Identifizierung und epigenetische Regulation von neuen geprägten Genen

von

Diplom-Ingenieur

Nico Ruf

von der Fakultät III - Prozesswissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender	Prof. Dr. Helmut Görisch
Berichter	Prof. Dr. Ulf Stahl
Berichter	Prof. Dr. Peter Nürnberg
Berichter	Dr. Ulrich Zechner

Tag der wissenschaftlichen Aussprache: 14. November 2006

Berlin 2006
D83

Teilergebnisse aus dieser Arbeit wurden vorab in folgenden Beiträgen veröffentlicht:

Publikation:

Ruf N, Dünzinger U, Brinckmann A, Haaf T, Nürnberg P, Zechner U. Expression profiling of uniparental mouse embryos is inefficient in identifying novel imprinted genes. Genomics. 2006 Apr;87(4):509-19

Vortrag:

Ruf N, Dünzinger U, Galetzka D, Bähring S, Haaf T, Nürnberg P, Luft FC, Zechner U. Expression profiling of uniparental mouse embryos is inefficient in identifying novel imprinted genes. Medizinische Genetik 1, März 2006, 17. Jahrestagung der Deutschen Gesellschaft für Humangenetik 8.-11. März 2006, Heidelberg, Deutschland

Poster:

Ruf N, Dünzinger U, Wagner M, Haaf T, Nürnberg P, Zechner U. Systematic identification of novel imprinted genes using Quantification of Allele-Specific Expression by Pyrosequencing, European Journal of Human Genetics 12 Supplement 1, European Human Genetics Conference, 12.-15. Juni 2004, München, Deutschland

Ruf N, Dünzinger U, Konersmann M, Haaf T, Nürnberg P, Zechner U. Quantification of Allele-Specific Expression by Pyrosequencing identifies three novel imprinted transcripts in the mouse transcriptome. Medizinische Genetik 1, März 2005, 16. Jahrestagung der Deutschen Gesellschaft für Humangenetik 9.-12. März 2005, Halle/Saale, Deutschland

Ruf N, Bähring S, Haaf T, Nürnberg P, Luft FC, Zechner U. Quantification of Allele-Specific Expression by Pyrosequencing identifies three novel imprinted transcripts in the mouse transcriptome. Environmental Epigenomics, Imprinting and Disease Susceptibility conference, 2.-4. November 2005, Durham, NC, USA

Inhaltsverzeichnis

1. EINLEITUNG 6

1.1. Epigenetik 7
1.1.1. Epigenetische Mechanismen 7

1.2. Genomische Prägung 9
1.2.1. DNA-Methylierung und genomische Prägung 12
1.2.2. Funktionelle Gemeinsamkeiten geprägter Gene 14
1.2.3. Gestörte Prägung und resultierende humane Erkrankungen 16

1.3. Hypothesen zur Evolution der genomischen Prägung 18
1.3.1. Genetische Konflikt-Theorie 18

1.4. Identifizierung neuer geprägter Gene 20
1.4.1. Experimentelle Ansätze 21
1.4.2. Sequenzbasierte Vorhersagen 24

1.5. Validierung von Imprinting-Kandidatengenen 26

1.6. Aufgabenstellung 27

2. MATERIAL & METHODEN 28

2.1. Material 28
2.1.1. Mausstämme 28
2.1.2. Humanes Material 28
2.1.3. Oligonukleotide 29

2.2. Methoden 30
2.2.1. Allgemeine molekularbiologische Methoden 30
2.2.2. RNA- und DNA-Isolierung 30
2.2.3. cDNA-Synthese 30
2.2.4. Polymerase-Kettenreaktion 31
2.2.5. DNA-Sequenzierung 31
2.2.6. Allelische Expressionsanalyse mit Pyrosequencing™ 32
2.2.7. Sondenpräparation zur Northern-Blot-Hybridisierung 35
2.2.8. Hybridisierung der Northern-Blots 35
2.2.9. Real-time-RCR mit TaqMan® zur Expressionsanalyse 36
2.2.10. Bestimmung der Transkriptlänge durch RACE-PCR 36
2.2.11. Datenbankrecherchen und Sequenz-Alignments 37

3. ERGEBNISSE 38

3.1. Imprinting-Kandidatengene der Microarray-Studie 38
3.1.1. Auswahl der Kandidatengene 38
3.1.2. Identifizierung von transkribierten SNPs 39
3.1.3. Allelische Expressionsanalyse durch QUASEP 39
3.1.4. Erhalt der Prägung im adulten Stadium 45
3.1.5. Genomische Organisation der geprägten Transkripte 46
3.1.6. Weitere Charakterisierung von AK043056 49
3.1.7. Identifizierung des humanen Orthologs von *Peg14* 50
3.1.8. Monoallelische Expression des humanen Orthologs 51
3.1.9. Vergleich der orthologen Transkripte 52
3.1.10. Allelische Expressionsanalyse von *USP29*, *ZIM3* und *ZNF264* auf HSA19q13.4 53

3.2. Imprinting-Kandidatengene der sequenzbasierten Studie ... 55
3.2.1. Auswahl und Analyse der Kandidatengene ... 55
3.2.2. Allelische Expressionsanalyse des humanen *KCNK9* ... 57
3.2.3. CpG-Inseln in der *Kcnk9*-Promotorregion ... 58

4. DISKUSSION ... 59

4.1. Microarray-Studie ... 60
4.1.1. Transkript AK039514 auf MMU17prox ... 60
4.1.2. Transkript AK018142 auf MMU12dist ... 60
4.1.3. Transkript AK039482 auf MMU7dist ... 61
4.1.4. *Peg14* (AK043056) auf MMU7prox ... 62
4.1.5. Präferentiell stammspezifisch exprimierte Transkripte ... 63
4.1.6. Biallelisch exprimierte Kandidatengene ... 63
4.1.7. Zusammenfassende Bewertung ... 64

4.2. Genomische Prägung auf MMU7prox und HSA19q13.4 ... 64
4.2.1. *Peg14* – eine nicht-kodierende RNA ... 65
4.2.2. Nicht-kodierende RNAs auf MMU7prox ... 66
4.2.3. *Peg14* und die Imprinting-Regulation ... 67
4.2.4. Zusammenfassende Bewertung ... 68

4.3. Sequenzbasierte Vorhersage neuer geprägter Gene ... 69
4.3.1. Die Effizienz von sequenzbasierten Vorhersagen ... 69
4.3.2. Genomische Prägung von *Kcnk9* ... 71
4.3.3. Zusammenfassende Bewertung ... 72

4.4. Vergleich beider Ansätze ... 73

4.5. Eignung von QUASEP zur Imprinting-Analyse ... 74

4.6. Ausblick ... 75

5. ZUSAMMENFASSUNG ... 76

6. LITERATURVERZEICHNIS ... 77

7. ANHANG ... 86

7.1. Danksagung ... 86

7.2. Publikationen ... 88

7.3. Oligonukleotide ... 89

7.4. Identifizierte tSNPs der Kandidatengenen ... 100

7.5. BLAST-Ergebnis AK043056 gegen das Humangenom ... 106

7.6. Abkürzungsverzeichnis ... 107

7.7. Abbildungsverzeichnis ... 108

7.8. Tabellenverzeichnis ... 109

1. Einleitung

Mit dem zuletzt entzifferten Chromosom 1 (Gregory *et al.*, 2006) ist die vollständige Sequenz des humanen Genoms bekannt und ein bedeutender Meilenstein in der Genomforschung erreicht. Gerade das Wissen über die DNA-Sequenz unseres Genoms sollte die medizinische Forschung durch die Aufklärung zahlreicher erblicher Erkrankungen revolutionieren. Bis auf seltene monogene Erkrankungen, bei denen zu über 2.300 Genen ein kausaler Zusammenhang belegt ist (siehe http://www.hgmd.cf.ac.uk/ac/), sind jedoch bisher kaum Erfolge bei den häufigeren komplexen Erkrankungen zu verzeichnen. So zählt die nachgewiesene Assoziation zwischen drei Varianten des Gens *NOD2* und Morbus Crohn zu den wenigen Erfolgen (Mathew und Lewis, 2004; Farrall und Morris, 2005; Hirschhorn und Daly, 2005).

Darüber hinaus erhoffte sich die Wissenschaft von der Entschlüsselung der Genome verschiedener Spezies auch Unterschiede zwischen oder innerhalb einer Spezies ursächlich erklären zu können. Die bisherigen Daten für *Homo sapiens* mit 21.206 Genen, *Mus musculus* mit 20.306 Genen, *Canis familiaris* mit 16.973 Genen und *Bos taurus* mit 17.238 Genen (siehe www.ensembl.org) wie auch die erstaunliche Homologie der meisten Gene zwischen den Spezies, weisen jedoch darauf hin, dass weder die DNA-Sequenz noch die Anzahl der Gene ausreicht, um überzeugend zu erklären, warum sich der Mensch von Nagetier, Hund oder Rind so deutlich unterscheidet.

Letztlich bleibt es auch trotz Kenntnis der Sequenz unseres Genoms ein Rätsel, warum sich beispielsweise Muskel- und Leberzellen eines Individuums deutlich unterscheiden und die Unterschiede stabil nach der Zellteilung bestehen bleiben, obwohl beide Zellarten das gleiche Genom beherbergen. Was hier auf zellulärer Ebene zu beobachten ist, lässt sich auch bei Organismen finden. So ist es bekannt, dass monozygote Zwillinge eine sehr unterschiedliche physische Konstitution aufweisen können und auch unterschiedliche Krankheiten entwickeln (Fraga *et al.*, 2005). Als Erklärung diente bisher die Interaktion von Umwelt und Genen. Welche Mechanismen dem zugrunde liegen, ist nach wie vor unklar (Qiu, 2006). Diese fehlende starre Verbindung zwischen Genotyp und Phänotyp könnte in einem Phänomen begründet sein, welches als Epigenetik bezeichnet wird. Hinter

dem Begriff steht ein kryptischer chemischer und physikalischer Code, der über die DNA-Sequenz geschrieben ist. Dieser Code ist vererbbar und steuert die Aktivität von Genen unabhängig von der DNA-Sequenz. Einfach ausgedrückt bedeutet dies, dass wir mehr sind als die Summe unserer Gene. Somit ist nicht allein die Sequenz der DNA wichtig, sondern auch wie die Aktivität von Genen beispielsweise durch die Interaktion mit der Umwelt reguliert wird (Qiu, 2006).

Das Wissen darüber, wie eine bestimmte Zelle ihr einzigartiges genomisches Programm absolviert, würde dazu beitragen, ein besseres Verständnis über die Entstehung oder die Entwicklung von Erkrankungen zu gewinnen, und letztlich eine Antwort auf die Frage zu finden, was unter biologischen Gesichtspunkten das Menschsein ausmacht (Qiu, 2006).

1.1. Epigenetik

Der Begriff „Epigenetik“ wurde erstmalig 1942 von Conrad Waddington (1905-1975) verwendet, um Interaktionen zwischen Umwelt und Genen und deren Bedeutung für die Embryologie zu beschreiben. Ohne molekularbiologische Beobachtungen beschrieb er dennoch sehr treffend den Epigenotyp so: „Ein allgemeines Merkmal ist es immerhin, dass er eine Verkettung von Prozessen darstellt, die derart zu einem Netzwerk verbunden sind, dass eine Störung in irgendeinem frühen Stadium weitgehende Abnormitäten in verschiedenen Organen und Geweben verursacht.“ (Waddington, 1942). Demnach determinieren Gene und Umwelt einen Phänotyp durch den Epigenotyp (Horsthemke, 2005).

Die moderne molekulare Epigenetik beschäftigt sich mit Vererbung und Variation von Genexpressionszuständen, die im Wesentlichen durch Chromatinmodifikationen und DNA-Methylierung reguliert werden. Derartige Modifikationen beeinflussen die Struktur des Chromatins und bestimmen somit beispielsweise, ob Bindungsstellen für Transkriptionsfaktoren zugänglich sind, d.h. ob Gene aktiv oder inaktiv sind (Horsthemke, 2005; Murrell *et al.*, 2005). Analog zur Genomik wird bei genomweiter Betrachtung der Epigenetik der Begriff „Epigenomik“ verwendet.

