ENU mouse mutagenesis project reveals kidney disease models. *Am J Physiol Renal Physiol*, 292(5):F1560–F1567, May 2007.
- [4] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *J Mol Biol*, 215(3):403–410, Oct 1990.
- [5] V. Arranz, C. Dreuillet, P. Crisanti, J. Tillit, M. Kress, and M. Ernoult-Lange. The zinc finger transcription factor, MOK2, negatively modulates expression of the interphotoreceptor retinoid-binding protein gene, IRBP. *J Biol Chem*, 276(15):11963–11969, Apr 2001.
- [6] M. M. Aslanian and V. A. Smirnova. [Analysis of mutations induced by alkylating compounds in *Drosophila* sperm. I. Character of visible changes induced by N-nitrosoethylurea and ethylmethanesulfonate in a succession of *Drosophila* generations]. *Genetika*, 13(4):646–654, 1977.
- [7] A. D. Attie. The mystery of PCSK9. *Arterioscler Thromb Vasc Biol*, 24(8):1337–1339, Aug 2004.
- [8] F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. *Short Protocols in Molecular Biology*. John Wiley and Sons, 1999.
- [9] P. Avner, L. Amar, L. Dandolo, and J. L. Guénet. Genetic analysis of the mouse using interspecific crosses. *Trends Genet*, 4(1):18–23, Jan 1988.
- [10] M. R. D. Barletta, E. Ricci, G. Galluzzi, P. Tonali, M. Mora, L. Morandi, A. Romorini, T. Voit, K. H. Orstavik, L. Merlini, C. Trevisan, V. Biancalana, I. Housmanowa-Petrusewicz, S. Bione, R. Ricotti, K. Schwartz, G. Bonne, and D. Toniolo. Different mutations in the LMNA gene cause autosomal dominant and autosomal recessive Emery-Dreifuss muscular dystrophy. *Am J Hum Genet*, 66(4):1407–1412, Apr 2000.
- [11] J. A. Beck, S. Lloyd, M. Hafezparast, M. Lennon-Pierce, J. T. Eppig, M. F. Festing, and E. M. Fisher. Genealogies of mouse inbred strains. *Nat Genet*, 24(1):23–25, Jan 2000.
- [12] L. Bengtsson and K. L. Wilson. Multiple and surprising new functions for emerin, a nuclear membrane protein. *Curr Opin Cell Biol*, 16(1):73–79, Feb 2004.
- [13] M. O. Bergo, B. Gavino, J. Ross, W. K. Schmidt, C. Hong, L. V. Kendall, A. Mohr, M. Meta, H. Genant, Y. Jiang, E. R. Wisner, N. V. Bruggen, R. A. D. Carano, S. Michaelis, S. M. Griffey, and

7. LITERATURVERZEICHNIS

- S. G. Young. Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. *Proc Natl Acad Sci U S A*, 99(20):13049–13054, Oct 2002.
- [14] O. Binda, J.-S. Roy, and P. E. Branton. RBP1 family proteins exhibit SUMOylation-dependent transcriptional repression and induce cell growth inhibition reminiscent of senescence. *Mol Cell Biol*, 26(5):1917–1931, Mar 2006.
- [15] S. Bione, E. Maestrini, S. Rivella, M. Mancini, S. Regis, G. Romeo, and D. Toniolo. Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. *Nat Genet*, 8(4):323–327, Dec 1994.
- [16] J. A. Blake, J. E. Richardson, M. T. Davisson, and J. T. Eppig. The Mouse Genome Database (MGD). A comprehensive public resource of genetic, phenotypic and genomic data. The Mouse Genome Informatics Group. *Nucleic Acids Res*, 25(1):85–91, Jan 1997.
- [17] V. C. Bode. Ethylnitrosourea mutagenesis and the isolation of mutant alleles for specific genes located in the T region of mouse chromosome 17. *Genetics*, 108(2):457–470, Oct 1984.
- [18] F. Bonhomme, F. Benmehdi, J. Britton-Davidian, and S. Martin. [Genetic analysis of interspecific crosses *Mus musculus* L. x *Mus spretus* Lataste: linkage of Adh-1 with Amy-1 on chromosome 3 and Es-14 with Mod-1 on chromosome 9]. *C R Seances Acad Sci D*, 289(6):545–548, Oct 1979.
- [19] G. Bonne, M. R. D. Barletta, S. Varnous, H. M. Bécane, E. H. Hammouda, L. Merlini, F. Muntoni, C. R. Greenberg, F. Gary, J. A. Urtizberea, D. Duboc, M. Fardeau, D. Toniolo, and K. Schwartz. Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat Genet*, 21(3):285–288, Mar 1999.
- [20] D. Botstein, R. L. White, M. Skolnick, and R. W. Davis. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32(3):314–331, May 1980.
- [21] V. L. Boyartchuk, M. N. Ashby, and J. Rine. Modulation of Ras and a-factor function by carboxyl-terminal proteolysis. *Science*, 275(5307):1796–1800, Mar 1997.
- [22] J. L. Broers, B. M. Machiels, H. J. Kuijpers, F. Smedts, R. van den Kieboom, Y. Raymond, and F. C. Ramaekers. A- and B-type lamins are differentially expressed in normal human tissues. *Histochem Cell Biol*, 107(6):505–517, Jun 1997.
- [23] J. L. V. Broers, H. J. H. Kuijpers, C. Ostlund, H. J. Worman, J. Endert, and F. C. S. Ramaekers. Both lamin A and lamin C mutations cause lamina instability as well as loss of internal nuclear lamin organization. *Exp Cell Res*, 304(2):582–592, Apr 2005.
- [24] K. W. Broman, H. Wu, S. Sen, and G. A. Churchill. R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19(7):889–890, May 2003.
- [25] S. D. Brown and R. Balling. Systematic approaches to mouse mutagenesis. *Curr Opin Genet Dev*, 11(3):268–273, Jun 2001.
- [26] K. H. Buetow, M. Edmonson, R. MacDonald, R. Clifford, P. Yip, J. Kelley, D. P. Little, R. Strausberg, H. Koester, C. R. Cantor, and A. Braun. High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci U S A*, 98(2):581–584, Jan 2001.

- [27] T. Carter and D. Falconer. Stocks for detecting linkage in the mouse and the theory of their design. *J. Genet.*, 50:307–323, 1951.
- [28] F. Caux, E. Dubosclard, O. Lascols, B. Buendia, O. Chazouillères, A. Cohen, J.-C. Courvalin, L. Laroche, J. Capeau, C. Vigouroux, and S. Christin-Maitre. A new clinical condition linked to a novel mutation in lamins A and C with generalized lipoatrophy, insulin-resistant diabetes, disseminated leukomelanodermic papules, liver steatosis, and cardiomyopathy. *J Clin Endocrinol Metab*, 88(3):1006–1013, Mar 2003.
- [29] L. Chen, L. Lee, B. A. Kudlow, H. G. D. Santos, O. Sletvold, Y. Shafeghati, E. G. Botha, A. Garg, N. B. Hanson, G. M. Martin, I. S. Mian, B. K. Kennedy, and J. Oshima. LMNA mutations in atypical Werner’s syndrome. *Lancet*, 362(9382):440–445, Aug 2003.
- [30] G. A. Churchill, D. C. Airey, H. Allayee, et al. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet*, 36(11):1133–1137, Nov 2004.
- [31] E. L. Coghill, A. Hugill, N. Parkinson, C. Davison, P. Glenister, S. Clements, J. Hunter, R. D. Cox, and S. D. M. Brown. A gene-driven approach to the identification of ENU mutants in the mouse. *Nat Genet*, 30(3):255–256, Mar 2002.
- [32] M. Columbaro, C. Capanni, E. Mattioli, G. Novelli, V. K. Parnaik, S. Squarzoni, N. M. Maraldi, and G. Lattanzi. Rescue of heterochromatin organization in Hutchinson–Gilford progeria by drug treatment. *Cell Mol Life Sci*, 62(22):2669–2678, Nov 2005.
- [33] D. Concepcion, K. L. Seburn, G. Wen, W. N. Frankel, and B. A. Hamilton. Mutation rate and predicted phenotypic target sizes in ethylnitrosourea-treated mice. *Genetics*, 168(2):953–959, Oct 2004.
- [34] D. P. Corrigan, D. Kuszczak, A. E. Rusinol, D. P. Thewke, C. A. Hrycyna, S. Michaelis, and M. S. Sinensky. Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. *Biochem J*, 387(Pt 1):129–138, Apr 2005.
- [35] P. Costet, B. Cariou, G. Lambert, F. Lalanne, B. Lardeux, A.-L. Jarnoux, A. Grefhorst, B. Staels, and M. Krempf. Hepatic PCSK9 expression is regulated by nutritional status via insulin and sterol regulatory element-binding protein 1c. *J Biol Chem*, 281(10):6211–6218, Mar 2006.
- [36] A. Darvasi and M. Soller. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics*, 141(3):1199–1207, Nov 1995.
- [37] T. Dechat, S. Vlcek, and R. Foisner. Review: lamina-associated polypeptide 2 isoforms and related proteins in cell cycle-dependent nuclear structure dynamics. *J Struct Biol*, 129(2-3):335–345, Apr 2000.
- [38] E. Delbarre, M. Tramier, M. Coppey-Moisan, C. Gaillard, J.-C. Courvalin, and B. Buendia. The truncated prelamin A in Hutchinson–Gilford progeria syndrome alters segregation of A-type and B-type lamin homopolymers. *Hum Mol Genet*, 15(7):1113–1122, Apr 2006.
- [39] W. F. Dietrich, J. Miller, R. Steen, M. A. Merchant, D. Damron-Boles, Z. Husain, R. Dredge, M. J. Daly, K. A. Ingalls, and T. J. O’Connor. A comprehensive genetic map of the mouse genome. *Nature*, 380(6570):149–152, Mar 1996.

7. LITERATURVERZEICHNIS

- [40] K. DiPetrillo, S.-W. Tsaih, S. Sheehan, C. Johns, P. Kelmenson, H. Gavras, G. A. Churchill, and B. Paigen. Genetic analysis of blood pressure in C3H/HeJ and SWR/J mice. *Physiol Genomics*, 17(2):215–220, Apr 2004.
- [41] T. Doetschman, R. G. Gregg, N. Maeda, M. L. Hooper, D. W. Melton, S. Thompson, and O. Smithies. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature*, 330(6148):576–578, 1987.
- [42] C. Dreuillet, J. Tillit, M. Kress, and M. Ernoult-Lange. In vivo and in vitro interaction between human transcription factor MOK2 and nuclear lamin A/C. *Nucleic Acids Res*, 30(21):4634–4642, Nov 2002.
- [43] H. Druckrey, S. Ivanković, and R. Preussmann. Teratogenic and carcinogenic effects in the offspring after single injection of ethylnitrosourea to pregnant rats. *Nature*, 210(43):1378–1379, Jun 1966.
- [44] J. T. Eppig, C. J. Bult, J. A. Kadin, et al. The Mouse Genome Database (MGD): from genes to mice—a community resource for mouse biology. *Nucleic Acids Res*, 33(Database issue):D471–D475, Jan 2005.
- [45] M. Eriksson, W. T. Brown, L. B. Gordon, M. W. Glynn, J. Singer, L. Scott, M. R. Erdos, C. M. Robbins, T. Y. Moses, P. Berglund, A. Dutra, E. Pak, S. Durkin, A. B. Csoka, M. Boehnke, T. W. Glover, and F. S. Collins. Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature*, 423(6937):293–298, May 2003.
- [46] H. V. Esch, A. K. Agarwal, P. Debeer, J.-P. Fryns, and A. Garg. A homozygous mutation in the lamin A/C gene associated with a novel syndrome of arthropathy, tendinous calcinosis, and progeroid features. *J Clin Endocrinol Metab*, 91(2):517–521, Feb 2006.
- [47] B. Ewing and P. Green. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*, 8(3):186–194, Mar 1998.
- [48] B. Ewing, L. Hillier, M. C. Wendl, and P. Green. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*, 8(3):175–185, Mar 1998.
- [49] D. Fatkin, C. MacRae, T. Sasaki, M. R. Wolff, M. Porcu, M. Frenneaux, J. Atherton, H. J. Vidaillet, S. Spudich, U. D. Girolami, J. G. Seidman, C. Seidman, F. Muntoni, G. Muehle, W. Johnson, and B. McDonough. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med*, 341(23):1715–1724, Dec 1999.
- [50] D. Z. Fisher, N. Chaudhary, and G. Blobel. cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. *Proc Natl Acad Sci U S A*, 83(17):6450–6454, Sep 1986.
- [51] J. Flint, W. Valdar, S. Shifman, and R. Mott. Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet*, 6(4):271–286, Apr 2005.
- [52] L. G. Fong, D. Frost, M. Meta, X. Qiao, S. H. Yang, C. Coffinier, and S. G. Young. A protein farnesyltransferase inhibitor ameliorates disease in a mouse model of progeria. *Science*, 311(5767):1621–1623, Mar 2006.

- [53] L. G. Fong, J. K. Ng, J. Lammerding, T. A. Vickers, M. Meta, N. Coté, B. Gavino, X. Qiao, S. Y. Chang, S. R. Young, S. H. Yang, C. L. Stewart, R. T. Lee, C. F. Bennett, M. O. Bergo, and S. G. Young. Prelamin A and lamin A appear to be dispensable in the nuclear lamina. *J Clin Invest*, 116(3):743–752, Mar 2006.
- [54] L. G. Fong, J. K. Ng, M. Meta, N. Coté, S. H. Yang, C. L. Stewart, T. Sullivan, A. Burghardt, S. Majumdar, K. Reue, M. O. Bergo, and S. G. Young. Heterozygosity for Lmna deficiency eliminates the progeria-like phenotypes in Zmpste24-deficient mice. *Proc Natl Acad Sci U S A*, 101(52):18111–18116, Dec 2004.
- [55] J. V. Frangioni and B. G. Neel. Use of a general purpose mammalian expression vector for studying intracellular protein targeting: identification of critical residues in the nuclear lamin A/C nuclear localization signal. *J Cell Sci*, 105 (Pt 2):481–488, Jun 1993.
- [56] J. M. Freije, P. Blay, A. M. Pendás, J. Cadiñanos, P. Crespo, and C. López-Otín. Identification and chromosomal location of two human genes encoding enzymes potentially involved in proteolytic maturation of farnesylated proteins. *Genomics*, 58(3):270–280, Jun 1999.
- [57] H. Fuchs, K. Schughart, E. Wolf, R. Balling, and M. Hrabé de Angelis. Screening for dysmorphological abnormalities—a powerful tool to isolate new mouse mutants. *Mamm Genome*, 11(7):528–530, Jul 2000.
- [58] K. Fujimura-Kamada, F. J. Nouvet, and S. Michaelis. A novel membrane-associated metalloprotease, Ste24p, is required for the first step of NH₂-terminal processing of the yeast a-factor precursor. *J Cell Biol*, 136(2):271–285, Jan 1997.
- [59] V. Gailus-Durner, H. Fuchs, L. Becker, et al. Introducing the German Mouse Clinic: open access platform for standardized phenotyping. *Nat Methods*, 2(6):403–404, Jun 2005.
- [60] M. Giacca, L. Zentilin, P. Norio, S. Diviacco, D. Dimitrova, G. Contreas, G. Biamonti, G. Perini, F. Weighardt, and S. Riva. Fine mapping of a replication origin of human DNA. *Proc Natl Acad Sci U S A*, 91(15):7119–7123, Jul 1994.
- [61] P. Gill, C. P. Kimpton, A. Urquhart, N. Oldroyd, E. S. Millican, S. K. Watson, and T. J. Downes. Automated short tandem repeat (STR) analysis in forensic casework—a strategy for the future. *Electrophoresis*, 16(9):1543–1552, Sep 1995.
- [62] R. D. Goldman, D. K. Shumaker, M. R. Erdos, M. Eriksson, A. E. Goldman, L. B. Gordon, Y. Gruenbaum, S. Khuon, M. Mendez, R. Varga, and F. S. Collins. Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A*, 101(24):8963–8968, Jun 2004.
- [63] D. Gordon, C. Abajian, and P. Green. Consed: a graphical tool for sequence finishing. *Genome Res*, 8(3):195–202, Mar 1998.
- [64] S. G. Gregory, M. Sekhon, J. Schein, et al. A physical map of the mouse genome. *Nature*, 418(6899):743–750, Aug 2002.
- [65] T. J. Griffin and L. M. Smith. Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *Trends Biotechnol*, 18(2):77–84, Feb 2000.