1.1.1. Epigenetische Mechanismen

Das Genom eukaryotischer Organismen lässt sich in Euchromatin und Heterochromatin unterscheiden. Es ist seit langem bekannt, dass diese getrennten Domänen Unterschiede in chromosomaler Architektur und Replikation, aber auch in transkriptioneller

Aktivität aufweisen (Li, 2002). Die Untersuchung dieser Chromatinstrukturen schließt die Entstehung des „higher order"-Chromatins, die Anlagerung an die Kernmatrix, die Bindung der DNA an die Nukleosomen sowie kovalente Modifikationen der Histonseitenketten und die Methylierung der DNA mit ein (Murrell *et al.*, 2005). Diese Vorgänge bilden ein komplexes epigenetisches Netzwerk und regulieren die Aktivität der Gene (Li, 2002).

Die Regulation der Gene[1] insbesondere durch die Interaktion von Histonmodifikationen und DNA-Methylierung führt dabei zu bemerkenswerten Sonderfällen wie der Inaktivierung des zweiten X-Chromosoms und der genomischen Prägung von Genen (siehe Kapitel 1.2). Dabei führt die Methylierung der DNA zur Bindung von methylierungssensitiven Proteinen, die wiederum mit histonmodifizierenden Proteinen assoziiert sind (Fuks *et al.*, 2003; Hashimshony *et al.*, 2003).

Kovalente Modifikationen an den N-terminalen Seitenketten der Kernhistone wie Acetylierung, Phosphorylierung, Methylierung und/oder Ubiquitinierung ergeben sehr komplexe Muster (siehe Abb. 1), die die Transkription aktivieren oder inhibieren (Li, 2002).

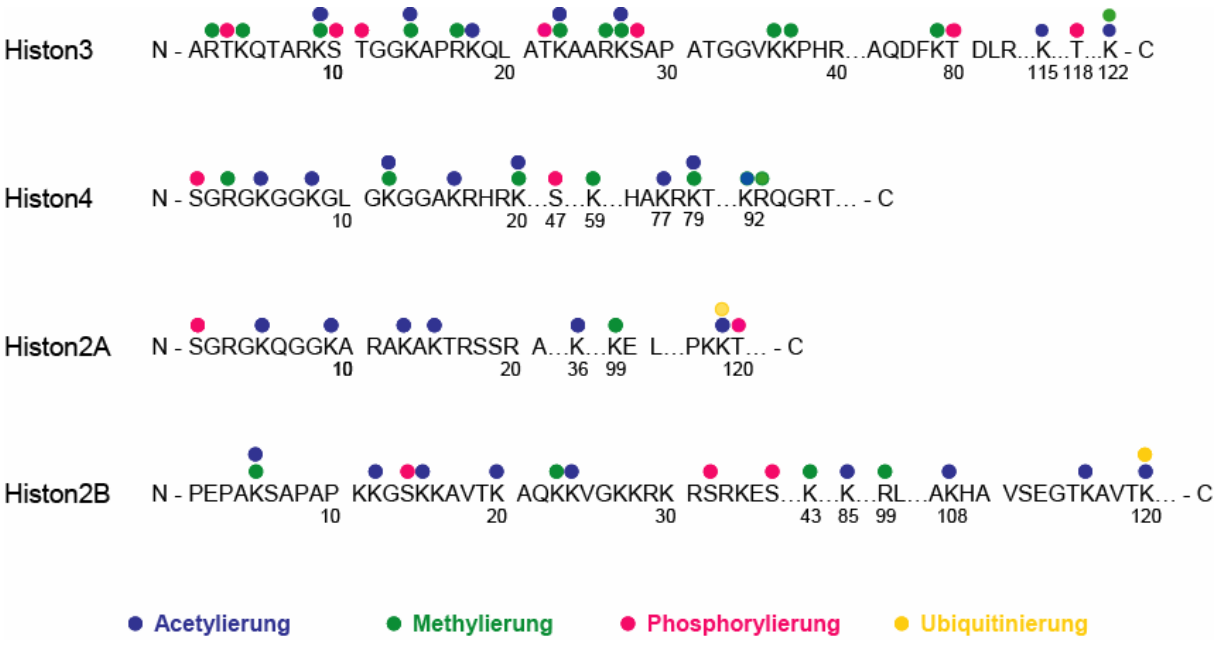

Abbildung 1. Modifikationen an den Seitenketten der vier Kernhistone. Verändert nach www.histone.com.

So ist bekannt, dass die Acetylierung der Histone die transkriptionelle Aktivität des Chromatins stets erhöht (Strahl und Allis, 2000), während eine Methylierung am vierten Lysin vom Histon3 (H3) mit erhöhter transkriptioneller Aktivität jedoch am Lysin9 vom H3 mit verringerter transkriptioneller Aktivität assoziiert ist (Jenuwein und Allis, 2001; Bird, 2002). Darüber hinaus sind Mono-, Di- oder Trimethylierung von Lysin zu beobachten, was die

[1] Mit Regulation der Gene wird hier und im Folgenden ausschließlich die Regulation der Transkription gemeint.

Komplexität dieser epigenetischen Markierung weiter erhöht (Strahl und Allis, 2000). Die Phosphorylierung des H3 scheint über chromosomale Umstrukturierungen, die zur Dekondensierung des Chromatins führen, eine direkte Rolle bei der Regulierung der Transkription zu spielen und agiert zusammen mit der Acetylierung am Lysin16 vom Histon4 (Strahl und Allis, 2000). Das komplexe Muster, welches durch die vielfältigen posttranslationellen Modifikationen der Histonseitenketten zu beobachten ist, führte zur Histon-Code-Hypothese, die besagt, dass die Modifikationsmuster von interagierenden Proteinen regelrecht gelesen werden und sich daraus unter anderem die transkriptionelle Aktivität von chromosomalen Bereichen ableitet (Strahl und Allis, 2000).

Wie bereits erwähnt, stehen Histonmodifikationen und DNA-Methylierung in unmittelbarem Zusammenhang. Bei Prokaryoten und Eukaryoten ist die enzymatisch katalysierte Methylierung am fünften Kohlenstoffatom des Cytosins eine häufige Modifikation. Bei Säugetieren kommt die Methylierung des Cytosins fast ausschließlich an symmetrischen CpG-Dinukleotiden vor, wobei die Cytosine beider DNA-Stränge methyliert sind (Paulsen und Ferguson-Smith, 2001). Im Allgemeinen ist die DNA-Methylierung mit einer Repression der Transkription korreliert, die beispielsweise in der Promotorregion durch eine Blockierung der Bindung von Transkriptionsfaktoren erreicht wird (Li, 2002). Die Methylierung der DNA erfolgt während oder nach der Replikation durch DNA-Methyltransferasen und wird somit stabil durch die mitotischen Zellteilungen vererbt (Paulsen und Ferguson-Smith, 2001). Die Verteilung der Methylgruppen im Genom ist dabei nicht zufällig, sondern folgt stabilen Mustern, die vererbbar, zell- und ortsspezifisch sind (Doerfler, 2005). Darüber hinaus gibt es genomische Loci mit allelspezifischen Methylierungsmustern, den so genannten Imprints. Bei den Imprints handelt es sich um essentielle epigenetische Markierungen, die für die Etablierung der genomischen Prägung von Genen verantwortlich sind (Li, 2002).

1.2. Genomische Prägung

Verschiedene Untersuchungen an Mäusen in den 1980er Jahren belegen, dass bei einem Individuum die mütterlichen und väterlichen Autosomen nicht äquivalent sind (McGrath und Solter, 1984; Surani *et al.*, 1984). Befruchtete Eizellen der Maus, deren Chromosomen vollständig maternalen Ursprungs sind, bilden keine Plazenta, während sich der Embryo zum Teratom entwickelt (Surani *et al.*, 1984; Surani *et al.*, 1986). Im umgekehrten Fall entwickeln befruchtete Eizellen mit Chromosomen rein paternalen Ursprungs keinen Embroyblast, während sich der Trophoblast zur Blasenmole umwandelt (McGrath und Solter, 1984). Der Grund dafür ist die genomische Prägung von Genen, die auch als Imprinting bezeichnet wird, einem Sonderfall epigenetischer Regulation. Durch

allelspezifische Methylierung in Abhängigkeit vom elterlichen Ursprung kommt es zur transkriptionellen Inaktivierung eines Allels und damit zur ausschließlichen oder präferentiellen Expression des anderen Allels (Reik und Walter, 2001). In der Literatur wird im Zusammenhang mit der genomischen Prägung von Genen zumeist der Transkriptionsstatus mit paternal bzw. maternal exprimiert angegeben. Dabei ist ein paternal exprimiertes Gen maternal geprägt und damit das maternale Allel transkriptionell inaktiv, und ein maternal exprimiertes Gen ist paternal geprägt und damit das paternale Allel transkriptionell inaktiv. Die Bezeichnung maternale Prägung anstatt paternaler Expression würde jedoch das Phänomen der genomischen Prägung präziser beschreiben (Morison *et al.*, 2005).

Das paternal exprimierte Gen *Igf2* wurde als erstes geprägtes Gen identifiziert (DeChiara *et al.*, 1991). Mittlerweile sind 77 geprägte Gene bei der Maus bekannt. Viele dieser Gene sind auch beim Menschen geprägt. Dazu zählen neben proteinkodierenden Genen (ca. 2/3) auch zahlreiche nicht-kodierende Gene (ca. 1/3), wie antisense-RNAs, snoRNA, microRNAs, Pseudogene, aber auch RNAs mit unbekannter Funktion (Morison *et al.*, 2005). Eine aktuelle Übersicht ist auf der Webseite der Mammalian Genetics Unit (siehe http://www.mgu.har.mrc.ac.uk/research/imprinting/) abrufbar. Neben den Imprints weisen geprägte Gene weitere bemerkenswerte epigenetische und genetische Eigenschaften auf, die im Folgenden erklärt werden.

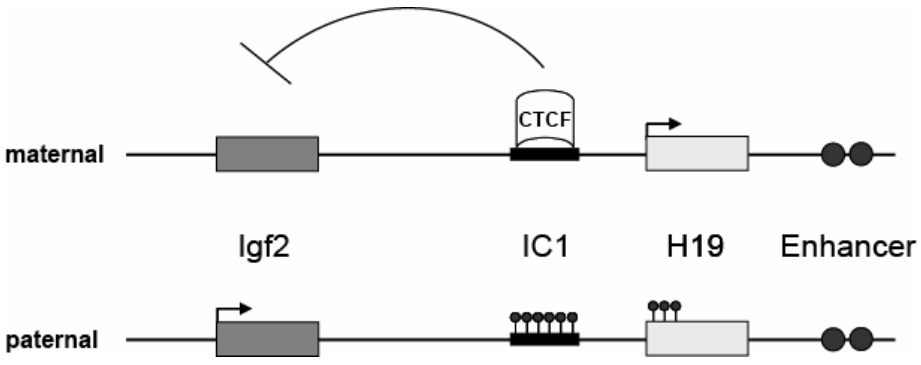

Abbildung 2. Die genomische Prägung von *Igf2* und *H19* auf dem distalen Maus-Chromosom 7 (MMU7dist). Die differentiell methylierte Region stromaufwärts von *H19*, hier Imprinting Center 1 (IC1) genannt, ist elternspezifisch methyliert und beinhaltet Bindungsstellen für den Insulator CTCF. Auf dem unmethylierten maternalen Allel bindet CTCF und stört die Wechselwirkungen zwischen dem *Igf2*-Promotor und den stromabwärts von *H19* liegenden Enhancer-Elementen. Das führt zur Inaktivierung des maternalen *Igf2*-Allels während *H19* transkribiert wird. Auf dem paternalen Allel sind IC1 sowie der *H19*-Promotor methyliert. CTCF ist methylsensitiv, kann nicht binden und wirkt somit nicht als Insulator. Folglich ist *Igf2* aktiv und *H19* inaktiv. Die aktive Transkription eines Allels wird durch einen Pfeil dargestellt. Verändert nach (Verona *et al.*, 2003).