7. LITERATURVERZEICHNIS

- [66] A. Grupe, S. Germer, J. Usuka, D. Aud, J. K. Belknap, R. F. Klein, M. K. Ahluwalia, R. Higuchi, and G. Peltz. In silico mapping of complex disease-related traits in mice. *Science*, 292(5523):1915–1918, Jun 2001.
- [67] H. Gu, J. D. Marth, P. C. Orban, H. Mossmann, and K. Rajewsky. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science*, 265(5168):103–106, Jul 1994.
- [68] I. G. Gut. Automation in genotyping of single nucleotide polymorphisms. *Hum Mutat*, 17(6):475–492, Jun 2001.
- [69] G. H. Hardy. Mendelian proportions in a mixed population. *Science*, 78:49–50, 1908.
- [70] S. M. Hebsgaard, P. G. Korning, N. Tolstrup, J. Engelbrecht, P. Rouzé, and S. Brunak. Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res*, 24(17):3439–3452, Sep 1996.
- [71] R. A. Hegele, H. Cao, D. M. Liu, G. A. Costain, V. Charlton-Menys, N. W. Rodger, and P. N. Durrington. Sequencing of the reannotated LMNB2 gene reveals novel mutations in patients with acquired partial lipodystrophy. *Am J Hum Genet*, 79(2):383–389, Aug 2006.
- [72] N. Herbach, B. Rathkolb, E. Kemter, L. Pichl, M. Klawns, M. H. de Angelis, P. A. Halban, E. Wolf, B. Aigner, and R. Wanke. Dominant-Negative Effects of a Novel Mutated Ins2 Allele Causes Early-Onset Diabetes and Severe beta-Cell Loss in Munich Ins2C95S Mutant Mice. *Diabetes*, 56(5):1268–1276, May 2007.
- [73] F. Hillenkamp, M. Karas, R. C. Beavis, and B. T. Chait. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem*, 63(24):1193A–1203A, Dec 1991.
- [74] S. Hitotsumachi, D. A. Carpenter, and W. L. Russell. Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. *Proc Natl Acad Sci U S A*, 82(19):6619–6621, Oct 1985.
- [75] D. Holtz, R. A. Tanaka, J. Hartwig, and F. McKeon. The CaaX motif of lamin A functions in conjunction with the nuclear localization signal to target assembly to the nuclear envelope. *Cell*, 59(6):969–977, Dec 1989.
- [76] P. Hozák, A. M. Sasseville, Y. Raymond, and P. R. Cook. Lamin proteins form an internal nucleoskeleton as well as a peripheral lamina in human cells. *J. Cell Sci.*, 108 (Pt 2):635–644, Feb 1995.
- [77] M. Hrabé de Angelis and R. Balling. Large scale ENU screens in the mouse: genetics meets genomics. *Mutat Res*, 400(1-2):25–32, May 1998.
- [78] M. Hrabé de Angelis, H. Flaswinkel, H. Fuchs, et al. Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat Genet*, 25(4):444–447, Aug 2000.
- [79] M. Hrabé de Angelis, J. McIntyre, and A. Gossler. Maintenance of somite borders in mice requires the Delta homologue Dll1. *Nature*, 386(6626):717–721, Apr 1997.
- [80] C. J. Hutchison and H. J. Worman. A-type lamins: guardians of the soma? *Nat Cell Biol*, 6(11):1062–1067, Nov 2004.

- [81] O. A. Iakoubova, C. L. Olsson, K. M. Dains, J. Choi, I. Kalcheva, L. G. Bentley, M. Cunanan, D. Hillman, J. Louie, M. Machrus, and D. B. West. Microsatellite marker panels for use in high-throughput genotyping of mouse crosses. *Physiol Genomics*, 3(3):145–148, Sep 2000.
- [82] R. Jaenisch. Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc Natl Acad Sci U S A*, 73(4):1260–1264, Apr 1976.
- [83] G. Jagatheesan, S. Thanumalayan, B. Muralikrishna, N. Rangaraj, A. A. Karande, and V. K. Par naik. Colocalization of intranuclear lamin foci with RNA splicing factors. *J Cell Sci*, 112 (Pt 24):4651–4661, Dec 1999.
- [84] H. Jakel, M. Nowak, A. Helleboid-Chapman, J. Fruchart-Najib, and J.-C. Fruchart. Is apolipoprotein A5 a novel regulator of triglyceride-rich lipoproteins? *Ann Med*, 38(1):2–10, 2006.
- [85] A. J. Jeffreys, V. Wilson, and S. L. Thein. Individual-specific 'fingerprints' of human DNA. *Nature*, 316(6023):76–79, 1985.
- [86] B. R. Johnson, R. T. Nitta, R. L. Frock, L. Mounkes, D. A. Barbie, C. L. Stewart, E. Harlow, and B. K. Kennedy. A-type lamins regulate retinoblastoma protein function by promoting subnuclear localization and preventing proteasomal degradation. *Proc Natl Acad Sci U S A*, 101(26):9677–9682, Jun 2004.
- [87] C. Jurinke, D. van den Boom, C. R. Cantor, and H. Köster. The use of MassARRAY technology for high throughput genotyping. *Adv Biochem Eng Biotechnol*, 77:57–74, 2002.
- [88] M. J. Justice, D. A. Carpenter, J. Favor, A. Neuhauser-Klaus, M. Hrabé de Angelis, D. Soewarto, A. Moser, S. Cordes, D. Miller, V. Chapman, J. S. Weber, E. M. Rinchik, P. R. Hunsicker, W. L. Russell, and V. C. Bode. Effects of ENU dosage on mouse strains. *Mamm Genome*, 11(7):484–488, Jul 2000.
- [89] M. Karas and F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*, 60(20):2299–2301, Oct 1988.
- [90] D. Karolchik, R. Baertsch, M. Diekhans, T. S. Furey, A. Hinrichs, Y. T. Lu, K. M. Roskin, M. Schwartz, C. W. Sugnet, D. J. Thomas, R. J. Weber, D. Haussler, W. J. Kent, and U. of California Santa Cruz. The UCSC Genome Browser Database. *Nucleic Acids Res*, 31(1):51–54, Jan 2003.
- [91] A. Kasprzyk, D. Keefe, D. Smedley, D. London, W. Spooner, C. Melsopp, M. Hammond, P. Rocca-Serra, T. Cox, and E. Birney. EnsMart: a generic system for fast and flexible access to biological data. *Genome Res*, 14(1):160–169, Jan 2004.
- [92] F. Kilic, M. B. Dalton, S. K. Burrell, J. P. Mayer, S. D. Patterson, and M. Sinensky. In vitro assay and characterization of the farnesylation-dependent prelamin A endoprotease. *J Biol Chem*, 272(8):5298–5304, Feb 1997.
- [93] J. B. Kim and B. M. Spiegelman. ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. *Genes Dev*, 10(9):1096–1107, May 1996.
- [94] F. Kirpekar, E. Nordhoff, L. K. Larsen, K. Kristiansen, P. Roepstorff, and F. Hillenkamp. DNA sequence analysis by MALDI mass spectrometry. *Nucleic Acids Res*, 26(11):2554–2559, Jun 1998.

7. LITERATURVERZEICHNIS

- [95] M. Klaften and M. Hrabé de Angelis. ARTS: a web-based tool for the set-up of high-throughput genome-wide mapping panels for the SNP genotyping of mouse mutants. *Nucleic Acids Res*, 33(Web Server issue):W496–W500, Jul 2005.
- [96] G. Krohne, I. Waizenegger, and T. H. Höger. The conserved carboxy-terminal cysteine of nuclear lamins is essential for lamin association with the nuclear envelope. *J Cell Biol*, 109(5):2003–2011, Nov 1989.
- [97] L. Kruglyak. The use of a genetic map of biallelic markers in linkage studies. *Nat Genet*, 17(1):21–24, Sep 1997.
- [98] J. Lammerding, L. G. Fong, J. Y. Ji, K. Reue, C. L. Stewart, S. G. Young, and R. T. Lee. Lamins A and C but not lamin B1 regulate nuclear mechanics. *J Biol Chem*, 281(35):25768–25780, Sep 2006.
- [99] J. Lammerding, P. C. Schulze, T. Takahashi, S. Kozlov, T. Sullivan, R. D. Kamm, C. L. Stewart, and R. T. Lee. Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. *J Clin Invest*, 113(3):370–378, Feb 2004.
- [100] E. S. Lander and D. Botstein. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1):185–199, Jan 1989.
- [101] E. S. Lander, L. M. Linton, B. Birren, et al. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, Feb 2001.
- [102] P. W. Lane, A. G. Searle, C. V. Beechey, and E. Eicher. Chromosome 18 of the house mouse. *J Hered*, 72(6):409–412, 1981.
- [103] J. Leushner. MALDI TOF mass spectrometry: an emerging platform for genomics and diagnostics. *Expert Rev Mol Diagn*, 1(1):11–18, May 2001.
- [104] J. Leushner and N. H. Chiu. Automated mass spectrometry: a revolutionary technology for clinical diagnostics. *Mol Diagn*, 5(4):341–348, Dec 2000.
- [105] G. Liao, J. Wang, J. Guo, J. Allard, J. Cheng, A. Ng, S. Shafer, A. Puech, J. D. McPherson, D. Foerzler, G. Peltz, and J. Usuka. In silico genetics: identification of a functional element regulating H2-Ealpha gene expression. *Science*, 306(5696):690–695, Oct 2004.
- [106] M. Lin, L.-J. Wei, W. R. Sellers, M. Lieberfarb, W. H. Wong, and C. Li. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics*, 20(8):1233–1240, May 2004.
- [107] K. Lindblad-Toh, E. Winchester, M. J. Daly, D. G. Wang, J. N. Hirschhorn, J. P. Laviolette, K. Ardlie, D. E. Reich, E. Robinson, P. Sklar, N. Shah, D. Thomas, J. B. Fan, T. Gingeras, J. Warrington, N. Patil, T. J. Hudson, and E. S. Lander. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat Genet*, 24(4):381–386, Apr 2000.
- [108] D. P. Little, A. Braun, M. J. O'Donnell, and H. Köster. Mass spectrometry from miniaturized arrays for full comparative DNA analysis. *Nat Med*, 3(12):1413–1416, Dec 1997.

- [109] B. Liu, J. Wang, K. M. Chan, W. M. Tjia, W. Deng, X. Guan, J. dong Huang, K. M. Li, P. Y. Chau, D. J. Chen, D. Pei, A. M. Pendas, J. Cadiñanos, C. López-Otín, H. F. Tse, C. Hutchison, J. Chen, Y. Cao, K. S. E. Cheah, K. Tryggvason, and Z. Zhou. Genomic instability in laminopathy-based premature aging. *Nat Med*, 11(7):780–785, Jul 2005.
- [110] D. J. Lloyd, R. C. Trembath, and S. Shackleton. A novel interaction between lamin A and SREBP1: implications for partial lipodystrophy and other laminopathies. *Hum Mol Genet*, 11(7):769–777, Apr 2002.
- [111] E. Logette, E. Solary, and L. Corcos. Identification of a functional DNA binding site for the SREBP-1c transcription factor in the first intron of the human caspase-2 gene. *Biochim Biophys Acta*, 1738(1-3):1–5, Dec 2005.
- [112] H. Maier, S. Döhr, K. Grote, S. O’Keeffe, T. Werner, M. Hrabé de Angelis, and R. Schneider. Lit-Miner and WikiGene: identifying problem-related key players of gene regulation using publication abstracts. *Nucleic Acids Res*, 33(Web Server issue):W779–W782, Jul 2005.
- [113] M. P. Mallampalli, G. Huyer, P. Bendale, M. H. Gelb, and S. Michaelis. Inhibiting farnesylation reverses the nuclear morphology defect in a HeLa cell model for Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A*, 102(40):14416–14421, Oct 2005.
- [114] K. F. Manly, R. H. Cudmore, and J. M. Meer. Map Manager QTX, cross-platform software for genetic mapping. *Mamm Genome*, 12(12):930–932, Dec 2001.
- [115] E. Markiewicz, T. Dechat, R. Foisner, R. A. Quinlan, and C. J. Hutchison. Lamin A/C binding protein LAP2alpha is required for nuclear anchorage of retinoblastoma protein. *Mol Biol Cell*, 13(12):4401–4413, Dec 2002.
- [116] E. Marshall. Genome sequencing. Celera assembles mouse genome; public labs plan new strategy. *Science*, 292(5518):822, May 2001.
- [117] T. Maruyama, Y. Miyake, S. Tajima, T. Funahashi, Y. Matsuzawa, and A. Yamamoto. A single point mutation in the splice donor site of the low-density-lipoprotein-receptor gene produces intron read-through, exon-skipped and cryptic-site-utilized transcripts. *Eur J Biochem*, 232(3):700–705, Sep 1995.
- [118] K. N. Maxwell and J. L. Breslow. Adenoviral-mediated expression of Pcsk9 in mice results in a low-density lipoprotein receptor knockout phenotype. *Proc Natl Acad Sci U S A*, 101(18):7100–7105, May 2004.
- [119] F. D. McKeon, M. W. Kirschner, and D. Caput. Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature*, 319(6053):463–468, 1986.
- [120] V. McKusick. *Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders*. Johns Hopkins University Press, 12th edition, 1998.
- [121] J. M. K. Mislow, J. M. Holaska, M. S. Kim, K. K. Lee, M. Segura-Totten, K. L. Wilson, and E. M. McNally. Nesprin-1alpha self-associates and binds directly to emerin and lamin A in vitro. *FEBS Lett*, 525(1-3):135–140, Aug 2002.