Von den bekannten geprägten Genen im Mausgenom sind 68 Gene (entspricht 88%) in Clustern organisiert. Die physikalische Kopplung deutet daraufhin, dass diese Gene koordiniert reguliert werden. Eine intensiv untersuchte Region liegt auf dem distalen Maus-

Chromosom 7 (MMU7dist). Die orthologe Region beim Menschen liegt auf Chromosom 11p15.5 (HSA11p15.5). Dieser ca. 1 Mb große Locus besteht aus zwei unabhängig regulierten Subdomänen, der *H19*- und der *Lit1*-Subdomäne. In der Abbildung 2 ist der Bereich mit *H19* und *Igf2* schematisch dargestellt und erklärt. Dieses vereinfachte Schema zeigt die koordinierte Regulation und reziproke Expression von *H19*, einem nicht-kodierenden putativen Wachstumssuppressor und *Igf2*, einem Wachstumsfaktor. Zwischen *Igf2* und *H19* liegt eine differentiell methylierte Region (DMR), die für die Regulation von *H19* sowie *Igf2* eine essentielle Rolle spielt. Üblicherweise wird eine DMR, die für die Regulation mehrerer geprägter Gene verantwortlich ist, als *imprinting center* (IC) definiert. Die DMR zwischen *Igf2* und *H19* wird als *imprinting center 1* (IC1) bezeichnet (Reik und Walter, 2001; Verona *et al.*, 2003). Nicht abgebildet sind drei weitere differentiell methylierte Regionen innerhalb von *Igf2* (DMR0, DMR1 und DMR2). In einer Untersuchung wurden Interaktionen zwischen IC1 und DMR1 auf dem maternalen Allel sowie zwischen IC1 und DMR2 auf dem paternalen Allel nachgewiesen, die zu unterschiedlichen elternspezifischen Chromatin-Schleifen und damit zur reziproken monoallelischen Expression von *Igf2* und *H19* führen (Murrell *et al.*, 2004).

Neben der häufigen physikalischen Kopplung gibt es zwar keine Ähnlichkeit in der Proteinsequenz, jedoch oft funktionelle Gemeinsamkeiten zwischen geprägten Genen. Zahlreiche geprägte Gene beeinflussen prä- und postnatales Wachstum und Entwicklung (Reik und Walter, 2001) sowie postnatales Verhalten von Mutter und Neugeborenen (Constancia *et al.*, 2004) (siehe Kapitel 1.2.2). Darüber hinaus weist auch die flankierende DNA-Sequenz von geprägten Genen Besonderheiten auf. So sind 88% der bekannten geprägten Gene mit CpG-Inseln und diese wiederum mit repetitiven Elementen gekoppelt. (Reik und Walter, 2001; Allen *et al.*, 2003).

Zu den epigenetisch bedingten Merkmalen geprägter Gene gehören die asynchrone Replikation während der S-Phase sowie unterschiedliche meiotische Rekombinationsfrequenzen (Reik und Walter, 2001). Die asynchrone Replikation wurde bereits kurz nach Entdeckung der ersten geprägten Gene nachgewiesen. Oft wird dabei das paternale Allel zuerst repliziert. Die asynchrone Replikation deutet darauf hin, dass geprägte Gene in größere DNA-Domänen eingebettet sind, die scheinbar eine elternspezifische Identität besitzen. Die genauen molekularen Ursachen sind jedoch nicht bekannt (Kitsberg *et al.*, 1993; Knoll *et al.*, 1994; Reik und Walter, 2001). Die beobachteten unterschiedlichen meiotischen Rekombinationsfrequenzen von Regionen mit geprägten Genen sind ursächlich ebenfalls nicht erforscht (Paldi *et al.*, 1995; Reik und Walter, 2001).

Inwieweit die oben genannten Eigenschaften mit DNA-Methylierung und Chromatinstrukturen gekoppelt und somit für die Prägung verantwortlich sind, ist bisher nicht geklärt. Es ist jedoch sehr wahrscheinlich, dass diese genetischen und epigenetischen

Modifikationen bzw. Merkmale eine wichtige Rolle spielen. Zumindest für die Imprints steht fest, dass sie unentbehrlich sind bei der Entwicklung und Erhaltung der Prägung (Reik und Walter, 2001).

1.2.1. DNA-Methylierung und genomische Prägung

Wie jeder Organismus und jede Zelle durchlaufen auch die für die Prägung essentiellen Imprints einen Zyklus (siehe Abb. 3). In der Keimbahn des Embryos der Maus werden sie zunächst wahrscheinlich durch aktive Demethylierung bis zum Tag 12 entfernt (Tada *et al.*, 1997), so dass geprägte Gene biallelisch exprimiert werden (Szabo *et al.*, 2002).

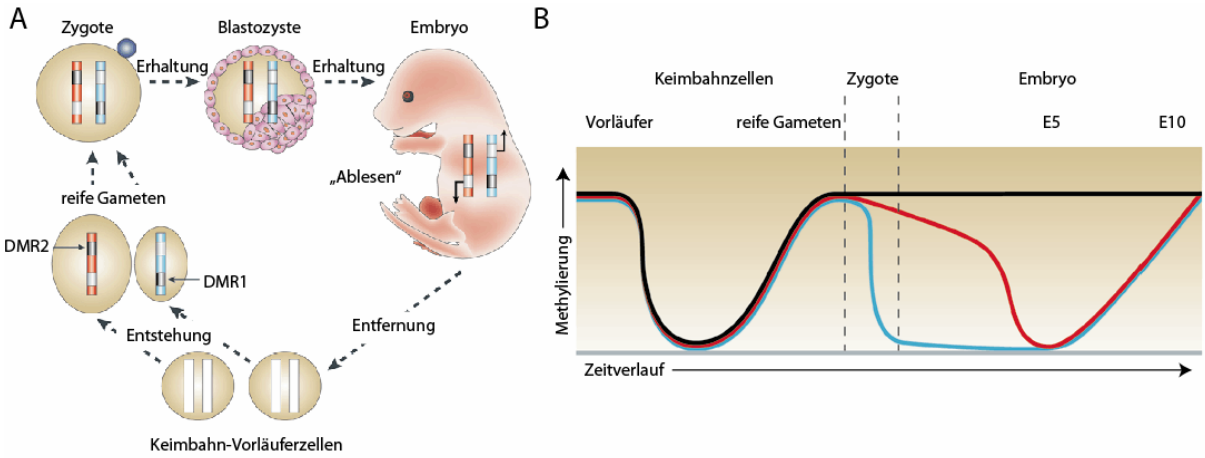

Abbildung 3. Der Lebenszyklus der Imprints. **(A)** - Die für die Prägung essentiellen Methylierungsmuster unterliegen einem Kreislauf aus Entfernung, Etablierung und Erhaltung. Die zwei dargestellten differentiell methylierten Regionen (DMR1 und DMR2) dienen als Beispiele für geschlechtsspezifische Methylierung. Das methylierte Allel ist in schwarz und das nicht methylierte Allel in hellgrau dargestellt. Das paternale Chromosom ist blau und das maternale Chromosom rot markiert. Die Transkription eines Gens wird durch einen Pfeil angezeigt. **(B)** - Hier ist der Methylierungsverlauf des methylierten (schwarz) und des nicht methylierten Allels geprägter Gene (hellgrau) innerhalb der Keimbahn und der Entwicklung des Embryos dargestellt. Die Methylierung der maternalen (rot) und paternalen Allele (blau) von nicht geprägten Genen ist ebenfalls eingezeichnet. E5 bzw. E10 stehen für *embryonic day* 5 bzw. 10. Verändert nach (Reik und Walter, 2001).

Anschließend werden die Imprints durch *de novo*-Methylierung wieder an den DNA-Strang gekoppelt. Die reifen Gameten weisen letztlich das geschlechtspezifische Imprint-Muster auf. Welche Enzyme in der Keimbahn dafür verantwortlich sind, ist nicht bekannt (Reik und Walter, 2001). Die DNA-Methyltransferase 1 (Dnmt1) und deren keimbahnspezifische Isoformen, aber auch Dnmt3a und Dnmt3b werden als Kandidaten genannt (Mertineit *et al.*, 1998; La Salle und Trasler, 2006). Untersuchungen an Dnmt1-defizienten Mäusen ergaben biallelische Expression geprägter Gene und belegen die Bedeutung der Methylierung für die

Etablierung der Prägung (Li *et al.*, 1993; Caspary *et al.*, 1998). Wie die Methyltransferasen jedoch ihr spezifisches Ziel - die DMR - erkennen, ist nach wie vor ungeklärt (Reik und Walter, 2001). Für die Erhaltung der DNA-Methylierung insbesondere durch die Methylierung des neu synthetisierten DNA-Stranges während der Replikation scheint Dnmt1 verantwortlich zu sein (Bird, 2002). Allerdings kommt es nach der Befruchtung zu einer genomweiten Demethylierung und einer darauf folgenden Remethylierung (siehe Abb. 3) (Morgan *et al.*, 2005). Die DMRs von geprägten Genen entgehen diesen Prozessen und bleiben unverändert. Die dafür verantwortlichen Schutzmechanismen sind noch nicht bekannt. Wahrscheinlich spielen aber Bestandteile des Chromatins eine wichtige Rolle (Reik und Walter, 2001).

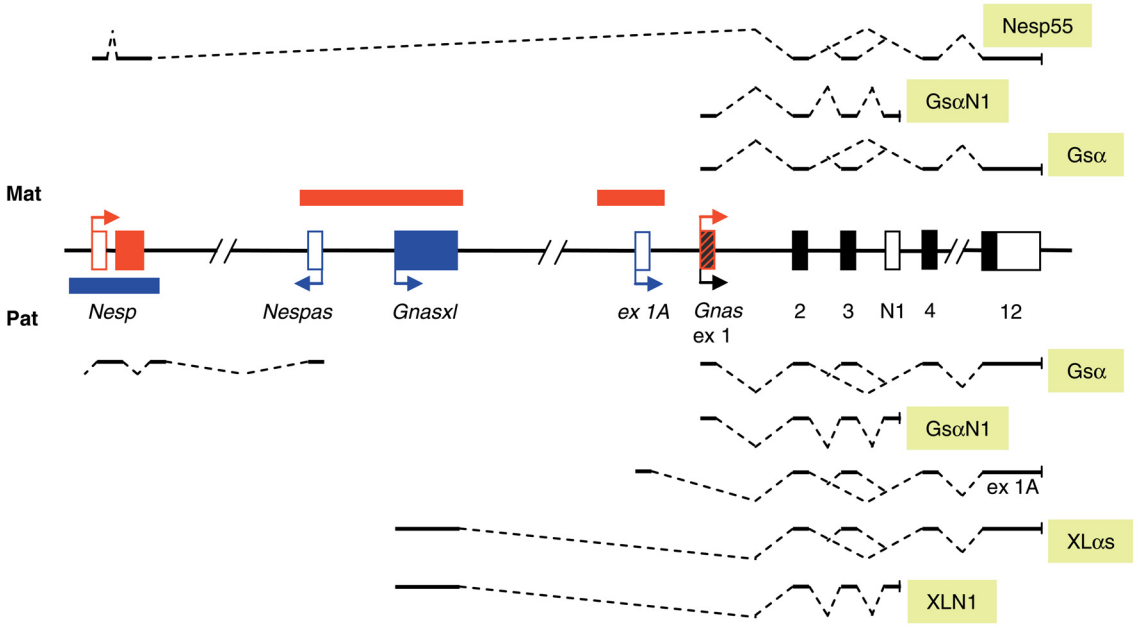

Abbildung 4. Schematische Darstellung des *Gnas*-Locus auf dem distalen Maus-Chromosom 2 (MMU2dist). Die allelspezifische Methylierung ist durch rote (für maternales Allel) und blaue (für paternales Allel) Balken markiert. Kodierende Exons sind durch ausgefüllte Rechtecke und nicht-kodierende Exons durch leere Rechtecke skizziert. Transkriptionsstart und -richtung sind durch einen Pfeil angegeben. Alle Transkripte einschließlich Spleißvarianten sind über bzw. unter dem chromosomalen Abschnitt dargestellt. Die gewebespezifische maternale Expression von *Gnas* ist durch ein schraffiertes Rechteck hervorgehoben. Die kodierenden Transkripte sind hellgrün unterlegt (von (Plagge und Kelsey, 2006)).