7. LITERATURVERZEICHNIS

- [122] J. L. Moran, A. D. Bolton, P. V. Tran, et al. Utilization of a whole genome SNP panel for efficient genetic mapping in the mouse. *Genome Res*, 16(3):436–440, Mar 2006.
- [123] K. Motoyama, S. Fukumoto, H. Koyama, M. Emoto, H. Shimano, K. Maemura, and Y. Nishizawa. SREBP inhibits VEGF expression in human smooth muscle cells. *Biochem Biophys Res Commun*, 342(1):354–360, Mar 2006.
- [124] C. L. Moulson, G. Go, J. M. Gardner, A. C. van der Wal, J. H. S. Smitt, J. M. van Hagen, and J. H. Miner. Homozygous and compound heterozygous mutations in ZMPSTE24 cause the laminopathy restrictive dermopathy. *J Invest Dermatol*, 125(5):913–919, Nov 2005.
- [125] A. Muchir, G. Bonne, A. J. van der Kooi, M. van Meegen, F. Baas, P. A. Bolhuis, M. de Visser, and K. Schwartz. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). *Hum Mol Genet*, 9(9):1453–1459, May 2000.
- [126] M. C. Mullins, M. Hammerschmidt, P. Haffter, and C. Nüsslein-Volhard. Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. *Curr Biol*, 4(3):189–202, Mar 1994.
- [127] K. B. Mullis and F. A. Faloon. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*, 155:335–350, 1987.
- [128] R. J. Mural, M. D. Adams, E. W. Myers, et al. A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome. *Science*, 296(5573):1661–1671, May 2002.
- [129] Z. M. Nagy and J. R. Misanin. Visual perception in the retinal degenerate C3H mouse. *J Comp Physiol Psychol*, 72(2):306–310, Aug 1970.
- [130] I. Nassi and B. Shneiderman. Flowchart techniques for structured programming. *ACM SIGPLAN Notices*, 8:12–26, 1973.
- [131] C. L. Navarro, J. Cadiñanos, A. D. Sandre-Giovannoli, R. Bernard, S. Courrier, I. Boccaccio, A. Boyer, W. J. Kleijer, A. Wagner, F. Giuliano, F. A. Beemer, J. M. Freije, P. Cau, R. C. M. Hennekam, C. López-Otín, C. Badens, and N. Lévy. Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. *Hum Mol Genet*, 14(11):1503–1513, Jun 2005.
- [132] C. L. Navarro, A. D. Sandre-Giovannoli, R. Bernard, I. Boccaccio, A. Boyer, D. Geneviève, S. Hadj-Rabia, C. Gaudy-Marqueste, H. S. Smitt, P. Vabres, L. Faivre, A. Verloes, T. V. Essen, E. Flori, R. Hennekam, F. A. Beemer, N. Laurent, M. L. Merrer, P. Cau, and N. Lévy. Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermopathy as a lethal neonatal laminopathy. *Hum Mol Genet*, 13(20):2493–2503, Oct 2004.
- [133] C. R. Newton, A. Graham, L. E. Heptinstall, S. J. Powell, C. Summers, N. Kalsheker, J. C. Smith, and A. F. Markham. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res*, 17(7):2503–2516, Apr 1989.
- [134] D. A. Nickerson, V. O. Tobe, and S. L. Taylor. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res*, 25(14):2745–2751, Jul 1997.

- [135] P. M. Nolan, J. Peters, M. Strivens, et al. A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet*, 25(4):440–443, Aug 2000.
- [136] G. Novelli, A. Muchir, F. Sangiuolo, A. Helbling-Leclerc, M. R. D’Apice, C. Massart, F. Capon, P. Sbraccia, M. Federici, R. Lauro, C. Tudisco, R. Pallotta, G. Scarano, B. Dallapiccola, L. Merlini, and G. Bonne. Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C. *Am J Hum Genet*, 71(2):426–431, Aug 2002.
- [137] J. K. Noveroske, J. S. Weber, and M. J. Justice. The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mamm Genome*, 11(7):478–483, Jul 2000.
- [138] C. Nüsslein-Volhard and E. Wieschaus. Mutations affecting segment number and polarity in *Drosophila*. *Nature*, 287(5785):795–801, Oct 1980.
- [139] M. Olson, L. Hood, C. Cantor, and D. Botstein. A common language for physical mapping of the human genome. *Science*, 245(4925):1434–1435, Sep 1989.
- [140] P. C. Orban, D. Chui, and J. D. Marth. Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci U S A*, 89(15):6861–6865, Aug 1992.
- [141] T. Ozaki, M. Saijo, K. Murakami, H. Enomoto, Y. Taya, and S. Sakiyama. Complex formation between lamin A and the retinoblastoma gene product: identification of the domain on lamin A required for its interaction. *Oncogene*, 9(9):2649–2653, Sep 1994.
- [142] Q. S. Padiath, K. Saigoh, R. Schiffmann, H. Asahara, T. Yamada, A. Koeppen, K. Hogan, L. J. Ptáček, and Y.-H. Fu. Lamin B1 duplications cause autosomal dominant leukodystrophy. *Nat Genet*, 38(10):1114–1123, Oct 2006.
- [143] W. Pargent, S. Heffner, K. F. Schäble, D. Soewarto, H. Fuchs, and M. Hrabé de Angelis. MouseNet database: digital management of a large-scale mutagenesis project. *Mamm Genome*, 11(7):590–593, Jul 2000.
- [144] W. R. Pearson and D. J. Lipman. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A*, 85(8):2444–2448, Apr 1988.
- [145] J. L. Peirce, L. Lu, J. Gu, L. M. Silver, and R. W. Williams. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet*, 5:7, Apr 2004.
- [146] A. M. Pendás, Z. Zhou, J. Cadiñanos, J. M. P. Freije, J. Wang, K. Hultenby, A. Astudillo, A. Wernerson, F. Rodríguez, K. Tryggvason, and C. López-Otín. Defective prelamin A processing and muscular and adipocyte alterations in *Zmpste24* metalloproteinase-deficient mice. *Nat Genet*, 31(1):94–99, May 2002.
- [147] P. M. Petkov, Y. Ding, M. A. Cassell, W. Zhang, G. Wagner, E. E. Sargent, S. Asquith, V. Crew, K. A. Johnson, P. Robinson, V. E. Scott, and M. V. Wiles. An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome Res*, 14(9):1806–1811, Sep 2004.
- [148] W. Pusch, J.-H. Wurbach, H. Thiele, and M. Kostrzewa. MALDI-TOF mass spectrometry-based SNP genotyping. *Pharmacogenomics*, 3(4):537–548, Jul 2002.
- [149] J. Ragoussis, G. P. Elvidge, K. Kaur, and S. Colella. Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. *PLoS Genet*, 2(7):e100, Jul 2006.

7. LITERATURVERZEICHNIS

- [150] B. Rathkolb, T. Decker, E. Fuchs, D. Soewarto, C. Fella, S. Heffner, W. Pargent, R. Wanke, R. Balling, M. Hrabé de Angelis, H. J. Kolb, and E. Wolf. The clinical-chemical screen in the Munich ENU Mouse Mutagenesis Project: screening for clinically relevant phenotypes. *Mamm Genome*, 11(7):543–546, Jul 2000.
- [151] D. E. Reich, M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti, D. J. Richter, T. Lavery, R. Kouyoumjian, S. F. Farhadian, R. Ward, and E. S. Lander. Linkage disequilibrium in the human genome. *Nature*, 411(6834):199–204, May 2001.
- [152] H. Ris. High-resolution field-emission scanning electron microscopy of nuclear pore complex. *Scanning*, 19(5):368–375, Aug 1997.
- [153] P. Ross, L. Hall, I. Smirnov, and L. Haff. High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nat Biotechnol*, 16(13):1347–1351, Dec 1998.
- [154] S. Rozen and H. Skaletsky. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*, 132:365–386, 2000.
- [155] I. Rubio-Aliaga, D. Soewarto, S. Wagner, M. Klaften, H. Fuchs, S. Kalaydjiev, D. H. Busch, M. Klempt, B. Rathkolb, E. Wolf, K. Abe, S. Zeiser, G. K. H. Przemeck, J. Beckers, and M. H. de Angelis. A genetic screen for modifiers of the delta1-dependent notch signaling function in the mouse. *Genetics*, 175(3):1451–1463, Mar 2007.
- [156] R. Rugh. *The Mouse: Its Reproduction and Development*. Oxford University Press, 1990.
- [157] F. Runkel, M. Klaften, K. Koch, V. Böhnert, H. Büsow, H. Fuchs, T. Franz, and M. Hrabé de Angelis. Morphologic and molecular characterization of two novel Krt71 (Krt2-6g) mutations: Krt71 (rco12) and Krt71 (rco13). *Mamm Genome*, 17(12):1172–1182, Dec 2006.
- [158] L. B. Russell and C. S. Montgomery. Supermutagenicity of ethylnitrosourea in the mouse spot test: comparisons with methylnitrosourea and ethylnitrosourethane. *Mutat Res*, 92(1-2):193–204, Feb 1982.
- [159] L. B. Russell and W. L. Russell. Frequency and nature of specific-locus mutations induced in female mice by radiations and chemicals: a review. *Mutat Res*, 296(1-2):107–127, Dec 1992.
- [160] W. L. Russell, P. R. Hunsicker, G. D. Raymer, M. H. Steele, K. F. Stelzner, and H. M. Thompson. Dose–response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. *Proc Natl Acad Sci U S A*, 79(11):3589–3591, Jun 1982.
- [161] W. L. Russell, E. M. Kelly, P. R. Hunsicker, J. W. Bangham, S. C. Maddux, and E. L. Phipps. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc Natl Acad Sci U S A*, 76(11):5818–5819, Nov 1979.
- [162] R. A. Röber, K. Weber, and M. Osborn. Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. *Development*, 105(2):365–378, Feb 1989.
- [163] J. Sambrook, T. Maniatis, and D. Russel. *Molecular Cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, 3rd edition, 2001.

- [164] A. D. Sandre-Giovannoli, M. Chaouch, S. Kozlov, J.-M. Vallat, M. Tazir, N. Kassouri, P. Szepe-towski, T. Hammadouche, A. Vandenberghe, C. L. Stewart, D. Grid, and N. Lévy. Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axo-nal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse. *Am J Hum Genet*, 70(3):726–736, Mar 2002.
- [165] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74(12):5463–5467, Dec 1977.
- [166] A. M. Sasseville and Y. Langelier. In vitro interaction of the carboxy-terminal domain of lamin A with actin. *FEBS Lett*, 425(3):485–489, Apr 1998.
- [167] K. Sax. The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHA-SEOLUS VULGARIS. *Genetics*, 8(6):552–560, 1923.
- [168] P. Scaffidi and T. Misteli. Lamin A-dependent nuclear defects in human aging. *Science*, 312(5776):1059–1063, May 2006.
- [169] W. K. Schmidt, A. Tam, K. Fujimura-Kamada, and S. Michaelis. Endoplasmic reticulum membra-ne localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal a-factor cleavage. *Proc Natl Acad Sci U S A*, 95(19):11175–11180, Sep 1998.
- [170] F. Schnütgen, N. Doerflinger, C. Calléja, O. Wendling, P. Chambon, and N. B. Ghyselinck. A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. *Nat Biotechnol*, 21(5):562–565, May 2003.
- [171] H. Seedorf, I. N. Springer, E. Grundner-Culemann, H.-K. Albers, A. Reis, H. Fuchs, M. Hrabé de Angelis, and Y. Açil. Amelogenesis imperfecta in a new animal model—a mutation in chromosome 5 (human 4q21). *J Dent Res*, 83(8):608–612, Aug 2004.
- [172] S. Shackleton, D. J. Lloyd, S. N. Jackson, R. Evans, M. F. Niermeijer, B. M. Singh, H. Schmidt, G. Brabant, S. Kumar, P. N. Durrington, S. Gregory, S. O’Rahilly, and R. C. Trembath. LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. *Nat Genet*, 24(2):153–156, Feb 2000.
- [173] S. T. Sherry, M. H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, and K. Sirotkin. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*, 29(1):308–311, Jan 2001.
- [174] T. Shibuya and K. Morimoto. A review of the genotoxicity of 1-ethyl-1-nitrosourea. *Mutat Res*, 297(1):3–38, Jul 1993.
- [175] L. M. Silver. *Mouse Genetics: Concepts and Applications*. Oxford University Press, New York and Oxford, 1995.
- [176] E. M. Simpson, C. C. Linder, E. E. Sargent, M. T. Davisson, L. E. Mobraaten, and J. J. Sharp. Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat Genet*, 16(1):19–27, May 1997.
- [177] B. Singer and M. K. Dosanjh. Site-directed mutagenesis for quantitation of base-base interactions at defined sites. *Mutat Res*, 233(1-2):45–51, 1990.
- [178] G. D. Snell. *Origins of Inbred Mice*, pages 1–33. Academic Press, New York, 1978.

7. LITERATURVERZEICHNIS

- [179] B. Sobrino, M. Brión, and A. Carracedo. SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci Int*, 154(2-3):181–194, Nov 2005.
- [180] D. Soewarto, V. Blanquet, and M. Hrabé de Angelis. Random ENU mutagenesis. *Methods Mol Biol*, 209:249–266, 2003.
- [181] M. Soller and T. Brody. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor Appl Genet*, 47(1):35–39, 1976.
- [182] T. P. Spann, A. E. Goldman, C. Wang, S. Huang, and R. D. Goldman. Alteration of nuclear lamin organization inhibits RNA polymerase II-dependent transcription. *J Cell Biol*, 156(4):603–608, Feb 2002.
- [183] W. L. Stanford, J. B. Cohn, and S. P. Cordes. Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet*, 2(10):756–768, Oct 2001.
- [184] V. Stierlé, J. Couprie, C. Ostlund, I. Krimm, S. Zinn-Justin, P. Hossenlopp, H. J. Worman, J.-C. Courvalin, and I. Duband-Goulet. The carboxyl-terminal region common to lamins A and C contains a DNA binding domain. *Biochemistry*, 42(17):4819–4828, May 2003.
- [185] N. Storm, B. Darnhofer-Patel, D. van den Boom, and C. P. Rodi. MALDI-TOF mass spectrometry-based SNP genotyping. *Methods Mol Biol*, 212:241–262, 2003.
- [186] N. Stuurman, S. Heins, and U. Aebi. Nuclear lamins: their structure, assembly, and interactions. *J Struct Biol*, 122(1-2):42–66, 1998.
- [187] A. I. Su, M. P. Cooke, K. A. Ching, Y. Hakak, J. R. Walker, T. Wiltshire, A. P. Orth, R. G. Vega, L. M. Sapinoso, A. Moqrich, A. Patapoutian, G. M. Hampton, P. G. Schultz, and J. B. Hogenesch. Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A*, 99(7):4465–4470, Apr 2002.
- [188] T. Sullivan, D. Escalante-Alcalde, H. Bhatt, M. Anver, N. Bhat, K. Nagashima, C. L. Stewart, and B. Burke. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J Cell Biol*, 147(5):913–920, Nov 1999.
- [189] A. Tam, F. J. Nouvet, K. Fujimura-Kamada, H. Slunt, S. S. Sisodia, and S. Michaelis. Dual roles for Ste24p in yeast a-factor maturation: NH₂-terminal proteolysis and COOH-terminal CAAX processing. *J Cell Biol*, 142(3):635–649, Aug 1998.
- [190] K. Tang, D. Fu, S. Kötter, R. J. Cotter, C. R. Cantor, and H. Köster. Matrix-assisted laser desorption/ionization mass spectrometry of immobilized duplex DNA probes. *Nucleic Acids Res*, 23(16):3126–3131, Aug 1995.
- [191] H. Taniura, C. Glass, and L. Gerace. A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. *J Cell Biol*, 131(1):33–44, Oct 1995.
- [192] K. R. Thomas and M. R. Capecchi. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, 51(3):503–512, Nov 1987.
- [193] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25(24):4876–4882, Dec 1997.