Die allelspezifische Methylierung ist zum Teil auf einzelne Gewebe beschränkt und führt dann zur gewebespezifischen monoallelischen Expression geprägter Gene (Li, 2002). Wie komplex das Zusammenspiel von DNA-Methylierung und genomischer Prägung sein kann, zeigt der *Gnas*-Locus auf MMU2dist (siehe Abb. 4). Neben *Gnas* beherbergt dieser Bereich die paternal exprimierten Gene *Gnasxl*, *Exon1A* und *Nespas*, sowie das maternal exprimierte Gen *Nesp*. *Gnas* wird in den meisten Geweben biallelisch, in einigen Geweben aber präferentiell maternal exprimiert (Peters *et al.*, 1999). Inzwischen konnten drei DMRs

identifiziert werden, die eine unterschiedliche Rolle spielen. Während die DMR bei *Exon1A* für die gewebespezifische Prägung von *Gnas* verantwortlich ist, wirken die DMRs bei *Nesp* und bei *Nespas* entgegengesetzt und begünstigen bzw. verhindern die Methylierung der *Exon1A*-DMR (Williamson *et al.*, 2006).

1.2.2. Funktionelle Gemeinsamkeiten geprägter Gene

Durch die oben beschriebenen Studien an uniparentalen Mäusen wurde bereits vor Entdeckung der geprägten Gene deren Bedeutung für eine normale pränatale Entwicklung erkannt. Aus den beobachteten Phänotypen leitete sich die Hypothese ab, dass paternal exprimierte Gene bei der Plazentaentwicklung eine zentrale Rolle spielen, während maternal exprimierte Gene bei der Embryonalentwicklung eine wichtige Funktion erfüllen (Barton *et al.*, 1984). Von einigen der 77 bekannten geprägten Gene gibt es bereits Studien mit Knockout-Mäusen, deren Ergebnisse diese Hypothese bestätigen und erweitern. So führt die Inaktivierung der paternal exprimierten Gene *Gnasxl*, *Peg1*, *Peg3*, *Igf2*, *Rasgrf1* oder *Dlk1* zu einem verzögerten Wachstum des Embryos, während die Inaktivierung der maternal exprimierten Gene *Gnas*, *Grb10* oder *Igf2r* zu embryonalem Überwuchs führt (siehe Tab. 1).

Bereits intensiv erforscht ist die antagonistische Funktion der Gene *Igf2* und *Igf2r*. *Igf2* und *Igf2r* kodieren für den Wachstumsfaktor Insulin-like Growth Factor Type 2 (IGF2) bzw. für den Insulin-like Growth Factor Type 2-Rezeptor (IGF2R). Während IGF2 das Wachstum von Embryo und Plazenta fördert (DeChiara *et al.*, 1991), inhibiert IGF2R die Wirkung von IGF2 und damit die Wachstumsförderung durch dessen Bindung mit nachfolgender lysosomaler Degradierung (Lau *et al.*, 1994). Auch die Gene einiger Transportproteine, die für den Austausch von Nährstoffen und Stoffwechselprodukten zwischen Embryo und Plazenta sorgen, sind geprägt. So ist bei einer erhöhten Expression des paternal exprimierten Gens *Slc38a4* ein gesteigerter Transport von Glucose und Aminosäuren von der Plazenta zum Fötus zu beobachten, was ein erhöhtes embryonales Wachstum zur Folge hat (Constancia *et al.*, 2005), während bei erhöhter Expression der maternal exprimierten Gene *Slc22a1l*, *Slc22a2* und *Slc22a3* ein gegenteiliger Effekt vermutet wird (Fitzpatrick *et al.*, 2002; Constancia *et al.*, 2004). Aus diesen Beispielen kann geschlussfolgert werden, dass paternal exprimierte Gene das Wachstum fördern, während maternal exprimierte Gene das Wachstum vermindern bzw. die Wirkung paternal exprimierter Gene inhibieren.

Neben der Bedeutung geprägter Gene für die pränatale Entwicklung sind postnatale Anpassungsprozesse, wie die Ausbildung des Glucose- und Temperaturgleichgewichts oder die Produktion der Muttermilch, das Saugverhalten des Neugeborenen sowie die emotionalen Interaktionen zwischen Mutter und Kind ein weiteres wichtiges gemeinsames

Funktionsfeld (Constancia *et al.*, 2004). Beispielhaft sei hierzu die Wirkung des gewebespezifisch maternal exprimierten Gens *Gnas* und des paternal exprimierten Gens *Gnasxl* erwähnt. *Gnasxl*-defiziente Mäuse zeigen schwaches postnatales Wachstum und geringe Überlebenschancen. Der Vergleich mit *Gnas*-defizienten Mäusen, die bei einem verminderten Metabolismus Adipositas und Insulinresistenz entwickeln, lässt die Vermutung zu, dass beide Genprodukte über eine rezeptorvermittelte Stimulation der Adenylat-Cyclase ebenfalls antagonistische Stoffwechselwege induzieren (Plagge *et al.*, 2004).

Tabelle 1. Geprägte Gene und die Phänotypen von Knockout-Mäusen.

Gensymbol	Name	exprimiert	Phänotyp bei Knockout-Mäusen	Quelle
Gnas	guanine nucleotide binding protein, alpha stimulating	maternal	bei tmiLsw-Isoform: Überwuchs, verkleinertes Herz bei tm4Lsw-Isoform: kongenitale Ödeme, 50% Sterblichkeit, Adipositas, Insulinresistenz	(Yu *et al.*, 2000) (Chen *et al.*, 2005)
Gnasxl	guanine nucleotide binding protein, alpha stimulating, extra large	paternal	postnatale Wachstumsverzögerung, abnormale Nahrungsaufnahme	(Plagge *et al.*, 2004)
Nesp	neuroendocrine secretory protein	maternal	abnormales Verhalten auf sich änderndes Umfeld	(Plagge *et al.*, 2005)
Peg 10	Paternally expressed gene 10	paternal	frühe embryonale Sterblichkeit, Plazentadefekte	(Ono *et al.*, 2006)
Asb4	Ankyrin repeat and suppressor of cytokine signaling	maternal	Defekt der Spermatogenese	(Kile *et al.*, 2001)
Mest	mesoderm specific transcript	paternal	Wachstumsverzögerung, mütterliches Verhalten	(Lefebvre *et al.*, 1998)
Peg3	paternally expressed gene 3	paternal	Wachstumsverzögerung, mütterliches Verhalten	(Li *et al.*, 1999)
Snrpn	small nuclear ribonucleoprotein polypeptide N	paternal	kein Effekt	(Yang *et al.*, 1998)
Snurf	Snrpn upstream reading frame	paternal	kein Effekt	(Yang *et al.*, 1998)
Ndn	necdin	paternal	letal oder lebensfähig je nach Mausstamm	(Tsai *et al.*, 1999)
Frat3	Frequently rearranged in advanced T-cell lymphomas	paternal	kein Effekt	(Jonkers *et al.*, 1999)
Ube3a	E6-AP ubiquitin protein ligase 3A	maternal	motorische Dysfunktion, Lerndefizit	(Jiang *et al.*, 1998)
H19	A cDNA clone isolated from a fetal hepatic library	maternal	kein Effekt	(Jones *et al.*, 1998)
Igf2	insulin-like growth factor type 2	paternal	Wachstumsverzögerung	(DeChiara *et al.*, 1991)
Ins2	insulin 2	paternal	kein Effekt	(Duvillie *et al.*, 1997)
Mash2	Mus musculus achaete-scute homologue 2	maternal	Plazentadefekt, Sterblichkeit	(Guillemot *et al.*, 1995)
Kcnq1	Potassium voltage - gated channel subfamily KQT member 1	maternal	Schwerhörigkeit, gastritische Hyperplasie	(Lee *et al.*, 2000)
Cdkn1c	cyclin-dependent kinase inhibitor 1C	maternal	reduzierte Überlebensrate, Wachstumsverzögerung	(Yan *et al.*, 1997)
Tssc3	Pleckstrin homology-like domain, Family A, member2	maternal	verstärktes oder verringertes Plazentawachstum	(Salas *et al.*, 2004)
Rasgrf1	Ras protein specific guanine nucleotide-releasing factor 1	paternal	gestörtes Langzeitgedächtnis, postnatale Wachstumsverzögerung	(Brambilla *et al.*, 1997)
Dcn	Decorin	maternal	Tumorgenese, Apoptose, unterentwickelte Haut	(Danielson *et al.*, 1997)
Grb10	growth factor receptor bound protein 10	maternal	Überwuchs	(Charalambous *et al.*, 2003)
U2af1- rs1	U2 small nuclear ribonucleoprotein auxiliary factor (U2AF), 35kDa, related sequence 1	paternal	kein Effekt	(Sunahara *et al.*, 2000)
Dlk1	delta like 1	paternal	Wachstumsverzögerung, Adipositas	(Moon *et al.*, 2002)
Gtl2	gene trap locus 2	maternal	Wachstumsverzögerung	(Schuster-Gossler *et al.*, 1996)
Slc22a3	Membrane spanning transporter protein	maternal	lebensfähig und fertil, aber gestörte Aufnahme von Neurotransmittern	(Zwart *et al.*, 2001)
Igf2r	insulin-like growth factor type 2 receptor	maternal	letal, Überwuchs	(Wang *et al.*, 1994)

Ein weiteres anschauliches Beispiel lässt sich durch die Wirkung des paternal exprimierten Gens *Peg3* beim Stillen gewinnen. So kontrolliert *Peg3* die Bereitschaft von Mutter und Säugling zum Stillen bzw. Saugen. Weibliche Mäuse ohne *Peg3*-Expression zeigen gegenüber ihrem Nachwuchs nur geringe Fürsorge und eine mangelnde Stillbereitschaft. Bei neugeborenen Mäusen ohne *Peg3*-Expression ist nur wenig Interesse am Saugen erkennbar. Sind beide Phänotypen kombiniert, sind die Auswirkungen für den Nachwuchs verheerend mit einer Sterblichkeitsrate von 94% (Curley *et al.*, 2004).

Darüber hinaus wird vermutet, dass geprägte Gene eine wichtige Rolle bei der neuronalen Entwicklung, besonders bei der Entwicklung des Gehirns, spielen (Davies *et al.*, 2005). Die Phänotypen von *Nesp*-, *Ube3A*-, *Kcnq1*- und *Slc22a3*-defizienten Mäusen weisen darauf hin, dass fehlregulierte maternal exprimierte Gene an der Entstehung von Verhaltensauffälligkeiten und abnormen neurologischen Phänotypen beteiligt sind (siehe Tab. 1). Diese Vermutung wird von der hohen Expression vieler geprägter Gene im zentralen Nervensystem (ZNS) untermauert. Da diese Gene aber auch in anderen Geweben exprimiert sind, konnte bisher die spezifische Funktion eines geprägten Gens bei der Entwicklung des ZNS nicht analysiert werden. Aus Studien mit Maus-Chimären, die weniger als 40% androgenetische bzw. parthenogenetische Zellen in sich tragen, konnte gefolgert werden, dass maternal exprimierte Gene das Wachstum des Gehirns fördern, während paternal exprimierte Gene eine reziproke Wirkung aufweisen (Allen *et al.*, 1995; Keverne, 1997; Davies *et al.*, 2005).