- [194] D. W. Threadgill, D. Yee, A. Matin, J. H. Nadeau, and T. Magnuson. Genealogy of the 129 inbred strains: 129/SvJ is a contaminated inbred strain. *Mamm Genome*, 8(6):390–393, Jun 1997.
- [195] A. R. Tobler, S. Short, M. R. Andersen, et al. The SNPlex genotyping system: a flexible and scalable platform for SNP genotyping. *J Biomol Tech*, 16(4):398–406, Dec 2005.
- [196] J. I. Toth, S. H. Yang, X. Qiao, A. P. Beigneux, M. H. Gelb, C. L. Moulson, J. H. Miner, S. G. Young, and L. G. Fong. Blocking protein farnesyltransferase improves nuclear shape in fibroblasts from humans with progeroid syndromes. *Proc Natl Acad Sci U S A*, 102(36):12873–12878, Sep 2005.
- [197] M.-Y. Tsai, S. Wang, J. M. Heidinger, D. K. Shumaker, S. A. Adam, R. D. Goldman, and Y. Zheng. A mitotic lamin B matrix induced by RanGTP required for spindle assembly. *Science*, 311(5769):1887–1893, Mar 2006.
- [198] W. Valdar, L. C. Solberg, D. Gauguier, S. Burnett, P. Klenerman, W. O. Cookson, M. S. Taylor, J. N. P. Rawlins, R. Mott, and J. Flint. Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet*, 38(8):879–887, Aug 2006.
- [199] I. Varela, J. Cadiñanos, A. M. Pendás, A. Gutiérrez-Fernández, A. R. Folgueras, L. M. Sánchez, Z. Zhou, F. J. Rodríguez, C. L. Stewart, J. A. Vega, K. Tryggvason, J. M. P. Freije, and C. López-Otín. Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. *Nature*, 437(7058):564–568, Sep 2005.
- [200] J. C. Venter, M. D. Adams, E. W. Myers, et al. The sequence of the human genome. *Science*, 291(5507):1304–1351, Feb 2001.
- [201] J. L. Vivian, Y. Chen, D. Yee, E. Schneider, and T. Magnuson. An allelic series of mutations in Smad2 and Smad4 identified in a genotype-based screen of N-ethyl-N-nitrosourea-mutagenized mouse embryonic stem cells. *Proc Natl Acad Sci U S A*, 99(24):15542–15547, Nov 2002.
- [202] S. Vreugde, A. Erven, C. J. Kros, W. Marcotti, H. Fuchs, K. Kurima, E. R. Wilcox, T. B. Friedman, A. J. Griffith, R. Balling, M. Hrabé de Angelis, K. B. Avraham, and K. P. Steel. Beethoven, a mouse model for dominant, progressive hearing loss DFNA36. *Nat Genet*, 30(3):257–258, Mar 2002.
- [203] C. M. Wade, E. J. Kulbokas, A. W. Kirby, M. C. Zody, J. C. Mullikin, E. S. Lander, K. Lindblad-Toh, and M. J. Daly. The mosaic structure of variation in the laboratory mouse genome. *Nature*, 420(6915):574–578, Dec 2002.
- [204] R. H. Waterston, K. Lindblad-Toh, E. Birney, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915):520–562, Dec 2002.
- [205] J. L. Weber and P. E. May. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet*, 44(3):388–396, Mar 1989.
- [206] W. Weinberg. Über den Nachweis der Vererbung beim Menschen. *Jahresh Verein f vaterl Naturk in Württ*, 64:368–382, 1908.
- [207] T. Wiltshire, M. T. Pletcher, S. Batalov, S. W. Barnes, L. M. Tarantino, M. P. Cooke, H. Wu, K. Smylie, A. Santrosyan, N. G. Copeland, N. A. Jenkins, F. Kalush, R. J. Mural, R. J. Glynne, S. A. Kay, M. D. Adams, and C. F. Fletcher. Genome-wide single-nucleotide polymorphism

7. LITERATURVERZEICHNIS

- analysis defines haplotype patterns in mouse. *Proc Natl Acad Sci U S A*, 100(6):3380–3385, Mar 2003.
- [208] W. Wurst, J. Rossant, V. Prideaux, M. Kownacka, A. Joyner, D. P. Hill, F. Guillemot, S. Gasca, D. Cado, and A. Auerbach. A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. *Genetics*, 139(2):889–899, Feb 1995.
- [209] S. H. Yang, M. O. Bergo, J. I. Toth, X. Qiao, Y. Hu, S. Sandoval, M. Meta, P. Bendale, M. H. Gelb, S. G. Young, and L. G. Fong. Blocking protein farnesyltransferase improves nuclear blebbing in mouse fibroblasts with a targeted Hutchinson-Gilford progeria syndrome mutation. *Proc Natl Acad Sci U S A*, 102(29):10291–10296, Jul 2005.
- [210] S. Ye, S. Dhillon, X. Ke, A. R. Collins, and I. N. Day. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res*, 29(17):E88–E88, Sep 2001.
- [211] S. G. Young, L. G. Fong, and S. Michaelis. Prelamin A, Zmpste24, misshapen cell nuclei, and progeria—new evidence suggesting that protein farnesylation could be important for disease pathogenesis. *J Lipid Res*, 46(12):2531–2558, Dec 2005.
- [212] M. S. Zastrow, S. Vlcek, and K. L. Wilson. Proteins that bind A-type lamins: integrating isolated clues. *J Cell Sci*, 117(Pt 7):979–987, Mar 2004.
- [213] F. L. Zhang and P. J. Casey. Protein prenylation: molecular mechanisms and functional consequences. *Annu Rev Biochem*, 65:241–269, 1996.

A. Liste der verwendeten Primer

Tabelle A.1: Primer zur Amplifikation vom Zmpste24

Name	Sequence forward (5'-3')	Sequence reverse (5'-3')
Exon1_1	CTGACGTTCCCAGGAAGAAG	AAGTCCCACATAGCGTCCAC
Exon1_2	CTGAAGAAGCCCTCGGAAG	GCTGGCCTAAGGCTGTCC
Exon2	AATGCTACAGACCATTGGGTG	ACTGGACTGCTAAAGGCTGG
Exon3	CCCATATACATTGTTTAGGCAAA	AATGGTAGGAGCCCCAAAAGG
Exon4	TGTGTACTTGCTGGTTTGTGC	TGGGACTGAAGCACAGAAAAC
Exon5	GCTGCCTACCTTATGTTCTTAAAGTG	TGGGAAGAATTGGAGCAAAG
Exon6	GTCCTGGGTGTTTGTGGAT	AAAGACTGACTGACTTAAGAACAGAGT
Exon7	TCCCAGTGTCCAAAGGACC	CACACACTGCATGGAACTTG
Exon8	TGATTAGATGGGTCAAGGCTAAA	TCAACTACATGAGACGCTGTTT
Exon9	TGGCAACATCCATAGTGAGC	CCCCAAAGAGTCTAGATTTGAAAC
Exon10_1	CCTGACCTGTACGGCAGAAT	GCTGCCAGGACAGAATAACC
Exon10_2	CCGGTTTCTGACTGGTTGTT	TGCAGATTCCACATAGTGCAG
Exon10_3	TTCCGGCTCTTGATGTTTTT	TGATTGGCTGGGCTAAATTAT
Exon10_4	TGGAATCTGCATAGAGGGTTTT	GGCCAGCAAGGACTAGAGTG
Exon10_5	TTAGCCCAGCCAATCAGTGT	AAAGCAAACCCAAAGTCAATG
Exon10_6	AGCACAAGGTGGCTAGAAC	CGATGGAAGGGAGGGTAAAT
Exon10_7	CAACTGAGGCAGAGGCAGAT	AAAGTAGAGACTTCTGTCTTGAAAAA
Exon10_8	ATTTACCCTCCCTTCCATCG	CACATACCTTTTCCCAGTGA
Exon10_9	TCAAGGACAGAAGTCTCTACTTTTATT	TGCTCTGCGTAGCTGTAGGA
Exon10_10	GGTAGCGAGTGCCTCCTG	AGGTCCATAGACGGCACTGT

Tabelle A.2: Primer zur Amplifikation vom Gli3

Name	Sequence forward (5'-3')	Sequence reverse (5'-3')
Exon7	CAACTTCTGTCCCTCTCCG	TGACCACTGGCACTACAGATG
Exon8	AAGGACCTGGTGTGATTTGG	GCCTTAACAATGATGGACCC
Exon9	TGCATCATCTGGACTCTCCC	GAAGTATCCCAAATCACCG
Exon10	TCCTAGTCTCTGGAGCAGGG	TCTCAGTGTGCCAGCTAAG
Exon11	TGAGCATTTCACTTCCTTTTG	TTTCAGATCCTGGGTCGAAG
Exon12	TTTTGAAAAGCTGCTGACCC	TCAGAGCCCACTCACAGTTG
Exon13	CCTGTGAGGTACCTAAATGCTG	GGATGGTGTATCCCTTGAC

A. LISTE DER VERWENDETEN PRIMER

Tabelle A.3: Primer zur Genotypisierung von SMA004

Name	Sequence forward (5'-3')	Sequence reverse (5'-3')
outer	GATAATGTCCTGGGTGTTTGTGGATG	CTCCTCTGGAAGAACACCCATGAGTCTT
inner	CCCTTTGACTAAGGTGTATGTTGTTGACGA	CTCTCTCACCTCCAGGTAGCCTGAC

Tabelle A.4: Primer zur Amplifikation möglicher Splicevarianten von SMA004

Name	Sequence (5'-3')
Primer_a	TGCCTGGCTGTTCACATTAG
Primer_b	GAGGGAAAGCTTAAGCAGGAA
Primer_c	AAACTGGATTGGCATCAGG
Primer_d	TCTTCACTGCCCCCTCACC

Tabelle A.5: Primer für die quantitative RT-PCR

Name	Sequence forward (5'-3')	Sequence reverse (5'-3')
ApoA5	TCGCAGTGTTTCGCAAGC	GCCTTTCAGGTTCTCCTGTG
Casp2	TGACAGGAGGAGCAGGATTT	CCCTTTGGCCTGGATGAG
Pax3	AAACCCAAGCAGGTGACAAC	CTCAGGATGCGGCTGATAG
Pcsk9	CTCCCCTGATGGAAGCAG	GTCACACTTGCTCGCCTGT
Rbp1	CAAGTGCATGACCACTGTGA	CTCAGCTCTCATTCCAGGTG
Rbp3	GATCATCGACATGAGGTTCAA	TCGTTACCTGTGAGCTGAG
Vegf	TCCTGGAGCGTTCACTGTG	GCTTGTCACATCTGCAAGTACG
Hprt1	CTGATTATGGACAGGACTGAAAGA	CCGTTGACTGATCATTACAGTAGC
β -Actin	GCCACCAGTTCGCCAT	CATCACACCCTGGTGCCTA
Gapdh	AAGGGTCATGACCACAGTC	ATCATACTTGGCAGGTTTCTCC

Tabelle A.6: Sequenzen der Amplifikations- und Extensionsprimer des hME Markersatzes. Die Locusangaben beziehen sich auf Mouse Genome Assembly 35.