Beim Menschen sind aus ethischen Gründen solche gezielten genetischen Manipulationen nicht möglich. Daher gibt es zur Funktion der humanen geprägten Gene nur wenige Studien. Die gestörte Regulation geprägter Gene zieht jedoch auch beim Menschen schwerwiegende phänotypische Konsequenzen nach sich. Neben den am besten untersuchten Imprinting-Erkrankungen wie Beckwith-Wiedemann-, Prader-Willi- und Angelman-Syndrom sind Imprinting-Defekte an der Entwicklung einer Vielzahl von Krebserkrankungen beteiligt (Paulsen und Ferguson-Smith, 2001; Reik und Walter, 2001).

1.2.3. Gestörte Prägung und resultierende humane Erkrankungen

Neben den in der Tabelle 2 aufgelisteten humanen Erkrankungen, die mit einer Störung der genomischen Prägung im Zusammenhang stehen, gibt es zahlreiche Erkrankungen, bei denen die Ursache in einer Imprinting-Störung vermutet wird. Zu dieser Vermutung führten beobachtete parentale Effekte, d.h. die prädominant parentale Vererbung. Dazu zählen beispielsweise Autismus, atypisches Asthma, Psoriasis, Diabetes Typ 1 und 2, Alzheimer und Schizophrenie (Morison *et al.*, 2005). Insgesamt sind derzeit 102

humane abnorme Phänotypen mit parentalen Effekten bekannt (siehe http://igc.otago.ac.nz/home.html).

Tabelle 2. Humane Imprinting-Erkrankungen (Morison *et al.*, 2005).

Phänotyp / Tumor	OMIM-Eintrag	chromosomaler Locus (prädominante Transmission[2])
Beckwith-Wiedemann-Syndrom	130650	11p15 (maternal)
Prader-Willi-Syndrom	176270	15q11-q12 (paternal)
Angelman-Syndrom	105830	15q11-q12 (maternal)
Silver-Russell-Syndrom	180860	7p11-p13, 7q31-qter (paternal)
Transienter neonataler Diabetes mellitus	601410	6q24 (maternal)
McCune-Albright-Syndrom	174800	20q13 (maternal)
Familiäres nicht chromaffines Paragangliom	601650	11q13 (maternal)
UPD14-Syndrom	608149	14 (maternal sowie paternal)
Neuroblastom	256700	1p36 (maternal) und 2 (paternal)
Akute myeloische Leukämie	252270	7 (paternal)
Wilms Tumor (WT2)	194071	11p15 (maternal)
Rhabdomyosarkom	268210	11p15 (maternal)
Sporadisches Osteosarkom	180200	13 (maternal)

Im Folgenden werden drei Imprinting-Erkrankungen vorgestellt, die ein charakteristisches Bild über die Auswirkungen fehlregulierter geprägter Gene vermitteln. Das Beckwith-Wiedemann-Syndrom (BWS) ist ein heterogenes Überwuchssyndrom und wird durch Makroglossie, abdominale Bauchwanddefekte, faziale Abnormitäten, Hypoglykämie und einem deutlich erhöhten Risiko für embryonale Tumore charakterisiert (Robertson, 2005). Die phänotypische Ausprägung der einzelnen Merkmale ist dabei variabel. Bei 10 bis 20 % der Fälle liegt eine paternale uniparentale Disomie (UPD) von Chromosom 11p15 vor. Typische Fälle mit BWS weisen eine biallelische Expression der geprägten Gene *LIT1*, *IGF2* und/oder Mutationen im Gen *CDKN1C* auf (Paulsen und Ferguson-Smith, 2001; Weksberg *et al.*, 2005).

Das Prader-Willi-Syndrom (PWS) und das Angelman-Syndrom (AS) sind unterschiedliche neurogenetische Erkrankungen, die durch die Fehlregulation der geprägten Gene in der gleichen chromosomalen Region 15q11-q13 verursacht werden. PWS ist gekennzeichnet durch Muskelhypotonie, Hypogenitalismus, Adipositas, Minderwuchs sowie leichte mentale Retardierung. In ca. 25% der Fälle liegt eine maternale UPD15 vor. In den meisten Fällen (70%) beruht die Erkrankung auf einer Deletion der paternalen Region von 15q11-q13. AS-Patienten weisen Merkmale wie Ataxie, Hyperaktivität, fehlende Sprachentwicklung, schwere mentale Retardierung sowie unmotivierte Lachanfälle auf. In bis zu 3% der Fälle liegt eine paternale UPD15 vor. 70% der Fälle beruhen auf einer Deletion der maternalen Region von 15q11-q13 auf (Paulsen und Ferguson-Smith, 2001; Horsthemke und Buiting, 2006).

Trotz der ursächlichen Komplexität dieser Imprinting-Erkrankungen wie auch der in Tabelle 2 aufgeführten Tumorarten sind die oben beschriebenen funktionellen Gemeinsamkeiten erkennbar. So lässt sich beispielsweise der Überwuchs beim Beckwith-

[2] Die Angaben zu prädominanter parentaler Transmission wurde der Webseite: http://igc.otago.ac.nz/home.html entnommen bzw. aus den Erläuterungen geschlussfolgert.

Wiedemann-Syndrom mit der biallelischen Expression von *IGF2* erklären, der ein Verlust der maternalen Prägung von *IGF2* zu Grunde liegt (Sparago *et al.*, 2004; Robertson, 2005).

Bedingt durch die funktionelle Haploidie der geprägten Gene ist der postulierte Vorteil diploider Organismen in Bezug auf die eingeschränkte Wirkung rezessiver Mutationen aufgehoben (Wilkins und Haig, 2003). So kann eine heterozygote Mutation in einem geprägten Gen leicht zu den oben beschriebenen Imprinting-Erkrankungen führen, da beispielsweise das kompensierende Wildtyp-Allel nicht exprimiert wird. Dieser Nachteil und die zum Teil antagonistische Funktion der geprägten Gene führten zu Fragen nach dem evolutionären Sinn der genomischen Prägung, aber auch dazu, wie ein derart komplexes epigenetisches System unter evolutionsbiologischem Aspekt der natürlichen Selektion widerstehen konnte und welche Vorteile die genomische Prägung letztlich mit sich bringt.

1.3. Hypothesen zur Evolution der genomischen Prägung

Seit Entdeckung der genomischen Prägung wurden verschiedene Hypothesen und Theorien über die Entstehung und den evolutionären Nutzen aufgestellt. Diese Hypothesen und Theorien sollen zum einen die Diversität der geprägten Gene erklären und zum anderen, warum die meisten Gene nicht geprägt sind. Unter anderem wird vermutet, dass eine Aufgabe der Prägung die Verhinderung der Parthenogenese sein könnte (Wilkins und Haig, 2003). Eine andere Hypothese besagt, dass die Prägung als Nebenprodukt eines epigenetisch basierten Verteidigungsmechanismus gegen retrovirale Elemente entstanden sein könnte (Barlow, 1993). In einer weiteren Hypothese wird ein evolutionärer Vorteil durch eine erhöhte Anpassungsfähigkeit an sich verändernde Umweltbedingungen durch die Entwicklung der genomischen Prägung diskutiert (Wilkins und Haig, 2003). Insgesamt kann mit diesen Hypothesen und Theorien nur bei einer kleinen Anzahl geprägter Gene erklärt werden, warum deren Prägung entstanden sein könnte.

1.3.1. Genetische Konflikt-Theorie

Die allgemein favorisierte Theorie für die Evolution der geprägten Gene basiert auf einem elterlichen Konflikt innerhalb eines Individuums. Die Konflikt-Theorie ist mit dem evolutionsbiologischen Prinzip der natürlichen Selektion verknüpft und stellt die genomische Prägung in einen Zusammenhang mit dem Leistungsvermögen bzw. der „Fitness“ eines Lebewesens. Der Begriff Fitness dient dabei als Maß für die Überlebens- und Fortpflanzungschancen eines Individuums oder, auf anderer Ebene, eines Allels ((Strachan

und Read, 2005) Seite 370). Für eine trächtige Maus und ihren ungeborenen Nachwuchs lassen sich nun zweierlei Behauptungen aufstellen. Erstens: je mehr Nährstoffe der Embryo von der Mutter akquirieren kann, umso besser ist dieser bei der Geburt entwickelt und umso besser sind seine Überlebens- und Fortpflanzungschancen, d.h. umso größer ist seine Fitness. Zweitens: je größer sich der Nährstoffbedarf bei einer Schwangerschaft entwickelt, umso mehr werden die mütterlichen Ressourcen ausgebeutet, was wiederum negative Auswirkungen auf zukünftige Schwangerschaften und damit eine verminderte Fitness der Mutter zur Folge hat (Moore und Haig, 1991).

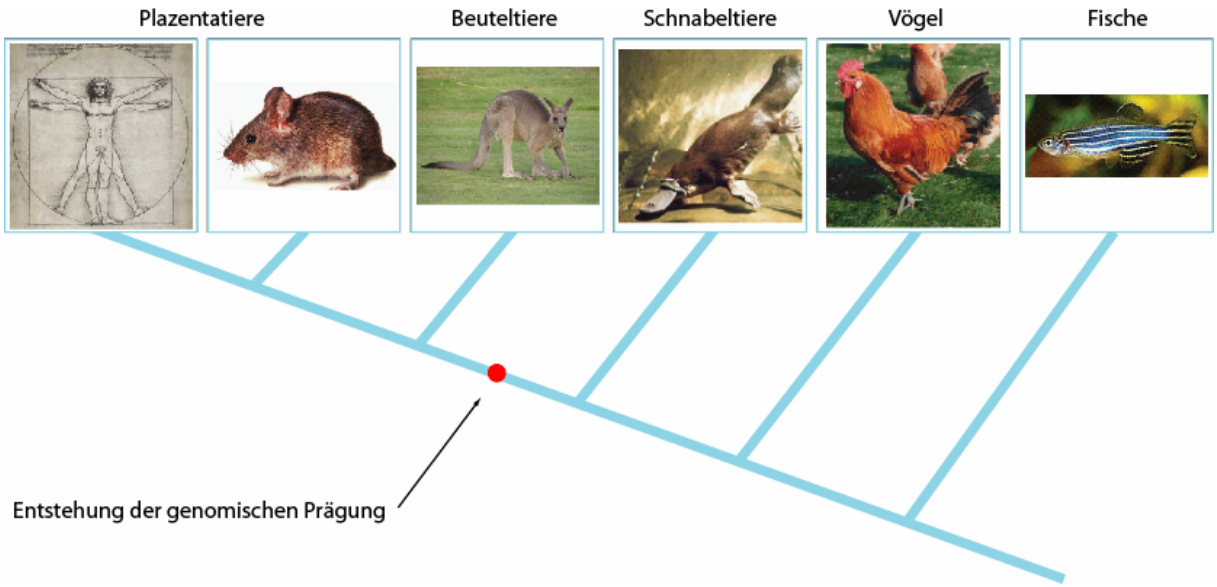

Abbildung 5. Der vermutliche Entstehungszeitpunkt der genomischen Prägung während der Evolution der Arten. Die phylogenetischen Distanzen sind modellhaft und nicht maßstabsgerecht. Verändert nach (Walter und Paulsen, 2005).

Bei polygamen Arten versucht das Weibchen ihre Ressourcen gleichmäßig auf alle Nachkommen, die von verschiedenen Vätern abstammen können, zu verteilen, während es im Interesse des Männchens liegt, dass die mütterlichen Ressourcen für seine Nachkommen maximal genutzt werden (Walter und Paulsen, 2005). Dadurch entsteht ein Konflikt zwischen Genen maternalen und paternalen Ursprungs (Moore und Haig, 1991). In der Weiterentwicklung der Konflikt-Theorie wird behauptet, dass die Evolution der genomischen Prägung in lebendgebärenden Arten einen Selektionsvorteil mit sich brachte und sich dadurch die Fitness erhöhte (Wilkins und Haig, 2003; Haig, 2004).