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
1	29.89	3706618	ACGTTGGATGGCCATACTCAACTGCTATCC	ACGTTGGATGGTGTCCACCCTTAGTTGTC	CTTCCTTTCTGCCATTGTT	A	T
1	43.98	13459182	ACGTTGGATGTGATGAAGATGCGGCCATTG	ACGTTGGATGGGCACCTCAGGAAGCTTAAT	ACCTCAGGAAGCTTAATAGTCAT	G	A
1	62.90	3688436	ACGTTGGATGCATCCATGTGACATAGGTGC	ACGTTGGATGACCCTGGGAAATGTATCCAC	GGACTTGGTTTTCTAATATGG	A	G
1	74.40	3090522	ACGTTGGATGAGAACGTGACCTCTTCTCC	ACGTTGGATGCATGGCTCAITACCAGAGAC	ATTGATGCAGCCCTCCTC	C	T
1	90.37	3684025	ACGTTGGATGGTCCCATCTTCTTATCCC	ACGTTGGATGCCCACAGAGCACTTAATGAG	AGTCTGACTTGCTTTTATCAT	A	G
1	109.17	13476046	ACGTTGGATGAAAGGCATGAGGTAGGTGAG	ACGTTGGATGCCTGCTTACTTGTAGATGTTG	GTTGACATGATGACCATCT	A	G
1	119.76	3672659	ACGTTGGATGTTGAGCTGTCCCTTCTTGG	ACGTTGGATGTCAACAACAAGGGGTGAGC	GAAGTCTGGTGACTTAGAGGCC	A	G
1	135.04	8250085	ACGTTGGATGTCAGAGCAAGAGGTTGCAG	ACGTTGGATGATGCTCCCTGAAAGCAGAAG	AGAGGAAATCTGAGGTCACCTA	A	T
1	149.63	6264289	ACGTTGGATGCCCAAGATCCTTTCAAAGGC	ACGTTGGATGTGTTCACTGCTAGGTTGCAG	AGACTCAGAACTGAGGG	A	C
1	156.77	4222732	ACGTTGGATGTGCAGGCTGGGATAGATAGG	ACGTTGGATGCCCTCAATTTCAAACGGTATG	GGATCCCTCATCCAAAG	G	A
1	176.90	3702990	ACGTTGGATGTAGTTCTGACATGGGCTTTC	ACGTTGGATGGTTTCAGCCGAGAAAGGATG	AAGGAACAATGAAATCCATCTACA	A	G
1	188.89	13476297	ACGTTGGATGTATAAACACGGTAAGGACCC	ACGTTGGATGTATGAGTCAACAGTGTAGGG	CCTGCTGGCTATGAGCCCA	T	C
2	13.38	13476352	ACGTTGGATGGGACCGTTTCAGAGATATGTG	ACGTTGGATGCCCTGATATCATAGGTACAC	TCTACCTTTTTAACATAGCTGGT	A	G
2	37.76	3022883	ACGTTGGATGCCTTTGCAGAGAGGAAAAGA	ACGTTGGATGTGACCCAGCACAACTGAGAG	AACTGAGAGCCTGGAAAA	A	C
2	71.12	4223216	ACGTTGGATGTCCTCTCTGATGACTTATG	ACGTTGGATGCCTAGTCGCGAATGGAAAAC	CTGAAACTGGATACATTCAAACA	G	A
2	102.78	13476672	ACGTTGGATGCTCCCCATGATATCTGGAC	ACGTTGGATGATAAATCTAGGCTCTCAGGG	GGCTCTCAGGGAATAGA	G	A
2	118.91	8279353	ACGTTGGATGCCAGGACTGAGGGTTAAAT	ACGTTGGATGTAAAGAGGCAGAAAGGGAGC	AGGGACGTGCTGTAATGCAGA	A	G
2	131.73	4223511	ACGTTGGATGTTCCACGTTACAAAGCTGGC	ACGTTGGATGAAGCACAAAACCCGTCATG	CACACAGCTAACACCAT	T	C
2	149.19	3675393	ACGTTGGATGAAGAGCTTTGGGTTGTTGG	ACGTTGGATGGGTGTACATATCTGTCATAGG	TCTTAGATTCAATGAAACACAC	C	T
3	10.77	3689058	ACGTTGGATGTCCACCTTACTTGATTCC	ACGTTGGATGCATTTGGCATAAGGTAGAAGC	TTTAAATGACACAAAGCATGA	A	G
3	30.06	3683804	ACGTTGGATGGTAGCTTCAAATGGGCACAC	ACGTTGGATGCAGATGAACAACACTGTGGAGG	GTTGAGCCCTCTCCAC	A	G
3	43.81	4223943	ACGTTGGATGCCTGTACTGGGACATACAC	ACGTTGGATGAGATGGCCTAATTGACCAGC	GACCAGCATTGGGTAGAGAC	T	C
3	54.69	3719338	ACGTTGGATGCAGCAAAGCCCTTCTCATC	ACGTTGGATGATCACTGTCCACAACCTTG	TCCCACAACCTTGAAATCCCTA	C	T
3	69.85	13477178	ACGTTGGATGAACGTGCATGACGGCAAATC	ACGTTGGATGAAAGAGAAGCCACAGTGGAC	ACAGTGAATGTGATGCT	T	G
3	79.44	4136771	ACGTTGGATGGGGCCAGTTTATATCAGAC	ACGTTGGATGCTCTGTAGACTGTTCTCTG	TGTAGCCAATCGGGACC	C	T
3	103.27	13477302	ACGTTGGATGCAGTTGATATCATCTGCTGGG	ACGTTGGATGAGAGGGGAAATGTCCCAATG	TCCCAATGTCTACACAGAAAG	C	T
3	112.56	6207837	ACGTTGGATGTGTTTTCCAGGATCCAGTGCC	ACGTTGGATGAGCAAAGATATGCAAGGCAG	ATGCAAGGCAGAITCAA	C	T
3	131.81	4136518	ACGTTGGATGAGTGGCCCTCTTGGAAATTG	ACGTTGGATGTCCATCTTCTGTCAGAG	AGCGAGGATGAAAAGAACAT	A	C
3	149.33	3718378	ACGTTGGATGACAGGAGGACCAACAGCAC	ACGTTGGATGTCTGCTCTCTTTTGGCAGAC	AACAATAGTTTCCCAACAATA	C	T
3	152.29	3686068	ACGTTGGATGCCAAACAACCAAGAAAGGTTG	ACGTTGGATGGCACCAGTGTACATAAAAATA	TTTCTATTACTTTGGTGTGA	A	G
4	35.45	3703981	ACGTTGGATGCTTGGCAAATACCCACGTGC	ACGTTGGATGTCAACATCTCCACGCTGCC	TCCCTCTCCACAACCTAG	A	C
4	62.20	3686083	ACGTTGGATGTGATCTTGAGTGCTGAGTCC	ACGTTGGATGACAGCAAACCTCAGTCTCTG	CGCACCCAGCTTCTCCC	C	T
4	92.47	3703756	ACGTTGGATGGAGCTTCAAGATGGTACATAG	ACGTTGGATGCTTCTCTTACATACTACC	TTCATTTGCAAGACATTTCTAGG	G	T
4	107.94	3664637	ACGTTGGATGCCTTTCACTCCCATACCTTC	ACGTTGGATGAAGCAGAGCACTGTTCTGG	TTGACCTAAATGTAGATAACA	A	C
4	123.39	3682306	ACGTTGGATGCACAGTCTTCTATCGCCTC	ACGTTGGATGTGATGCTAAGTTGCACAGGC	CCAACCTACCAAGAGGG	A	G
4	144.75	13478035	ACGTTGGATGCCTCCAAGGGATATTCTAAG	ACGTTGGATGACATCAGGCACGTGTTATTC	CAAGGAAAGACCCTGCCAC	C	G
5	21.84	3023036	ACGTTGGATGCTCTCTTAAGCTGATGTTGG	ACGTTGGATGAAAGTCTACAGAGGTTGGG	TTAGCATAATAAAGGCAGATTACA	A	G
5	31.09	3718492	ACGTTGGATGGAGCTAGCCATGTAACCTGTG	ACGTTGGATGTGATAGTTGAGGAAGACACAG	TCCCGCCCTTGAAAAGC	A	G
5	61.08	3724321	ACGTTGGATGTAAGAGCAGTTCAAGCGACC	ACGTTGGATGAGGAGGCTTCTGTACAGATG	CTAAGTTTCCACAACAATGCCCG	A	C
5	76.85	4225300	ACGTTGGATGAAATCTTACACAACAGGAG	ACGTTGGATGGCAGGCTGGTGACATTTTA	GCAGATTTAGTGGATTGAG	A	G
5	90.49	3700063	ACGTTGGATGGACTCAGGCTCGATGAGATG	ACGTTGGATGGCATCAGGATCTGAAATGGG	GAGAGCAGAGGAGAACCT	G	A
5	91.96	3717237	ACGTTGGATGGCTTCGTGACATCGTTGATC	ACGTTGGATGATAACCGCTGCTCAAAATG	CGCTGCTCAAAATGTGGCTC	A	G
5	103.63	8256225	ACGTTGGATGTAAGCACTCCGGTTGATC	ACGTTGGATGGCAACAATAAATGCTGTTCC	TGCTGTTCTAGCAACCTTT	T	C
6	13.31	3678339	ACGTTGGATGCGATTCCTTTCTGTGCTC	ACGTTGGATGCTCTCTCTTCTAAAGCTTC	GAGGAAAATCTACTACACT	A	C
6	29.95	3676693	ACGTTGGATGGGCAGGTAGGGAATTGTGTC	ACGTTGGATGCATTTCTATGCAGACCACAC	AGACCACACTCTCCTA	A	G
6	43.89	3703696	ACGTTGGATGAGCAGCTCATGAAGTTGGAG	ACGTTGGATGAAAGCAGGAGCAGGTGTAAG	CCAGTGTGCTGTACAGA	G	C
6	55.14	3688612	ACGTTGGATGCAAATCTGGGTATACCCCTC	ACGTTGGATGCCTCAGGAGTGAAGAAGCTTG	TCAGCCTCAGTAGTACACT	G	A
6	72.79	13478817	ACGTTGGATGCTCTCCTAGCTAGATGCTAC	ACGTTGGATGGCAGACTTGGTAAAGTCGG	TTTGGGATGCTCAGAATC	A	C

Tabelle A.6: hME Markersatz Fortsetzung

Chromosom	Locus [Mb]	rsSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
6	86.15	3656641	ACGTTGGATGTTGCATCTTTGGTGGTGTGG	ACGTTGGATGAGAGGTAC2AT1GTG2CAGG	GTTGGAGCCCAATGTTAGGAAC2T	T	G
6	98.40	3702663	ACGTTGGATGAGCCCTCATTTCCAATGAAG	ACGTTGGATGGTTAAGTACTTGAACGGCTG	AATGCTGTACAGGAGTTG	A	G
6	114.61	3720170	ACGTTGGATGTGAACATGATGTGTCCAGG	ACGTTGGATGAGGCTCACATATGGTGAGAG	CACACTGAAGAGTGTTCCTTAC	G	A
6	125.14	3701869	ACGTTGGATGCTGCCTTTGAAACTTTTAGTC	ACGTTGGATGATCTTCTGGTCTCCTCAGGG	CCAGGAGTAATTTGGGT	A	G
6	139.57	3679077	ACGTTGGATGAAAGACGGCAAATGGACAGG	ACGTTGGATGGAGA2ACTTGGCACACTTTGG	TAGGAACATCCAGCCACA	T	C
7	14.43	4226424	ACGTTGGATGGGGTTTTTCTGTGCTGATAG	ACGTTGGATGCCATTACAAACAGCAACTAAC	AGCAACTAACATTTGCCTCT	T	G
7	36.60	3668793	ACGTTGGATGGGCCCTCTTAAGGATCACATC	ACGTTGGATGCCATAACCATTGACCTCAGAG	TTCTTACCAGGTTTCTAGGC	T	C
7	45.85	4226613	ACGTTGGATGCTCTGCTTGTATGTCCTTCTG	ACGTTGGATGAGAGT2CAGACAAGTCCACAC	GCAC2TCTGAAAAGGCTCCC	A	G
7	61.91	4226645	ACGTTGGATGCCAAAAGACATCTGATGGCC	ACGTTGGATGAGCAGTTGTGTCCATGGAG	CTGGCTTGCATGTTTATGGA	T	C
7	76.21	4226725	ACGTTGGATGACAGGCTAAAAACGCTCTGTG	ACGTTGGATGAGAAAAGGCTGTTAGTGCCTG	TTCCATATCCGGGCCCA	C	G
7	86.36	3656285	ACGTTGGATGACCTCTCTAGCATCTGTGG	ACGTTGGATGCTAATCAGAAATGTCAGGCAC	GTGTATGACTCTTTTGCATG	G	A
7	95.89	6387510	ACGTTGGATGTACCTGCATCCTTTTGGTCC	ACGTTGGATGTCTAGCGCATCAGAAGTACC	TCAGAAGTACCTAAGAGACA	G	A
7	111.37	3654133	ACGTTGGATGTGCTCTTCCCTTTTGTCTGTC	ACGTTGGATGAACACAACAATGGGTGGGTG	GTGGGTGTGATGGATGAA	T	C
7	138.89	13479566	ACGTTGGATGAGTCATCTTGTCTGTAGGACC	ACGTTGGATGAGGCTGTAGGATTTGGGAAC	AGGAGTTCCTGACTGACA	T	C
8	9.73	13479604	ACGTTGGATGGGTTGAAATTCTGGCTCTGG	ACGTTGGATGTCCTTATCCTCTCTCAGTCC	TGACCTTGAATGTGAATGGTT	C	G
8	30.15	3661862	ACGTTGGATGTTGCATAGCCAGGAAAGGG	ACGTTGGATGGGTCTGTTTGTGAGCAAGC	ATATTGATAGGTGCATACAGTC	G	A
8	45.88	13479741	ACGTTGGATGATCTTGTCTCCGTTTGTCTCC	ACGTTGGATGTCTAC2TGAAGGACTGTGC	TGTGCATGTGAAGAAAAGA	A	G
8	54.18	3723574	ACGTTGGATGAAGACTCTGTACAGCAAAGG	ACGTTGGATGATCTGTGGATATAGTATTG	CCTTCCCAATCTGTAGTC	C	T
8	75.39	6282879	ACGTTGGATGATATCTTTGGACAGGCCTC	ACGTTGGATGGCAGATACAACCTGGACAAC	AGA2CTGTTATCCAGCCAG	T	G
8	94.23	4227350	ACGTTGGATGAGCCAGACTCCATCTTAAGC	ACGTTGGATGAGACATCCGAGGAATCTGAG	GACTGTGTATGCAACATGTCT	T	C
8	106.55	3666037	ACGTTGGATGTAGCAAAGCGTATCAGGTCC	ACGTTGGATGACCAGGACTCCAGAACTTTG	GGACTCCAGAACTTTGCCCTCAC	G	T
8	121.91	4227428	ACGTTGGATGCAGCCACCATCTCAAGTGTCT	ACGTTGGATGGTAAACAACCCAGAGTGTCC	CTTAGATCAAA2TGAAGCCA	G	A
9	32.61	3716375	ACGTTGGATGACTCAACCTCTTAAGGTCC	ACGTTGGATGAAAGTGGCTTGTCTGTAGG	AATGTGCATTTTCTTTGGAGA	A	G
9	64.48	3664300	ACGTTGGATGGATTGCAAACGCATCAAAGG	ACGTTGGATGCCACACCTTTTGGCTTTGTG	GCCCTGTCTCAGTCTTATA	C	T
9	79.67	6213724	ACGTTGGATGACGCAAACCTTGACAGAATC	ACGTTGGATGCATTTGTCTCCAGGAAAGCG	TCAATACACCAGCACTCACA	C	T
9	94.24	3699737	ACGTTGGATGTCACTTAGTGTGGCAGAGG	ACGTTGGATGCTCAGGGAACATACCAAAAC	CAAGATCATACTGACCCACGT	A	G
9	113.73	3693662	ACGTTGGATGTGTACAGAGAGAGGT2CAGAG	ACGTTGGATGTGACCTGAGTTTCTACCCC	CCCTGACCCATGAGTGT	G	A
10	8.32	4228101	ACGTTGGATGCCAGAAAGCCTCAGAATGTCA	ACGTTGGATGGAAC2TCTTTGTAGGAAAGG	TATCAGATGTAAGTTGCAAAG	C	T
10	65.56	3725109	ACGTTGGATGCTGATTGTCTGTGGTGTGG	ACGTTGGATGTCTTGGCACTT2AGGACAAG	CAAGTTCAGATATCTGCT	T	A
10	76.37	13480657	ACGTTGGATGAGTGCCTTTTCTCTCTCTC	ACGTTGGATGACTTATCATTG2CACAGGGAC	CACAGGGACTCCCAAGA	T	G
10	89.05	3673574	ACGTTGGATGGAAGAGAAAATACCATTCCGC	ACGTTGGATGTATCATAACGTGACCCCTGC	CAGGGAAGGAGGATGCA	A	C
10	89.20	4139881	ACGTTGGATGTGCCCATGTTGTAGGATTG	ACGTTGGATGATCCAACCC2CAGAACCAATG	AGCCAGAGGTGGTGGCATA2T	A	G
10	99.01	4228444	ACGTTGGATGCGTGAGACACTCATGTGATA	ACGTTGGATGTGACTTACTCATG2CCCAAG	CTCATGCCCAAGTACTCAA	G	C
10	117.80	13480784	ACGTTGGATGAAAGCAGATAAACCTAAGGG	ACGTTGGATGCAGTTTCTA2ACTCCAGCTCG	CCGCCCCATCCAAGCC	A	G
11	14.33	13480881	ACGTTGGATGAAGAAATGTAATCTGTGGCC	ACGTTGGATGTCTGAGAACAGAATGCTGCC	CATTATTCTATACAGGGAGGA	A	G
11	45.10	13481004	ACGTTGGATGTGGAGAGCTCACATTTTCCC	ACGTTGGATGGCCACTGTAATACCTTGGTC	CCTTGGTCTGGGCTGCTGTG	G	A
11	59.47	3691800	ACGTTGGATGTCAAGGTTGGAGGCACTTTG	ACGTTGGATGTTCTCATCTG2CTCACTCAC	GCCTGGATAGTCAACAAG	G	C
11	76.55	3693943	ACGTTGGATGGCCTCAATGATTTCAATGCC	ACGTTGGATGTTACG2TGTGATGTCTGCTC	TCATGTCCCATTGTGGC	G	A
11	94.48	16783394	ACGTTGGATGTTGCTGGCCGGACTGATGTT	ACGTTGGATGAGGCCATCACAAAGATGCAG	ATGCAGTGGTAAATTTGGTGTAA	G	C
11	104.44	3153214	ACGTTGGATGAGCTCTTGAATGCTGCCTTG	ACGTTGGATGACATTCCTCTGCTGGC2ATAC	TGGCATACTACCC2CCGACAA	A	T
12	33.75	3723945	ACGTTGGATGTGTGAAGTCCGAATTTGGCC	ACGTTGGATGAAGGTATATGCCACTCCAC	CACCC2CAGCCAGATA	T	C
12	63.42	3686891	ACGTTGGATGCTGAAGACGCTAATTTCTCTC	ACGTTGGATGGAAACCTAACA2ACTACCAGG	GAGCTGTGACTTTCTGCAAT	A	G
12	74.85	13481543	ACGTTGGATGGAGAGGGATTTGGAGATTGG	ACGTTGGATGTCTGACACAGCTCAGGGAAC	TCCAGGA2ACTGACACAC	T	C
12	90.83	3684493	ACGTTGGATGACTGCCCTAAACATCATAGG	ACGTTGGATGAATTTGGACTCAGCAGGTTG	GTGAATAGTTAGCTTGTATATA	A	G
12	106.20	6390948	ACGTTGGATGACAGCTTGTGTGAACCTTGC	ACGTTGGATGTCTATAACCTGAG2GTGAGC	GGAGTGAAGT2GAGCCTGT	T	C
13	18.76	6345767	ACGTTGGATGGGTTTCACTGCTCCACTAAG	ACGTTGGATGGCATTACTCTCACTGAAGAC	ACGTAAGACTGGGCTGA	G	A
13	44.54	6249046	ACGTTGGATGCTTGAACCTCTGCTGTAC	ACGTTGGATGGAGCAACGATTCCCAAGTTC	AGACA2ATTGCCCTGAACTCC	A	G
13	61.70	3712411	ACGTTGGATGAGACTGTGACCAACTCTAAG	ACGTTGGATGGTCTGTCA2TCAATTCAGGG	TAGTCTGATACTTGCACCAAT	A	C
13	71.15	4229908	ACGTTGGATGCATCCAGTGACAATACTGGG	ACGTTGGATGGGATTTTCTACAGAGGATGC	TTTCTACAGAGGATGCAAAAT	T	C

Tabelle A.6: hME Markersatz Fortsetzung

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
13	89.96	3709305	ACGTTGGATGTCAAATGGGAGTCTTAGTC	ACGTTGGATGGTAACTTCTAGCCTGTACAC	GCCTGTACACATTACCCAA	A	C
14	11.93	4230190	ACGTTGGATGGAAATCGGTCCCATGTAAG	ACGTTGGATGACTACCAACTCCTCCCTTC	TTCCTGCACATACAGAG	G	A
14	63.62	3089070	ACGTTGGATGGGGAGAGTTGTTATAGACCG	ACGTTGGATGCCAGGTCCTTGAATTACTG	TCCTAGTTATTCTGTGGTCTT	T	G
14	81.15	3668028	ACGTTGGATGGGGCATGGGATTTAGTAGC	ACGTTGGATGCAAAAGTCAGGGGTTTCAGG	GTAAAACATTCAAATCCACTG	C	T
14	93.73	3688255	ACGTTGGATGTGTAGTCTGAAGGGTCTTG	ACGTTGGATGACACCATAATCCAAGTAGCC	GTAGCCAAGAGGCAGAG	T	C
15	5.01	4230638	ACGTTGGATGCCCATCCTTTGGATCTTTG	ACGTTGGATGTTCTTTCCCAAGATGCCCC	ACATATCTTTACACACCCAG	C	T
15	29.55	3656493	ACGTTGGATGGCAGGTAGCAGAAAAGATAG	ACGTTGGATGCCTGTGTAGATGATTTGAGTC	GAATGGTACATGCAATCAC	G	A
15	45.50	13482548	ACGTTGGATGCAGTCACTGGATAAGGCTTG	ACGTTGGATGCACCACACTTTAGGCTTTCC	GGCTTTCCACAATGCA	T	G
15	57.23	3702158	ACGTTGGATGCTCATCCAAGCAACTTGGC	ACGTTGGATGACTGGATACTGAGAAAGGCC	GGATACTGAGAAAGGCCAGTGT	A	G
15	79.20	3667621	ACGTTGGATGTACAGACTGGTGTGGGAAAC	ACGTTGGATGTCCACTGGTTCCACAGATG	TGGTTTCCACAGATGGATGAC	A	G
16	10.55	4161352	ACGTTGGATGTATTGAGAACCAGCAGGAG	ACGTTGGATGCCTCTGCTGCTTTCTTAG	TCTTTCCTTAGGGCCTCTC	T	C
16	27.21	3694564	ACGTTGGATGCTCACAGGCTCAGATATTTG	ACGTTGGATGACCTACTCCAAGTGAAGTAC	CATCTTCTAATCTACCCAAA	G	A
16	46.38	4182798	ACGTTGGATGAAGTGCCAGGTCGTTTCTC	ACGTTGGATGAGTTTGTCTTGCCTCAGGATG	AGGATGCAATGATGGGTACGG	T	C
16	54.50	3687272	ACGTTGGATGGGTCAGTCAACTGTGAGAAG	ACGTTGGATGAGCAACTAGAGCTGCATATG	TAGAGCTGCATATGAAGCATGA	G	A
16	74.56	3663889	ACGTTGGATGATTTTTGTCTAGACCCCGCC	ACGTTGGATGAAAAGAGCAAGGGAGTGAGG	TGAGTCAAGCACACTTTGCTGTTG	G	A
16	85.80	4212526	ACGTTGGATGCTATTCTGGACTTGTCCCC	ACGTTGGATGTTTTCAITCCAGCACCCGAG	GTTGGTGGCTCCGCACT	C	G
17	9.30	13482869	ACGTTGGATGCCATTTGGGTGGGCTGTATC	ACGTTGGATGCCATGAGAGAAGTCCTAACC	AGCTCACCTTGACTCCT	G	A
17	41.21	4231493	ACGTTGGATGAACTTTCGCCCTTCCCTCTG	ACGTTGGATGGGAGAAGGGTTTCAGAACAC	ATCCTTCCCAAGAGATGAA	C	T
17	79.29	2020834	ACGTTGGATGTGAGATGCTCAGACAAACCC	ACGTTGGATGGCTACCACAGTGGCTGTCAT	GTGGCTGTCATTAACCT	A	G
18	9.42	3725581	ACGTTGGATGATCTGGTGGGTAATCCTGTG	ACGTTGGATGGAAGTACTGCGAGTAGCCTG	TAGCCTGCTATGGCCTC	T	C
18	30.23	3664296	ACGTTGGATGAAAGATGCTCTGGCTGCTG	ACGTTGGATGCTGACCTCACAGAAATCTTC	CTTCTTTCTATGTACATGTT	A	G
18	43.66	13483329	ACGTTGGATGGTGGACTCTCAACTCCATAG	ACGTTGGATGTCCCTCTTACTCTIAGACAC	TAGACACATTCCCTTCCCC	C	T
18	58.85	3713935	ACGTTGGATGGACTTCTGCATGAGGATTTAG	ACGTTGGATGTGAAGGGAACAGAGATGAAC	ATTAATGGATATTCTTGACCA	G	A
18	77.15	3705890	ACGTTGGATGATTAGACCCTTAGTGTTACG	ACGTTGGATGTGGCATTGATACTGTGCC	GTGCCCATACAGATAAGAC	A	G
19	10.75	13483528	ACGTTGGATGGTAGTGCATCTCTACTTTTC	ACGTTGGATGACAAATACCTCCTCACCCAC	TCGTACAAGAAGCTATGATTA	C	T
19	24.80	3717368	ACGTTGGATGTTGAGATGGGATGGCTACTC	ACGTTGGATGATTGCCATGAGCCGTTATGG	ACTCTTCTACTAAGAACTTC	C	G
19	29.98	6238842	ACGTTGGATGGCAGCAGAGGCTTTACA AAC	ACGTTGGATGTTTCAGAGAATGCAGGGTGAG	GTGAGCTTAATCTGGGC	T	G
19	48.01	3674914	ACGTTGGATGAGATTGCAGTGGGCAACCG	ACGTTGGATGTTAAGGAGGCCCTCGGAGTC	TCGGAGTCTGTAACCCAA	T	C

Tabelle A.7: Sequenzen der Amplifikations- und Extensionsprimer des iPlex Markersatzes. Die Locusangaben beziehen sich auf Mouse Genome Assembly 35.

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
1	5.12	13475703	ACGTTGGATGGGCGTTTGTAGTAACAAGTG	ACGTTGGATGGCTCACACATTGAGGACTG	GACTGAATGCCAGGAGCCATA	G	A
1	23.69	13475764	ACGTTGGATGAAGAAATGGCCCTTCACGTC	ACGTTGGATGGCATTTTAACTGGACATGG	ATTTAACTGGACATGGTATGGG	A	G
1	38.28	13475818	ACGTTGGATGTCCTTTGAAGACAGGGTCG	ACGTTGGATGGCTACAATCTAGCACTGAG	CTAGCACTGAGGGACAGA	G	A
1	48.37	13475854	ACGTTGGATGAGGCCACAGAGACTTCAAAT	ACGTTGGATGAGGGCTTAACTTCCATAAC	AGGAATGCGCATATGAAAATTCCT	G	A
1	65.78	32716288	ACGTTGGATGCCACATCAAAATGCCAAACAG	ACGTTGGATGCTCATTGTCTGTATCTGTG	tTCTGTATCTGTGTGCAGTA	C	T
1	76.42	3678148	ACGTTGGATGAGTCCCCCAACAGATTTC	ACGTTGGATGGTACTGCTGGTTAGGATTTG	gagtAATTGAGAGTTTGACCCAATTC	A	G
1	93.17	13475986	ACGTTGGATGGAGACAGGAACAGGTGTAAG	ACGTTGGATGATGCCTCAATGGTGGCAAG	GGTGGCAAGTGCATTGATGACTG	C	T
1	114.72	13476065	ACGTTGGATGGCAATTTCTTGTAGCTCTCCTG	ACGTTGGATGTAAGACCTCCATTATGTGAC	CATATCATATGCATGTTATGCCCTGA	C	T
1	124.91	30551255	ACGTTGGATGTGCAGAGAGACTGTATAATC	ACGTTGGATGGCGCTGTGTTTGTCTATGG	ATAATAAAGCAACATAAAAATAAATGC	T	C
1	142.14	30942489	ACGTTGGATGTTTGGGTTAGCATGTAGGAG	ACGTTGGATGATTGAGGGATCCCAGAATG	AGGGATCCCAGAATGTAAACAGGA	A	G
1	158.02	31593281	ACGTTGGATGTGTAGTGCAGATGCACCTC	ACGTTGGATGGAAC TGCCCGTAAATCATCC	TGCTTGCTGAAATGTCAC	C	T
1	170.70	33777727	ACGTTGGATGGATCAGAAACATTCCATAGTC	ACGTTGGATGTACCGAAACCTTAGGAGTGC	TTCTCTCAGAGGCTCAG	A	G
1	188.78	13476296	ACGTTGGATGCACAGTCTTAATTACCACATG	ACGTTGGATGGGGAAAGTTGCTGACAATG	CTTCTTGGTGTITTTCTTCTTAGA	G	A
1	192.81	13499691	ACGTTGGATGCGGGAGGATGTGATTTATGG	ACGTTGGATGCTCGACACTTAGAAGTAGGC	TGCCATGTACCATGTC	A	G
2	14.34	13476355	ACGTTGGATGCTTGTCTTGGTGGCAGAC	ACGTTGGATGGTTGGACAGACATCTTAGC	CTCTAGCATTGCATTTTTTCA	G	C
2	20.35	27120459	ACGTTGGATGCATAGTTCCTCTAAATAC	ACGTTGGATGAGCTCTGACCACTATCTTG	CCAGTAATGACTAGCAAAAG	C	T
2	37.15	13476434	ACGTTGGATGTGCACCTTCTATGAGTGAAGC	ACGTTGGATGTCATCTATAGCCTGAGGAAC	GAAGAGAGGTTTAAAGGAAGTTAA	A	G
2	50.80	13476490	ACGTTGGATGCTGACTTTCCTGCTCTCTTG	ACGTTGGATGAGGATGCCACAAC TAGAATG	ACACACAGCATGGACGAA	C	T
2	70.83	13476567	ACGTTGGATGAGAGGAAAGACAGGAGCAAG	ACGTTGGATGACAGTAGCCAGAAGAAACAC	AAACACTTAGAAAAGGCTGTCA	C	T
2	81.19	28316257	ACGTTGGATGGTGTGACAAATGGTGCCAAC	ACGTTGGATGAGCTGTGGCCTAATATTCTG	aggTGTGAAAAAGCAAGAGGG	G	A
2	87.03	27459798	ACGTTGGATGCAAATTTTGGACATGTTAC	ACGTTGGATGAAAACACACTAGAGTTTGC	ATATTAGAAATATACTTGGCATCA	C	T
2	95.62	3679193	ACGTTGGATGCTATGGTTCTCTGTTTCTCTG	ACGTTGGATGATTTCCAGCCTTCTCAGGAC	ACCCTTCTCATGCTGA	C	T
2	114.15	27441842	ACGTTGGATGGTCTGTAACACTGAATAGGG	ACGTTGGATGGCTTAITCTCTGGTCAGTC	AAAGCTCATGAGTTC AACTGCAATACC	C	T
2	129.40	27257388	ACGTTGGATGTCTGCTATCCACTGCATTCC	ACGTTGGATGGGTTTGCTTATAGAGGCTGG	GAGGCTGAAAAGCTAGAAA	C	G
2	164.10	3696248	ACGTTGGATGTCCCACCTTGTTCCTCAAAGG	ACGTTGGATGTGTGGAGCCCAGTACTTAAC	CCCAGTACTTAACACATAGCCTTTT	G	A
2	169.30	13476909	ACGTTGGATGCAGGAATAGAGCCAAAGTGTG	ACGTTGGATGAAACAGCTTGTGAACCTACCC	AGTTACTCTGGGGTCTGGGG	C	A
2	182.39	3691120	ACGTTGGATGACCAAGCTGTGCTTAACATC	ACGTTGGATGAACAAGCACTGGCTTTTGGG	CATCCCATTAGTGGCTGATGCT	C	A
3	7.69	29657774	ACGTTGGATGAAACATCCGGTCAAGGTTAGG	ACGTTGGATGACAGGTACTTTGAGGAACCC	TTGAGGAACCCAGATTTT	C	T
3	26.20	13477026	ACGTTGGATGCCATAACCACAGGCTTGCTT	ACGTTGGATGAAACTCAGCAGCTTCATCAC	TTCATCACTGTTCTCTGA	C	T
3	36.80	3151604	ACGTTGGATGGGAAAAGTCAAGTGTGCACTC	ACGTTGGATGAGGTGTTGTAATAGCGCTTG	GAAAGGAAAAGGCACTGGT	A	G
3	52.83	3685081	ACGTTGGATGCAATCAACTCCCAGGATGC	ACGTTGGATGCTGTTCACCATCATGATGC	CATCATGATGCAATAATCCA	C	T
3	69.85	13477178	ACGTTGGATGACGTGCATGCAGGCAAATC	ACGTTGGATGAAAGAGAAGCCACAGTGGAC	ACTTACAGTGAATGTGATGCT	G	T
3	89.36	8259135	ACGTTGGATGTAAGTGTGGACTGGAGAACTG	ACGTTGGATGTTT CAGAAGACTCCGCCAGC	ACTCCGCCAGCTCTCCAGTAG	G	A
3	103.27	13477302	ACGTTGGATGCAGTTGTATCATCTGTCTGGG	ACGTTGGATGAGAGGGAAATGTCCCCAATG	CCCCAATGCTTACACAGAAAG	C	T
3	109.01	13477321	ACGTTGGATGCAGTTGTGTGCACAGAAC	ACGTTGGATGCGAATAGCTCATATACGAGG	CTGAAATTAACAACGTGGCATTTA	C	T
3	129.70	16799508	ACGTTGGATGAGGAGGGACTTAACTTTTC	ACGTTGGATGCCTCTTAGATGTTTTCTGCC	TGCCATAGTTGTGCTT	G	A
3	142.91	13477460	ACGTTGGATGATTCTCCAGCTGGCTCAITG	ACGTTGGATGTCAAAGGTAAGGTTGGGAGG	GCAGCTTAAATACAACCTTCT	C	T
4	10.97	27731305	ACGTTGGATGCTGTGTGTTTCAGCACAGG	ACGTTGGATGAGGACTCCAGGCTTATACAG	CAGCTCTTGAATTTGGTC	C	G
4	22.52	13477603	ACGTTGGATGGTTCCATGAAGTCTCTGACAA	ACGTTGGATGGGTAAGTTTCCACCATCAATC	TTCACCATCAATCAAATGTA	C	T
4	35.41	27781503	ACGTTGGATGTTGATTCACAGACTCTAGC	ACGTTGGATGGGGTACTATGTGACATTTG	TGTGACATTTGACATAAAGAAAAGGA	T	A
4	52.83	27780226	ACGTTGGATGTCATCTGCTGTTCCTACTAATC	ACGTTGGATGGTTTTTAAATCTTCTCATGG	ccTCTTCTCATGGTATGCTC	C	T
4	54.34	27794497	ACGTTGGATGGGTTTTGTGTAATGGCAGGG	ACGTTGGATGCGGAAGGAAAACAGTATCCC	GGATTTGTGCCAGAGA	T	C
4	62.51	3718270	ACGTTGGATGCCTCCTCTGTAGCAATG	ACGTTGGATGAACCTGCTTGCCCATATCC	CCTCCCTGAGCCCTC	T	C
4	86.57	28056583	ACGTTGGATGGAATGACTGGGCCTAAGAC	ACGTTGGATGAAATACGAAGGTGTCAGGGC	gTACGAAGGTGTCAGGCGAATCCT	C	T
4	100.87	28307021	ACGTTGGATGTACACACCTTCAACCCAGC	ACGTTGGATGTCAGCTTAGCCAACTTTGTG	gggTGTATCTATGTCTGTAAGGTA	A	G
4	117.25	13469808	ACGTTGGATGCAACCCTTAGAGCAGCTGTC	ACGTTGGATGTGACACTAACGGACCAAGGC	CCAAGGCCCTCCATGCT	C	T
4	132.94	13477989	ACGTTGGATGCAAGCCGCTCTTCAACATTC	ACGTTGGATGCATACCCATGTTATGCTG	TGGGAACCTAATAAATACAATGGG	C	T
4	141.72	3711383	ACGTTGGATGTTGCCCTTTTGTATCAGTGGAG	ACGTTGGATGTCCCCAGAAAGAGAATCAC	AGGAAGCCAAGTCCCTA	C	T