Die Konflikt-Theorie setzt voraus, dass zwischen Mutter und Embryo ein direkter Austausch von Nährstoffen stattfindet, wie er durch die Plazenta verwirklicht und durch den Embryo aktiv beeinflusst wird. Daher würde man nach dieser Theorie die genomische Prägung nicht bei Vögeln, Reptilien oder Fischen vermuten. Nach den bisherigen

Erkenntnissen ist die Prägung von Genen auf Plazenta- und Beuteltiere beschränkt (siehe Abb. 5) und entwickelte sich vermutlich zuerst in einem gemeinsamen Vorfahr (Walter und Paulsen, 2005). Mit der Konflikt-Theorie kann man demnach erklären, warum sich die genomische Prägung nur bei Säugetieren entwickelt hat. Darüber hinaus erklärt sie die antagonistische Wirkung geprägter Genpaare wie *Igf2/Igf2r* und *Gnasxl*/*Gnas*. Mit Hilfe der Konflikt-Theorie lassen sich auch die Beobachtungen bei androgenetischen Mausembryonen (Blasenmole und kein Embryoblast) erklären. Denn je größer, d.h. je besser die Plazenta entwickelt ist, umso mehr Nährstoffe können dem Embryo zur Verfügung gestellt werden und umso besser kann der Embryo wachsen (Constancia *et al.*, 2004).

Mit der Konflikt-Theorie lassen sich auch die bereits oben erwähnten postnatalen Effekte erklären, die durch geprägte Gene wie *Peg3*, *Gnas* oder *Gnasxl* beeinflusst werden. Denn bei vielen geprägten Genen bleibt die monoallelische Expression zeitlebens erhalten. Auch nach der Geburt ist das Neugeborene auf die weitere Nahrungsversorgung und Pflege durch die Mutter angewiesen. Aus den beobachteten postnatalen Effekten kann geschlussfolgert werden, dass paternal exprimierte Gene wie *Peg3* die mütterlichen Investitionen erhöhen, während maternal exprimierte Gene wie *Gnas* diese verringern.

Mit Hilfe der Konflikt-Theorie lässt sich die paternale oder maternale Prägung vieler Gene erklären, aber auch, weshalb die meisten Gene nicht geprägt sind. Darüber hinaus erklärt diese Theorie postnatale Effekte und warum diese epigenetische Sonderform nur auf Plazenta- und Beuteltiere beschränkt ist.

Ein kritischer Punkt der Konflikt-Theorie und anderer Theorien und Hypothesen ist, dass sich nicht alle bisher identifizierten geprägten Gene erklären lassen. So kann beispielsweise mit keiner Theorie die Prägung der Gene *Snrpn* oder *Kcnq1* erklärt werden (Morison *et al.*, 2005). Weitere Untersuchungen zur Identifizierung neuer geprägter Gene wie auch ein besseres Verständnis über deren Funktion und Regulation sind notwendig, um die Frage nach Entstehung und Nutzen geprägter Gene präziser beantworten zu können.

1.4. Identifizierung neuer geprägter Gene

Die Anzahl der geprägten Gene, ist bisher nicht bekannt. In zwei bedeutenden Untersuchungen aus den 1990er Jahren wird eine Gesamtzahl von 100 bzw. 100 bis 200 geprägten Genen vermutet. Beide Annahmen basieren auf genomweit beobachteten parentalen Effekten (Hayashizaki *et al.*, 1994; Barlow, 1995). In einem aktuellen Übersichtsartikel wird die Vermutung geäußert, dass sich die Anzahl der geprägten Gene nur noch wenig erhöhen wird (Morison *et al.*, 2005). Darüber hinaus lassen sich in der Literatur sehr verschiedene Vorhersagen mit bis zu 2101 neuen Kandidatengenen finden (Nikaido *et*

al., 2003). Die zwei grundlegenden Herangehensweisen zur Identifizierung und Untersuchung geprägter Gene werden im Folgenden vorgestellt.

1.4.1. Experimentelle Ansätze

Die meisten experimentellen Untersuchungen zu geprägten Genen wurden an Mäusen durchgeführt. So wurden auch die ersten Erkenntnisse zu den funktionellen Unterschieden der parentalen Autosomen von Mausexperimenten abgeleitet (McGrath und Solter, 1984; Surani *et al.*, 1984). Mausstämme mit verschiedenen chromosomalen Translokationen wurden dabei benutzt, um Nachkommen zu erzeugen, die eine uniparentale Disomie von ganzen Chromosomen oder eine uniparentale Duplikation von spezifischen chromosomalen Regionen aufweisen. So konnten auf acht Chromosomen mindestens elf Bereiche identifiziert werden, die bei uniparentalem Ursprung zu abnormen Phänotypen führen (Oakey und Beechey, 2002). Diese Loci sind zum Teil intensiv erforscht. Es wurden aber auch mehrere geprägte Gene (*Zac1, Dcn, Htr2a, Peg13* und *Slc38a4*) außerhalb dieser Regionen identifiziert. Diese Gene haben vermutlich keine schwerwiegenden Auswirkungen und lassen demnach bei einer Fehlregulation keine phänotypischen Unterschiede erkennen.

Die experimentellen Ansätze zur Identifizierung geprägter Gene basieren meistens auf differentieller allelischer Expression oder Methylierung. Die subtraktive Hybridisierung ist eine weit verbreitete Methode um mRNA-Populationen zu vergleichen. Für die Untersuchung von differentieller allelischer Expression werden uniparentale Embryonen mit normalen Mausembryonen oder parthenogenetische mit androgenetischen Embryonen verglichen. Zunächst wird eine zu untersuchende RNA-Präparation in einzelsträngige cDNA umgeschrieben und anschließend diese mit einem Überschuss von poly(A)-RNA einer Vergleichs-RNA vereinigt und möglichst erschöpfend hybridisiert. Die Hybriden werden anschließend abgetrennt. Dadurch bleiben theoretisch nur solche einzelsträngige cDNA-Moleküle übrig, bei denen in der zweiten RNA-Präparation keine korrespondierenden Transkripte existieren. Die übrig gebliebenen cDNA-Moleküle werden anschließend durch Klonierung und Sequenzierung analysiert ((Lottspeich und Zorbas, 1998), Seite 873). Durch subtraktive cDNA-Hybridisierung von normalen und parthenogenetischen Mausembryonen wurde das paternal exprimierte Gen *Peg3* identifiziert (Kuroiwa *et al.*, 1996). Ein Vorteil dieser Methode ist die genomweite Analyse. Da uniparentale Embryonen jedoch in frühen Stadien der Entwicklung sterben, sind Gene mit einer gewebespezifischen monoallelischen Expression von der Untersuchung ausgeschlossen (Oakey und Beechey, 2002). Die weiteren Nachteile der Methode liegen in dem hohen Arbeits- und Zeitaufwand. Außerdem können pro Experiment nur zwei Zellpopulationen untersucht werden ((Lottspeich und Zorbas, 1998), Seite 874).

Bei der Untersuchung mittels *methylation-sensitive representational difference analysis* (Kelsey *et al.*, 1999) werden allelische Methylierungsunterschiede genomweit untersucht. Zunächst wird die DNA von uniparentalen sowie von normalen Mausembryonen mit dem methylierungssensitiven Restriktionsenzym *Hpa*II verdaut. Dadurch entstehen zwei unterschiedliche Populationen von *Hpa*II-Fragmenten. In der anschließenden PCR des vollständigen Genoms werden diese Fragmente amplifiziert und die spezifisch unmethylierten Fragmente durch eine subtraktive Hybridisierung identifiziert. Auf diesem Wege konnten drei neue DMRs und daraus folgend auch die Gene *Nap1l5*, *Peg13* und *Slc38a4* als geprägt identifiziert werden (Smith *et al.*, 2003). Ein Vorteil der Methode besteht darin, dass die Methylierungsunterschiede, die zu verschieden großen *Hpa*II-Fragmenten führen, mittels PCR und Gel-Elektrophorese einfach sichtbar gemacht werden können. Der Ausschluss von Genen, die in uniparentalen Embryonen nicht untersucht werden können, ist ein Nachteil der Methode.

Bei der PCR-basierten Methode des *differential display* werden Expressionsniveaus verschiedener RNA-Populationen verglichen (Liang und Pardee, 1992). Im ersten Schritt werden zu vergleichende RNA-Populationen mit oligo(dT)-Primern in cDNA umgeschrieben. Durch den Einsatz verschiedener Kombinationen von Primern in einer PCR wird im zweiten Schritt ein kleiner Bestandteil der cDNA-Moleküle amplifiziert. Diese werden im dritten Schritt mittels Polyacrylamid-Gel-Elektrophorese aufgetrennt. Differentiell exprimierte Transkripte werden so sichtbar gemacht. Mit dem *differential display* kann gleichzeitig eine größere Anzahl an RNAs verglichen werden ((Lottspeich und Zorbas, 1998), Seite 868). Mit dieser Methode wurde erstmalig das geprägte Gen *Impact* auf dem Maus-Chromosom 18 identifiziert (Hagiwara *et al.*, 1997). Auch der Ansatz des *differential display* schließt das gesamte Genom ein, wird jedoch durch die Sequenzkombinationen der eingesetzten PCR-Primer limitiert (Oakey und Beechey, 2002).

Mit der Methode der *serial analysis of gene expression* (SAGE) werden Expressionsprofile in mRNA-Proben ermittelt und verglichen (Velculescu *et al.*, 1995). Dieser Ansatz ist sequenzunabhängig, und es können dabei selbst geringe Transkriptmengen detektiert werden. SAGE basiert auf zwei Prinzipien. Ein 9 bis 12 bp langer so genannter *sequence tag* (oder *tag*) ist zum einen informativ genug, um ein Transkript eindeutig zu identifizieren, und zum anderen lässt sich eine Verkettung dieser *sequence tags* in einem Klon durch eine serielle Sequenzierung effektiv analysieren. Im ersten Schritt wird vom 3´-Ende jeder mRNA ein *tag* synthetisiert. Anschließend werden diese mit Adaptoren versehen und miteinander verkettet. Im nächsten Schritt werden die verketteten *tags* in ein Plasmid kloniert und sequenziert. Insgesamt müssen zwischen 200.000 und 300.000 *tags* sequenziert werden, um auch seltene Transkripte analysieren zu können. Bei der Auswertung kann durch die Häufigkeit, mit der ein *tag* vorkommt, ein Rückschluss auf die

Expressionshöhe gezogen werden. Ein wichtiger Vorteil von SAGE ist die Möglichkeit auch unbekannte Transkripte analysieren zu können. Nachteilig ist der hohe Aufwand durch die Sequenzierung von bis zu 300.000 *tags*. Wenn nur geringe Mengen an zu untersuchendem Gewebe oder Zellen vorhanden sind, stellt SAGE jedoch die Methode der Wahl dar (Velculescu *et al.*, 2000). So wurde durch SAGE in Neuroblastom-Zelllinien eine ungewöhnlich hohe *DLK1*-Expression mit ca. 50% der gesamten mRNA bei einem Erhalt der genomischen Prägung nachgewiesen (Van Limpt *et al.*, 2003).

Mit der Entwicklung der cDNA-Microarray-Technologie lassen sich mit geringerem Aufwand Expressionsprofile erstellen und vergleichen. Zur Identifizierung von neuen geprägten Genen werden dabei meist mRNA-Populationen von androgenetischen und parthenogenetischen Mausembryonen verglichen (Oakey und Beechey, 2002). Auf einem Objektträger sind einige tausend DNA-Fragmente geordnet immobilisiert, die als Zielsequenz für zu untersuchende Gene dienen. Die zu vergleichenden RNAs werden zunächst in cDNA umgeschrieben und dabei mit unterschiedlich fluoreszierenden Farbstoffen markiert. Beide cDNA-Präparationen werden auf den Objektträger aufgetragen und an die immobilisierten Ziel-DNAs hybridisiert. Dabei ergeben sich spezifische Muster, die miteinander verglichen werden (Kurella *et al.*, 2001). So wurde beispielsweise das paternal exprimierte Gen *Dlk1* identifiziert (Kobayashi *et al.*, 2000). Dieser Ansatz ist für eine Hochdurchsatzsuche prädestiniert, jedoch durch die frühe Sterblichkeit uniparentaler Mausembryonen limitiert. Außerdem können nur bekannte Transkripte untersucht werden (Oakey und Beechey, 2002).