A. LISTE DER VERWENDETEN PRIMER

Tabelle A.7: iPlex Markersatz Fortsetzung

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
5	13.51	13481347	ACGTTGGATGTCAGGAGGCTGACTTCTGG	ACGTTGGATGCGCCAGCTATTCAITGAGAA	AGCTATTCAITGAGAAAACAGTCG	T	C
5	24.40	13478148	ACGTTGGATGGTGCACCTTGTAGCTTCTCAG	ACGTTGGATGACCAAACACACCTGAAACCC	CTGAAACCCAAAGCCC	C	A
5	40.00	13478204	ACGTTGGATGACACAACGGACATTGGAGAC	ACGTTGGATGCTGAGAATAAGACACTTCCC	TCCCTTAACCTTGATCTCTTTAGTGAG	A	G
5	54.62	13478263	ACGTTGGATGCCAGCACAAAATGCAACAAG	ACGTTGGATGGGCAACTTCAAATGAGCTTC	AAATTGAGCTTCATAGCCAT	G	A
5	66.93	29635956	ACGTTGGATGGGCTTAGCTGTCTAAGCTG	ACGTTGGATGATTCTCTTAGTGCTCGCTCG	TTATGATCTACAGAACAAACATGCCAA	G	A
5	78.89	31585424	ACGTTGGATGCCATAITTTAGCTCCCCAGG	ACGTTGGATGAGGAATGATGTATCATGGC	ataATGGCACAITTTCTGAAAACTT	T	A
5	102.10	13478429	ACGTTGGATGTCAGATTACCTCGTCTTCCC	ACGTTGGATGGATCTCTGTCCAAGGGTGC	ACTAGGAGACGTGCTAC	C	T
5	110.44	32067291	ACGTTGGATGGCGGTACCTGTAAACACTTG	ACGTTGGATGATCGCCCTGGCTAAGTGTCA	GAGACTGGATGTTGGG	G	A
5	125.91	13478514	ACGTTGGATGTTCCCCAGCACCAGGCAACT	ACGTTGGATGCATGTACAGGCACATACATTG	TGTTAATATGTCATTGCTACAGGAAA	G	C
5	143.28	4225559	ACGTTGGATGTGCTACTTTGATTTTGAC	ACGTTGGATGTCTGTCTCACCAGATCC	CCCCCTCCACACCCCT	C	A
6	4.78	13478606	ACGTTGGATGGAGAAGCAAGTGCATGCTG	ACGTTGGATGGCTACGTTTGTGTGAGCATC	ATCACATCTCATAAGAAAATCAC	G	T
6	26.12	13478670	ACGTTGGATGAGATCATCAGGGATCGCTTG	ACGTTGGATGGACTTTAGCCCGAGATTCAG	CCACACAGTTGTAAAATGTCAG	G	T
6	35.95	3023067	ACGTTGGATGTGCAATTGAGGTGTGATTATC	ACGTTGGATGGTGTGCTTGTCTTCTTTTG	TGCTTGTCTCTTTTGAGCTGAG	T	A
6	52.75	13478756	ACGTTGGATGAAAGCACTGACTCCTGTGAG	ACGTTGGATGATCCCACTGTACTGGACCAC	TGGATGGCGTGAGTCCATTGCAAAACA	C	T
6	72.40	13478816	ACGTTGGATGTGGAATCGAGAGCATGATCC	ACGTTGGATGAGCATGTGAGGAAAGGGATG	AGGGATGGGTTGAGAG	T	A
6	86.02	13478872	ACGTTGGATGTGGATAATCTACTCCCCCTG	ACGTTGGATGCAAGTGTCTACAGCAGAAC	AAGTGATCTACAGCAGAACAGGAAG	A	G
6	102.37	13478935	ACGTTGGATGCCTCTACATAGCTAAGGTGG	ACGTTGGATGCTCGTATTGTCCACCGTGC	TGCTGGTACCACACA	A	T
6	115.31	13478987	ACGTTGGATGGAAGACAGATCCACTGGG	ACGTTGGATGGCACGCTGTGTAAAGAATG	GTACCTTGAATATCACATCCTGA	C	T
6	137.65	16815348	ACGTTGGATGCACTTGACTATGGCTCTG	ACGTTGGATGCATGCAGCAAGTATGATACC	AGCAAGTATGTACCCACATTA	C	A
6	144.60	13479084	ACGTTGGATGTGGCACAGTGTGACCTATC	ACGTTGGATGGCAGCATAGAAAACCCACAG	CCACAGAAGCCTTGAA	T	C
7	4.87	13479108	ACGTTGGATGTGAAAGCTAGGAGCAGGAGG	ACGTTGGATGCGCCTGGTTCTAAATTAC	AAAGAAATTAATATCAACTGATTAGG	A	T
7	24.25	13479164	ACGTTGGATGCTTCTACAGTGTGCATCCG	ACGTTGGATGGACAATCATGAACACCCAG	ggCCTCTCCTCGGCAGCA	G	T
7	40.75	16793422	ACGTTGGATGTGCGGCGAGCAGGAAAAAGA	ACGTTGGATGGCTTATTGTGTGCCCAGG	CAGGGGAGGTGCAGGG	A	G
7	49.68	13479256	ACGTTGGATGTAGAGCAAGGCATTATGTG	ACGTTGGATGGAGGAAGAAAGTTAACAGGGC	GCCTGGGATAAATGGTT	G	A
7	61.88	16805799	ACGTTGGATGGCCCTGTACTCATAAGAG	ACGTTGGATGGGTTGGAAATGATGTAATG	TGATGTAATGACAAAGAGCCTAATGT	C	A
7	88.94	4226783	ACGTTGGATGTGCTTGTGGGATCCACAAGG	ACGTTGGATGAACTACCATTGTGCAAGG	CTGCAAGGGAGGGGAAA	G	C
7	99.24	13479437	ACGTTGGATGGCTCTGTCTTGATAACTACC	ACGTTGGATGGGATTTCTCTGCCTTTCTGC	TATCTTGTGGCCCTAATGG	G	T
7	112.81	13479476	ACGTTGGATGAGCAAGCACACATGGTAAGG	ACGTTGGATGGCCTGTGTATAGATCCTTCTC	gCCTTCTCATGGCATGCT	T	C
7	129.28	13479537	ACGTTGGATGACATGTCCAGCCACAATGC	ACGTTGGATGTGGCTAGCTTTGCTGTCTC	GCAGAAAGCCTTGTGA	A	G
8	9.73	13479604	ACGTTGGATGTGAAATCTGGCTCTGGTGG	ACGTTGGATGTCTTATCCTCTCTCAGTCC	CCTTGAATGTGAATGGTT	C	G
8	25.66	13479662	ACGTTGGATGCTTTGGTATACAGTGTGGTC	ACGTTGGATGTTCTTAAATGGTGACCTCCC	CCTAAATGGTGACCTCCATAATT	A	G
8	45.88	13479741	ACGTTGGATGATCTGTCTCCGTTTGTCTCC	ACGTTGGATGCTGAAAGGACTGTGCATGTG	GTGCATGTGAAGAAAAGA	A	G
8	56.43	13479782	ACGTTGGATGACTTGACCAAGAGAAAAACAG	ACGTTGGATGGCACACACATACACACAC	CACATACACACACTCTACTG	G	A
8	66.78	13479814	ACGTTGGATGTGGATGTGTGTTTTCAGTC	ACGTTGGATGGCCTATAGAGCAGGTTCAATTG	TATTTGACCAGTTCTGGAAAT	T	C
8	86.73	6264181	ACGTTGGATGAGAAACCAAACCTCCTTCAG	ACGTTGGATGCCTCTTGGTTTAAACATTGGC	gCATTGGCATTGATTTGCTATTTACA	T	C
8	99.50	13479952	ACGTTGGATGCATGAAAGTTGATCTAGTTG	ACGTTGGATGTACTTCTGTGTACATGTGC	ccCTTCTGTGTACATGTGCACAGTGT	T	C
8	112.76	13479998	ACGTTGGATGTTTCTTCTCATGTGCAGGG	ACGTTGGATGGCATGGGCTATTTGGTAATT	cTCATTCAATGTACACCTTTGTAT	C	G
9	13.27	13480073	ACGTTGGATGCTTTTGTCTGACTTCCCTTCC	ACGTTGGATGACGGAACAGGGTGTGAGAAAC	gGGAAGAGTTTGTAGTGCACACAG	G	A
9	24.50	6183014	ACGTTGGATGGGTTTCCAAAGGTCTGTTTA	ACGTTGGATGCAAGGCTTCAAATATATGGC	TCAAATATATGGCATGAGTACTA	C	T
9	37.62	3023207	ACGTTGGATGCTCAGACTCTTCTAGCTACC	ACGTTGGATGCCCGCAGAAAGCAGAATTAG	CTGAAGAGCTGGGGAG	A	G
9	57.85	13480217	ACGTTGGATGAACCTCCAGCTGTTTCTTG	ACGTTGGATGAGTCTCCTAATCTCTCAGCC	TCAGCCTTTTCCCTCTTA	G	A
9	65.23	13480245	ACGTTGGATGGCCCTAGGCTCTGACATTTT	ACGTTGGATGACTTTGGTTCTGCCACTCTC	CCACTCTCTGATCTCTACT	C	T
9	80.29	13480299	ACGTTGGATGCACCTTCCAAGCTCTTATCG	ACGTTGGATGTGCTGTTTGCAGAAAGAGTG	GGGTGGATCTGAGATGCTGTGAGAC	C	T
9	95.97	3673055	ACGTTGGATGCATCGAGTACTTTAGCTGG	ACGTTGGATGGAGTCTACTTTCAGATCCAC	TACTTTCAGATCCACTTTAATT	C	T
9	110.76	13480418	ACGTTGGATGCAGTCTGGTACTTTGAGAG	ACGTTGGATGATCTCCAGTACCCCAATTC	CTTCTACCTTTTAAACGGTC	C	T
10	8.17	13480484	ACGTTGGATGAGCACCTTCCAGAAAGTTCAC	ACGTTGGATGGTAGGAGATCTCCATCCTG	cCCATCCTGGAGATTATTTTTTTT	T	A
10	22.68	13480541	ACGTTGGATGCGAGTAACTGTATGATGGAG	ACGTTGGATGCTTCTGCTTAAAGCCAACAAG	cGAGGTTTGTGGGCACA	C	T
10	34.36	13480578	ACGTTGGATGGAACAGTTGAGCAGTCCAGG	ACGTTGGATGCCTCAGAGGCAGCATAGAAA	TCTTTTCCATGCTGTATCAGAAA	C	A
10	57.57	13480619	ACGTTGGATGACTGGGCTTGTCTTATGTTG	ACGTTGGATGAGCTTTGACAACAGGCTTCC	CCTACTCTTTTGTGTTTGT	C	T

A. LISTE DER VERWENDETEN PRIMER

Tabelle A.7: iPlex Markersatz Fortsetzung

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
10	69.05	13480638	ACGTTGGATGTGTTACAGGTAACCAACGC	ACGTTGGATGGAGGTCACCTGTGAGAATTG	CTGCTTTTAAATGCACCTCT	G	T
10	84.23	13480678	ACGTTGGATGAAAGGGCACATGAAGCTCAG	ACGTTGGATGCAGCTGGCCATTAAGAAATC	GCCATTAAGAAATCCTGTACCTC	A	G
10	99.58	8258500	ACGTTGGATGAAAGCGAACAATGGCCATAG	ACGTTGGATGCTTGCACAACACCTACTACG	TAATAAAAGATCCATCACATAATG	C	A
10	117.80	13480784	ACGTTGGATGCAAGTTTCTAATCCAGCTCG	ACGTTGGATGAAAGCAGATAAACCTAAGGG	CAGATAAACCTAAGGGATGCCGAC	G	A
11	7.16	13480851	ACGTTGGATGCAGGGTGAGAGGAGACTGTA	ACGTTGGATGTGTGCTCTCAAGGAAAGCTC	ACAGCGCATGTCCACCCCGT	A	G
11	21.67	13480905	ACGTTGGATGGGATGCTGATGGCCTCATAT	ACGTTGGATGCCTGGTAGACCTCATCTTAC	ACCTCATCTTACAGTACATTCACTACA	A	G
11	32.28	26822879	ACGTTGGATGTACCATCATCTTGGCTCTG	ACGTTGGATGTAGAAACCAGAGACACAGGC	AACAGATTGTACAGCCAGGA	A	T
11	54.13	26982471	ACGTTGGATGTAACCTTTTTGGCGCTGGTG	ACGTTGGATGGGAGAGAGGCTAGCTTTTG	AGGCTAGCTTTTGTCTTAAAAAGT	T	A
11	63.07	13481061	ACGTTGGATGTTCCGGTGATTCGGGTTACTG	ACGTTGGATGACTCCCTCTGATACTGCTTG	cTACTGCTTGAAAAACACAACCC	G	T
11	83.48	13481127	ACGTTGGATGAGGCAAGCAGAGGGAGCAA	ACGTTGGATGCTCCCGGCTGCATATTC	TCAGCCACCGTCAGGA	G	T
11	98.80	27041242	ACGTTGGATGGGTCATTCATGGTGCTCAAC	ACGTTGGATGGGCTTGTGTGCATTCAGAG	AGAGCCCTCGGACATAG	A	G
11	114.58	27000576	ACGTTGGATGGCCCTGGTTGTAAACATGTG	ACGTTGGATGAGCAGAGGAAAGCACAAAAGC	AGGAAAGCACAAAAGCAGCCGCC	C	T
12	13.21	13481307	ACGTTGGATGCAAGAAGGAGAGGCCACAAG	ACGTTGGATGCCACATCTAGTCTTCAGGAG	CAGGAGCCTCCCAACCACTC	C	A
12	22.94	13481351	ACGTTGGATGAGTTTATGGCACCCGGGATTC	ACGTTGGATGCAGGACAGGCGATGAAAAC	CTGCCGCCAACACTCTCAG	G	A
12	37.66	13481405	ACGTTGGATGTGGAATTTCTCAAAGGTCAG	ACGTTGGATGAACTGTAAACGAGTAAGCGG	ATATTTCTTCTGCATCTGTATGA	T	A
12	52.66	13481462	ACGTTGGATGTGGCTGGCATAGGAATTTT	ACGTTGGATGTTTCAGGCGGCACAAAATCAG	CAGCAAGGATGCAGGCAAGGA	A	G
12	71.89	8259450	ACGTTGGATGTGGCAGTGATGGTAGGTTTC	ACGTTGGATGAAATTTCTCCAAGCCCTCCAAG	CCAAGTATGAGCACAGTTAC	C	T
12	80.33	6194112	ACGTTGGATGCCAGGACAATCTAGGAAGAG	ACGTTGGATGAGAATTTCTGCTGAGCACC	gTAGCACCTACCCTGGGATA	G	T
12	95.57	13481604	ACGTTGGATGGCCATTAGGTAAGAACCAC	ACGTTGGATGATTCTCACTACCTTGGAGTC	CCTTGGAGTCTTAATGTGAA	C	T
12	110.35	13459138	ACGTTGGATGAAATGAGCCAGGATGGTGTG	ACGTTGGATGACTGTACGACACAGGATCTC	CTGTGGACACCAGCGG	G	A
13	6.19	13481676	ACGTTGGATGCTCTCTGTCTGAAGGTC	ACGTTGGATGTTGTGAACCTGGGCACAATG	TGCCACATGCAGCTAC	G	A
13	18.76	6345767	ACGTTGGATGGGTTTCAGCTGTCCACTAAG	ACGTTGGATGGCATTACTCTACGTAAGAG	CGTAAGACTGGGCTGA	G	A
13	42.28	13481783	ACGTTGGATGCCATTATATTGAATGGCC	ACGTTGGATGGCTGGTTAGAAAATGCATAG	aattCATAGCAGGTAGGACCAA	C	T
13	53.48	13481815	ACGTTGGATGCAGGAGATCTGGAAGTAAGC	ACGTTGGATGCTGCCTCTTGTCTCTTACC	CCCCCTCAGCCTTGGCTCT	C	T
13	66.12	13481863	ACGTTGGATGAAATGCCATTTCTGGATGTGCC	ACGTTGGATGCTATGTACACTCAAGAGTGC	ACTCAAGAGTGCCTTCTCC	C	A
13	80.10	13481910	ACGTTGGATGTGATGTCGCTCAGCAAGAGG	ACGTTGGATGTGTCTAATATCACGCTTGGCC	aacgGGCTCACCAGAGAACAGGGA	G	C
13	93.55	29566800	ACGTTGGATGCCATGCCACACTCTTTATG	ACGTTGGATGTACCTGTGTGCATCTCTGTG	AGCAGATGCAGTTAGAAT	G	A
13	107.57	30511458	ACGTTGGATGATCCTTACGCTTCTCTGAC	ACGTTGGATGGAGCCTGGCATTCTTAAGAG	TAAGAGTAGGCAGATAGAAA	C	A
14	19.76	30406796	ACGTTGGATGCTCTGGTGCATCATATGG	ACGTTGGATGAGAGCAGTCTCATGTTCTC	CAACTACGTAGATTGTGC	T	C
14	41.38	13482161	ACGTTGGATGGAATATACACACAGCTTGGC	ACGTTGGATGAGGCAGCAGACAAAATCAG	TCCCTCCTAAGACCTG	C	T
14	53.85	30895903	ACGTTGGATGGTGTATGGCAGGATACCATT	ACGTTGGATGCCTTTCACAATCTCTCTCCTG	TTCACAATCTCTCTGGACTCA	G	A
14	68.36	30865397	ACGTTGGATGGAGAGTGAAGGAAGCAGGTC	ACGTTGGATGCACACTACTACACTTACCAC	ccCACTTACCACATTTGTATCA	T	C
14	82.88	13482292	ACGTTGGATGGGTTAATGAAACGTTACC	ACGTTGGATGTAGTTACTGAGAGACACATC	AGCTTATTTCTTAAAAGAGACT	C	T
14	99.74	30482696	ACGTTGGATGGCCTAAAGTACTAAGTTATCC	ACGTTGGATGTCTTATCCAAAGACTGCTC	TCATCCAAGACTGCTCTATAGTTGGT	A	G
14	112.16	30947935	ACGTTGGATGCTGTCTCCAGAGAATATGGC	ACGTTGGATGAGAGAAGGCAACCAAGGAGG	CTCGTGGATGGGAAGT	C	T
15	25.38	13482484	ACGTTGGATGGTGTATGCCCTGCCTATTTT	ACGTTGGATGCCTGTTGTAGTGGCAGTTAC	AGTGGCAGTTACTGGGTT	G	T
15	39.00	13482528	ACGTTGGATGCCGAATCTCAGTGCTAGAGG	ACGTTGGATGATGACGCCCTCCAATGCATC	CTGACCTGGATGGGATG	G	A
15	50.64	13482574	ACGTTGGATGGGAAGGGCTTATTTCTATGGG	ACGTTGGATGGCTTATAGTAGCCAGAGCAC	cGCCAGGCACTACCATCGTT	T	A
15	67.81	13482627	ACGTTGGATGTGGTAGAATGGTGTGGTAGG	ACGTTGGATGCATGAGACACAAAGTAGGGC	gACCTGGATGGTCTATGGAA	G	T
15	85.87	16820334	ACGTTGGATGCTCAGCTGGGAGAGAACTTG	ACGTTGGATGACCATGCCCTGGCTTATGGG	TTCTGAGGAGCCAACC	C	G
15	97.89	16804751	ACGTTGGATGATCATTGCAGGAGAGCAGAG	ACGTTGGATGCCACTTCTATCTGGATCTTGG	TCTGGTGAAGCACTGG	G	A
16	10.55	4161352	ACGTTGGATGCCTCTGTCTTCTCCTTAC	ACGTTGGATGTATTGAGAAGCAGCAGGAG	AGTGTCCCCAGAGAA	C	T
16	27.10	4165602	ACGTTGGATGCAACTGGAAGATGTGCTATC	ACGTTGGATGGCCAACAACCTCAGGGATT	CAATTCCAGTTGATTTCTATAAGA	C	T
16	31.86	4170048	ACGTTGGATGGTCAAAGGAAAGTCAGGGC	ACGTTGGATGCTCCCTTCACTGTGATCAAC	TGTCTGTTGGGTGTGG	G	C
16	51.24	4186801	ACGTTGGATGCGCTAAGCTTCCCATGTTG	ACGTTGGATGAGCTGGTGTATTTGACCTG	ggcgCATGAAGCAGAGCCT	A	T
16	69.75	4199268	ACGTTGGATGCTAAGAGGCACTGCTAATT	ACGTTGGATGTTGCCCTGATTTGGTGGAA	ACAAGTCAIAGTCACTCTTAACT	G	A
16	84.13	4211770	ACGTTGGATGGTAAATTTCTCTGACATTC	ACGTTGGATGCAAATCCAAATGGAAGAGC	GGAAGAGCTCATGTGTACACTG	T	C
16	93.59	4220159	ACGTTGGATGGGTTAGGACACAATAGATAAG	ACGTTGGATGTACACCATTAGAGTTCTCC	TTCTAAAATTACCTTGCCCTT	C	T
17	10.11	33418817	ACGTTGGATGAGGGTTTTACTGGACGAC	ACGTTGGATGAGTGCATTTGTGCATCTGTG	TCAAAGTGAATCCTGCGAATTTAC	C	T

Tabelle A.7: iPlex Markersatz Fortsetzung

Chromosom	Locus [Mb]	refSNP	Sequence forward (5'-3')	Sequence reverse (5'-3')	Extensionsprimer	Allel 1	Allel 2
17	26.23	33259283	ACGTTGGATGTGTGGCTGGCGTTCAGTAG	ACGTTGGATGTTCTGGAATAAGCTGTGTC	ACATTTTCCTAGTGTGGAG	C	T
17	38.85	33428427	ACGTTGGATGACTAATAATTAAGTGACTG	ACGTTGGATGCAGTAAACATGGCAATATTGG	GTATAATCAAGTCTTGACGAT	T	A
17	53.33	29504995	ACGTTGGATGAACCAGGCTGTTGGTCATAC	ACGTTGGATGAGGAGTGTCTGTTCCAATGC	CTGTTCCAATGCAACCTGCAGACC	C	T
17	70.65	13483097	ACGTTGGATGAAGTGGATCCTGAGAAACCC	ACGTTGGATGAGTGATTCTCTCCCTCCC	CCCCACCTTTTCTCTGTCTG	A	T
17	83.36	13483140	ACGTTGGATGTAACGTCTCTCCTCGTGGTTC	ACGTTGGATGATTCTCAGCTCTGGAGGAAG	GATTTAGCACAGTTGATGA	G	A
18	8.87	29556953	ACGTTGGATGTCTTCTCCTCCATGCATTC	ACGTTGGATGCCTGTGCTGAAGTCAATACC	TCAATACCACACCTGC	C	A
18	25.73	29827614	ACGTTGGATGCCCCAGCTAGTATTGAAGAC	ACGTTGGATGACTGAGCAGTCACTATCTCC	AGTCACTATCTCCCCAGTCAACAGT	C	T
18	38.46	29823686	ACGTTGGATGCAGAGAGAGAGGTGTTTCAG	ACGTTGGATGATGTGACTGTGGCAGTTGTG	TTCTGCCTAGGGCAGT	T	C
18	58.05	13483379	ACGTTGGATGATGCAAGCTCTCTACAAGGG	ACGTTGGATGGGACCAGAGCATGATTTTCC	gCCGGTCACCTGTTCTCCGCTGTCCTA	C	T
18	70.86	13483427	ACGTTGGATGAAGGAGCAAGGACTGTCATC	ACGTTGGATGGAGTTGCTTATCCTGAAGGC	TCTGCTAAAGACTTACAAAACC	G	A
18	86.93	13483484	ACGTTGGATGTGACAGCAATGGGAAGAAGG	ACGTTGGATGTCCTTCTGCTTCCCTGTTCC	gCCAGCAAGGCTGCTATGG	G	A
19	13.56	6247194	ACGTTGGATGGATCTTGCTTTGAAACAAGG	ACGTTGGATGGGAAAGCTCTCAGTTTGTG	AGTTGAATTGAATACTTACGA	C	T
19	25.63	13483576	ACGTTGGATGGGTCCACAGCAATATTTTCC	ACGTTGGATGGTGTTCATGTGTGAGAGGC	GAGGCCACTACTTGCG	T	C
19	43.39	4232188	ACGTTGGATGAGTTGCCTCCAGTCTGACAC	ACGTTGGATGGGCCAGTTCTTACTATGC	CATTTCTAGCGTCTTCCATTACAGGC	A	G
19	56.18	6339594	ACGTTGGATGAACCGCTGGATCAGATGAGG	ACGTTGGATGTCTCCTCAGATACTGTCTCG	tcGACCCTCACACACA	G	A

B. Danksagung

Mein besonderer Dank geht an Herrn Prof. Dr. Martin Hrabé de Angelis für die Überlassung des Themas und für die Hilfestellung und die Diskussionsbereitschaft bei dessen Bearbeitung. Hervorheben möchte ich dabei insbesondere das in mich gesetzte Vertrauen bei der Durchführung dieses Projekts und für die Freiheit, die ich bei dessen Umsetzung genoss.

Desweiteren möchte ich mich besonders bei Herrn apl. Prof. Dr. Jerzy Adamski bedanken, einerseits in seiner Funktion als Betreuer dieser Arbeit an der TU München und andererseits als Leiter des Genom Analyse Zentrums der GSF, in dem insbesondere die SNP-Genotypisierungen und die DNA-Sequenzierungen stattfanden.

Frau Dian Michel möchte ich für die unermüdliche Diskussionsbereitschaft speziell zu Beginn meiner Arbeit danken. Herrn Dr. Koichiro Abe möchte ich ebenfalls für die vielen anregenden Diskussionen danken. Mein großer Dank geht außerdem an alle Technischen Assistenten der Arbeitsgruppe und der Deutschen Mausklunik und an Frau Dr. Sibylle Wagner und das gesamte Team aus der Mauszucht unter ihrer Leitung.

Herrn Dr. Norman Klopp und Herrn Dr. Peter Lichtner möchte ich für die Diskussionsbereitschaft und die Hilfestellung bei der experimentellen Umsetzung der Genotypisierung danken. Bei Frau Dr. Isabel Rubio-Aliaga und Herrn Dr. Johannes Beckers bedanke ich mich für die kritische Durchsicht der Rohfassung dieser Arbeit. Ebenso geht mein Dank an Thomas Lisse, der alle englischen Texte, die im Rahmen dieser Arbeit entstanden sind, korrigiert hat. Herr Dr. Helmut Fuchs und Frau Dr. Birgit Rathkolb haben sehr viele Kollaborationen organisiert, zu denen ich meinen Teil beitragen konnte. Dafür möchte ich mich recht herzlich bedanken. Frau Christina Hoefler möchte ich gerne für die vielen Diskussionen und den Austausch von wissenschaftlichen „Spinnereien“ danken.

Ganz besonders möchte ich meiner Freundin Nadine Lindemann danken, die es in der letzten Zeit nicht immer leicht mit mir hatte, mich aber stets liebevoll und geduldig unterstützte. Meinen Eltern möchte ich ganz herzlich danken, weil sie immer an mich glaubten und mir auch in schwierigen Zeiten stets zur Seite standen.