In einer Microarray-Studie wurden Expressionsprofile von 27663 cDNA-Klonen, die 18609 Transkriptionseinheiten repräsentieren, zwischen androgenetischen (AG) und parthenogenetischen (PG) Mausembryonen verglichen und dabei 2101 neue Imprinting-Kandidatengene identifiziert (Nikaido *et al.*, 2003). Die untersuchten Transkripte wurden als Kandidatengene deklariert, wenn ein mindestens zehnfacher Expressionsunterschied zwischen AG und PG messbar war (siehe Abb. 6). Es wird angenommen, dass von diesen Kandidatengenen 1403 Transkripte maternal und 698 Transkripte paternal geprägt sind. In der Gesamtzahl der untersuchten Klone waren auch elf bekannte geprägte Gene als Positivkontrollen enthalten, die durch den Versuchsansatz ebenfalls bestätigt werden konnten. Alle Kandidatengene wurden nach verschiedenen Kriterien wie genomische Lokalisation, kodierende Eigenschaft, und Zugehörigkeit zu vorhergesagten Imprinting-Krankheits-Loci analysiert und die Ergebnisse in der EICO-Datenbank zusammengestellt (Nikaido *et al.*, 2004). 39 Kandidatengene kartieren dabei in bekannten geprägten Regionen. Unter den Kandidatengenen konnte in nachfolgenden Versuchen die genomische Prägung von zwei nicht-kodierenden RNAs (PX00010K13 und PX00113D24), die innerhalb des zum Menschen homologen Prader-Willi-Syndrom-Locus auf MMU7cen kartieren, verifiziert werden. Die Autoren der Microarray-Studie vermuten auch eine unbekannte Anzahl falsch

positiver Kandidatengene, die unter der Kontrolle geprägter Gene, d.h. durch so genannte *downstream*-Effekte, fälschlicherweise als Kandidatengene identifiziert wurden.

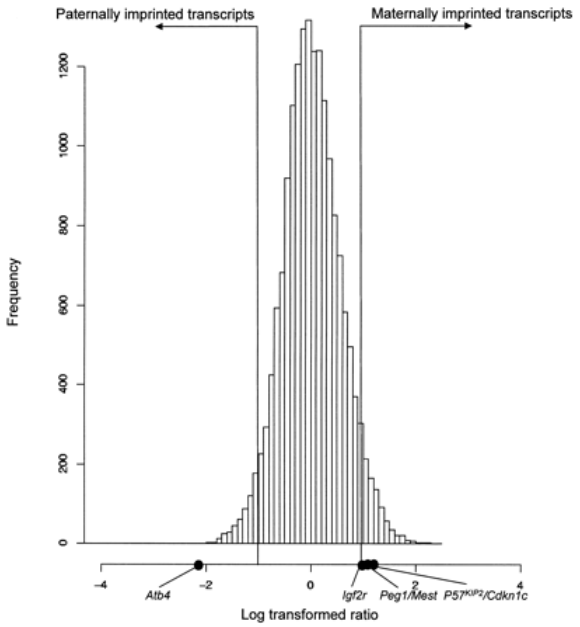

Abbildung 6. Vergleich der Expressionsniveaus von androgenetischen und parthenogenetischen Mausembryonen dargestellt als Histogramm. Die Expressionsunterschiede sind durch Logarithmierung transformiert. Ein Expressionsunterschied größer als 10 entspricht logarithmiert einem Wert von >(±1) und grenzt die Kandidatengene ab (von (Nikaido *et al.*, 2003)).

In einer vorangegangenen Studie von Mizuno *et al.* (2002) wurden vergleichbare cDNA-Microarrays benutzt und ein Grenzwert für Expressionsunterschiede von log = ±1,88 gewählt (entspricht 75fach). Dieser Grenzwert leitete sich aus den acht enthaltenen geprägten Genen ab, die als Positivkontrollen dienten. Aus der Gesamtzahl der 18.816 untersuchten cDNA-Klone wurden auf diesem Wege 25 Kandidatengene identifiziert. Darunter befanden sich die acht bekannten geprägten Gene, drei X-chromosomale Gene sowie drei bislang unbekannte geprägte Gene (*Slc38a4*, *Asb4* und *Dcn*).

1.4.2. Sequenzbasierte Vorhersagen

Der genomische Kontext geprägter und nicht geprägter Gene weist statistisch signifikante Unterschiede auf. Um die Promotorregion und innerhalb eines 100 bis 200 kb großen Sequenzbereiches sind bei geprägten Genen ein geringerer Anteil von SINE-Transposons, vor allem Alu- und MIR-Elemente, ein höherer GC-Gehalt sowie ein höherer Anteil an LINE-1-Transposons als bei nicht geprägten Genen zu finden (Greally, 2002; Allen *et al.*, 2003; Smith *et al.*, 2004). Daraus leitete sich die Idee ab, die Sequenzmuster

bekannter geprägter Gene zur Suche neuer Imprinting-Kandidatengene mit Hilfe bioinformatischer Programme zu nutzen. So wurden in einer Studie über 1000 neue Kandidatengene für monoallelische Expression vorhergesagt (Allen *et al.*, 2003). Auch ein getrennter Vergleich von maternal und paternal exprimierten Genen mit einer Gruppe von Kontrollgenen weist auf deutliche Unterschiede hin. Die maternal exprimierten Gene haben einen höheren GC-Gehalt und eine größere Anzahl an CpG-Inseln. Die paternal exprimierten Gene zeigen hingegen einen höheren Anteil an LINE-1-Transposons (Greally, 2002). Zusammengefasst ergibt die Analyse des genomischen Kontextes geprägter Gene statistisch signifikante Unterschiede zu nicht geprägten Genen und lässt sogar Unterschiede zwischen maternal und paternal exprimierten Genen erkennen (Smith *et al.*, 2004). In einer vergleichenden Sequenzanalyse der humanen und murinen Beckwith-Wiedemann-Syndrom-Region auf HSA11p15.5 bzw. MMU7dist wurden konservierte CpG-Inseln sowie konservierte nicht exonische Sequenzen identifiziert (Onyango *et al.*, 2000).

In einer weitergehenden Studie wurden die spezifischen Sequenzmuster geprägter Gene erstmalig für eine Suche neuer Imprinting-Kandidatengene im gesamten Mausgenom herangezogen (Luedi *et al.*, 2005). Im ersten Schritt der Untersuchung wurden spezifische Sequenzmotive wie repetitive Elemente, Transkriptionsfaktorbindungstellen und CpG-Inseln hinsichtlich Vorkommen und Position an 44 bekannten geprägten Gene analysiert und folgend ein Vorhersagealgorithmus trainiert. Dabei zeigte sich, dass die Anzahl und Lage der repetitiven Elemente wichtigstes Kontextmerkmal ist. So sind intronisch gelegene repetitive Elemente gefolgt von 10 kb stromaufwärts und 100 kb stromabwärts vom Promotor gelegenen Elementen besonders charakteristische Merkmale. Auch die Art der Elemente ist dabei von großer Bedeutung. Endogene Retrovirus Elemente gefolgt von LINE-1-Elementen kommen besonders häufig vor. Von den untersuchten Bindungsstellen für Transkriptionsfaktoren spielen SRF-, NFuE1- und AP2-Bindungsstellen die größte Rolle für die Vorhersage geprägter Gene. Von 23.788 analysierten Genen wurden 600 als neue Imprinting-Kandidatengene identifiziert, die ein vergleichbares Sequenzmuster wie die anfänglich analysierten 44 geprägten Gene aufweisen. Die Analyse der Kandidatengene mit einem zweiten Algorithmus hinsichtlich parentaler Expression ergab, dass 64% der 600 potenziell geprägten Gene maternal exprimiert sind. In nachfolgenden Analysen zur Verifizierung der Ergebnisse wurden die 44 geprägten Gene, die zum Training des Algorithmus benutzt wurden, vollständig und richtig identifiziert und 93% der nicht geprägten Gene (530 nicht geprägte Gene untersucht) wurden korrekt als nicht geprägt identifiziert. Auch bekannte geprägte Gene, wie *Nnat* und *U2af1-rs1*, die nicht zum Training benutzt wurden und ein sehr spezifisches Prägungsmuster aufweisen, wurden durch das Programm korrekt identifiziert.

1.5. Validierung von Imprinting-Kandidatengenen

Die oben geschilderten Vorgehensweisen zur Identifizierung neuer geprägter Gene, die auf laborexperimentellen oder bioinformatischen Ansätzen (z.B. Microarray und SAGE) beruhen, sind allein für eine zuverlässige Aussage nicht ausreichend. Es ist folglich notwendig, die monoallelische Expression der Kandidatengene durch eine allelspezifische Expressionsanalyse zu validieren (Oakey und Beechey, 2002).

Die Verwendung von heterozygoten transkribierten Polymorphismen als Marker ist eine einfache Möglichkeit zur allelischen Diskriminierung. Bei der Validierung von murinen Imprinting-Kandidatengenen werden interspezifische F1-Hybriden verwendet, die man durch die Kreuzung verschiedener Subspezies erzeugt. Meistens wird der *Mus musculus*-Inzuchtstamm C57BL/6 mit einer anderen *Mus musculus*-Subspezies wie z.B. *Mus musculus castaneus* (Cast/Ei) gekreuzt (Oakey und Beechey, 2002).

Geeignete Polymorphismen sowie informative Individuen werden durch direkte Sequenzierung von PCR-Produkten aus genomischer DNA identifiziert. Zu den oft verwendeten methodischen Ansätzen zur allelischen Expressionsanalyse zählen neben *RNAse protection assays* (Bartolomei *et al.*, 1991) vor allem RT-PCR mit anschließender RFLP-Analyse (RFLP steht für Restriktionsfragmentlängenpolymorphismus) (Oakey und Beechey, 2002). So wurde beispielsweise die monoallelische Expression der geprägten Gene *Peg3* (Kuroiwa *et al.*, 1996) und *Peg1* (Kaneko-Ishino *et al.*, 1995) durch RT-PCR gefolgt von einer RFLP-Analyse bestätigt.

Allerdings sind mit diesen Methoden aufgrund ihrer geringen Genauigkeit keine exakten quantitativen Aussagen zu erhalten. Mit Hilfe einer neuen Anwendung, die auf der Pyrosequencing™-Technologie basiert und die im Kapitel 2.2.6 näher erläutert wird, lassen sich hingegen auch geringere allelische Expressionsunterschiede genau detektieren. Neben der Möglichkeit monoallelisch exprimierte Gene zu analysieren, kann man mit dieser neuen Anwendung auch präferentiell allelisch exprimierte Gene genau analysieren (Rutsch *et al.*, 2003; Sun *et al.*, 2005).

1.6. Aufgabenstellung

Inwieweit experimentelle Studien wie von Nikaido *et al.* (2003) oder auch sequenzbasierte Vorhersagen wie von Luedi et al. (2005) geeignet sind, um effizient neue geprägte Gene im Mausgenom zu identifizieren, ist bislang nicht in weiteren Studien untersucht.

Das Hauptziel der vorliegenden Arbeit war es, die Kandidatengene aus beiden Studien zu priorisieren, jeweils eine hinreichend große Subgruppe auszuwählen und diese experimentell zu überprüfen, um neue geprägte Gene zu identifizieren und die Effizienz des experimentellen und sequenzbasierten Ansatzes zu bewerten.

Dazu sollten durch Sequenzierung genomischer DNA von (C57BL/6 x Cast/Ei)F1-Hybriden geeignete transkribierte Einzelnukleotidpolymorphismen (tSNPs) in den Kandidatengenen identifiziert werden. Anhand dieser tSNPs sollte anschließend mittels Pyrosequencing™ von RT-PCR-Produkten aus (C57BL/6 x Cast/Ei)F1-Embryonen (E11.5) die vorhergesagte parentale monoallelische Expression verifiziert werden. Allelische Expressionsunterschiede sollten durch die Analyse von informativen Hybriden der reziproken Kreuzung bestätigt werden.

Des Weiteren sollte die genomische Prägung neben der Untersuchung im Embryonalstadium auch im adulten Stadium analysiert und weiter charakterisiert werden. Die orthologen Gene von den identifizierten geprägten Genen im Mausgenom sollten auch beim Menschen auf monoallelische Expression untersucht und weiter charakterisiert werden. Die geprägten murinen und humanen Loci sollten analysiert und die Bedeutung der neu identifizierten geprägten Gene hinsichtlich Imprinting-Regulation untersucht und bewertet werden.

2. Material & Methoden

2.1. Material

Soweit nicht gesondert aufgeführt, wurden für alle Experimente handelsübliche Enzyme, Chemikalien und sonstige Reagenzien im analytischen Maßstab sowie handelsübliches Material und Standardlaborgeräte verwendet.

2.1.1. Mausstämme

Für die Untersuchungen am Mausmodell wurden Mäuse aus dem *Mus musculus*-Inzuchtstamm C57BL/6 und der *Mus musculus*-Subspezies *Mus musculus castaneus* (Cast/Ei) verwendet (Charles River Laboratories, Wilmington MA, USA). Die Zucht der Mäuse erfolgte unter der Koordination von Dr. Ulrich Zechner am Institut für Humangenetik der Universität Mainz. Die beiden Mausstämme, deren F1-Hybriden und die entsprechenden reziproken Kreuzungen wurden als Quelle zur Gewinnung von RNA sowie genomischer DNA benutzt. Durch natürliche Verpaarung wurden zeitlich festgelegte Embryonen erzeugt, wobei der Tag nach der Empfängnis als Tag 0,5 definiert wurde. Bei der Benennung der F1-Hybriden sowie der reziproken Kreuzung wird zuerst das Muttertier angegeben (z.B. (C57BL/6 x Cast/Ei)F1 mit C57BL/6 als Weibchen und Cast/Ei als Männchen). Für die allelspezifische Expressionsanalyse wurden die Embryonen am Tag 11,5 nach der Empfängnis (E11.5) und verschiedene Gewebe aus adulten Tieren entnommen.

2.1.2. Humanes Material

Für die Untersuchungen am Menschen wurden RNA und genomische DNA aus Abortmaterial (fetales Gehirn von 11 Individuen und fetaler Testis von 10 Individuen) sowie RNA und genomische DNA aus Biopsien (adulter Testis von 6 Individuen) verwendet. Die Proben wurden von Dr. Ulrich Zechner und Dr. Danuta Galetzka vom Institut für

Humangenetik der Universität Mainz zur Verfügung gestellt. Des Weiteren wurden Proben aus verschiedenen fetalen und adulten Fibroblasten, Lymphoblastoiden und Amniozyten für Expressionsexperimente eingesetzt.

2.1.3. Oligonukleotide

Zur Identifizierung transkribierter SNPs durch DNA-Sequenzierung und zur allelspezifischen Expressionsanalyse mittels QUASEP (Quantification of Allele-Specific Expression by Pyrosequencing™) wurden die im Anhang aufgelisteten Oligonukleotide verwendet (siehe Tab. 13 und 14). Bei der QUASEP-Analyse wurden für jeden Versuchsansatz drei Primer eingesetzt. Dazu zählt neben dem Vorwärts- und Rückwärtsprimer noch ein Sequenzierprimer, der in unmittelbarer Nähe zum untersuchten tSNP bindet. Dabei ist entweder der Vorwärts- oder der Rückwärtsprimer mit Biotin markiert. Der Sequenzierprimer ist komplementär zu dem biotinmarkierten Einzelstrang. Die QUASEP-Primer wurden mit Hilfe der Assay Design-Software des Herstellers ausgewählt. Das Programm gibt für jeden Versuchsansatz einen Qualitätswert zwischen 0 und 100 aus, wobei 100 den bestmöglichen Wert darstellt.

Zur Herstellung der Sonden für die Northern-Blot-Hybridisierung wurden für das murine AK043056 (NB-43056) und für das humane Ortholog von AK043056 (hC19-43056) die in Tabelle 3 aufgelisteten Oligonukleotide eingesetzt.

Tabelle 3. Primer zur Amplifikation der Sonden für die Northern-Blot-Hybridisierung.

Primername	F-Primer (Sequenz 5´- 3´)	R-Primer (Sequenz 5´- 3´)	T_{melt} [C°]	Länge [bp]
hC19-43056	GGCCATGACTCTTTGTTGAGA	TTGCTAACGTGTGCCAACAT	60	286
NB-43056	AGGCGTCCTTTCTCTTAGGG	ATCTGCTCCCTGGTGTCATT	60	319

Für die RACE-PCR des humanen und murinen AK043056-Transkripts wurden neben den Primern des Gene Racer-Kits (Invitrogen GmbH, Karlsruhe, Deutschland) geeignete genspezifische Oligonukleotide eingesetzt (siehe Tab. 4). Bis auf die QUASEP-Primer erfolgte das Design für alle Oligonukleotide mit Hilfe der Primer3-Software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi).

Tabelle 4. Primer für die RACE-PCR.

	Primername	Sequenz 5´- 3´	Tmelt [C°]
5´-Ende Mensch	h5GSP-R2	TTCACTGGATCCAAAGTCAGCGTAT	72
	h5nestGSP-R2	CCAGGGTATGCTGATCACAAGGTCA	76
3´-Ende Mensch	h3GSP-F	TCCATTCTGGAGAATTCTTGGGCCAGTG	84
	h3nestGSP-F	TGACCCGTCTTAATCCATCCTCCATTTG	82
5´- Ende Maus	m5GSP-R2	GCACGGTTAAGTAGGGCTTAGGAATG	78
	m5nestGSP-R2	TCGATGGCAAATATGACTGTCCTTTC	74
3´-Ende Maus	m3GSP-F2	GGCAGAATTCTTAGCCAGCCCTCT	74
	m3nestGSP-F2	TCTGGTCCATCTCTCTAGGGTCTGC	78

2.2. Methoden

2.2.1. Allgemeine molekularbiologische Methoden

Molekularbiologische Standardmethoden wie Polymerase-Kettenreaktion, Gel-Elektrophorese, cDNA-Synthese, Klonierung, enzymatischer Verdau sowie das Herstellen von Standardlösungen wurden nach Sambrook et al. (1989) durchgeführt und werden methodisch nicht im Detail beschrieben.

2.2.2. RNA- und DNA-Isolierung

Die RNA- (Gesamt-RNA) sowie die genomischen DNA-Proben wurden überwiegend von Dr. Ulrich Zechner und seinen Mitarbeitern am Institut für Humangenetik der Universität Mainz präpariert. Die Isolierung von genomischer DNA aus verschiedenen Geweben und Embryonen erfolgte durch Aussalzen nach Standardprotokoll (Sambrook *et al.*, 1989).

Die Präparation von Gesamt-RNA aus Embryonen, adulten Geweben und Zellmaterial wurde mit dem TRIzol-Reagenz (Invitrogen, Karlsruhe, Deutschland) nach dem Protokoll des Herstellers durchgeführt. Da die Kontamination der RNA-Proben mit genomischer DNA zu einer Verfälschung der Expressionsuntersuchungen führen würde, wurden alle RNA-Proben zweifach mit DNAse (DNAfree, Ambion Ltd., Huntington, UK) nach dem Protokoll des Herstellers verdaut und die Qualität des Verdaus mittels Test-PCR überprüft. Die Konzentration der RNA- und DNA-Proben wurde photometrisch bestimmt. Um die Qualität der RNA-Präparationen abzuschätzen, wurde jeweils 1 bis 5 µL auf ein mit Ethidiumbromid versetztes einprozentiges Agarose-Gel aufgetragen und dokumentiert.

2.2.3. cDNA-Synthese

Um die Expression der Kandidatengene zu untersuchen, musste die RNA in cDNA umgeschrieben werden. Mit dem cDNA-Synthese-Kit von Fermentas (Fermentas GmbH, St. Leon-Rot, Deutschland) wurden jeweils 3 µg der RNA nach dem Protokoll des Herstellers umgeschrieben. Durch die Verwendung von oligo$(dT)_{18}$-Primern sollte präferentiell mRNA transkribiert werden. Nach Angaben des Herstellers beträgt die Ausbeute bei der cDNA-Synthese mindestens 50%, so dass die minimale cDNA-Konzentration mit 50% der eingesetzten RNA-Menge angenommen wurde.

2.2.4. Polymerase-Kettenreaktion

Die PCR wurde standardmäßig in 25µL angesetzt. Die Amplifikation erfolgte in 40 bis 45 Zyklen jeweils mit Denaturierung (10 – 30 Sekunden), Primeranlagerung (10 – 30 Sekunden) und Elongation (10 – 60 Sekunden) in Abhängigkeit von der Fragmentlänge. Als Matrize wurde mindestens 6ng genomische DNA bzw. mindestens 12ng cDNA eingesetzt. Des Weiteren wurden 1 bis 3mM $MgCl_2$, 1 Unit Taq-Polymerase oder Immolase (Bioline GmbH, Luckenwalde, Deutschland), 200nM je dNTP, 200nM je Primer (siehe Anhang) sowie 1xPCR-Puffer (10mM Tris-HCl, pH 8,3; 50mM KCl) eingesetzt.

Die Qualität und Quantität der PCR-Produkte wurde in einem zwei- bis vierprozentigen Agarose-Gel in 1xTAE-Puffer und unter Verwendung verschiedener quantitativer DNA-Leitern kontrolliert.

2.2.5. DNA-Sequenzierung

Zur Bestimmung der Sequenz eines DNA-Abschnitts gibt es zwei methodische Ansätze. Während es für das chemische Verfahren (Maxam und Gilbert, 1980) nur wenige Anwendungsmöglichkeiten gibt, wird für DNA-Sequenzierungen hauptsächlich die enzymatische Sanger-Methode verwendet (auch als Kettenabbruch-Verfahren bekannt (Sanger *et al.*, 1977)). Die DNA-Sequenzierungen der vorliegenden Arbeit basierten auf dem Sanger-Verfahren unter Verwendung des BigDye™ Terminator v1.1 Cycle Sequencing-Kit und anschließender Auftrennung der Produkte auf einem ABI 3730 DNA Analyzer (beide Applied Biosystems, Foster City, MA, USA).

Die PCR-Produkte wurden vor der Sequenzreaktion zum Abbau der überschüssigen Oligonukleotide mit 1,5 Units Exonuklease 1 (New England Biolabs, Frankfurt, Deutschland) und 0,3 Units alkalischer Phosphatase (SAP, Promega, Mannheim, Deutschland) behandelt (25 Minuten bei 37°C und 15 Minuten bei 72°C). Für die Sequenzreaktion wurden in 10 µL Gesamtvolumen jeweils 0,5 µL BigDye-Prämix, 250 nM Primer, 2,0 µL DNA-Template (gereinigtes PCR-Produkt), 1xSequenzier-Puffer und ddH_2O (HPLC gereinigtes Wasser) eingesetzt. Die Sequenzreaktion erfolgte in 30 Zyklen nach dem Protokoll des Herstellers. Nach der Sequenzreaktion wurden die Proben mittels Sephadex G-50 Gelfiltration (Amersham Bioscience, Uppsala, Schweden) nach Standardprotokoll aufgereinigt. Die Auswertung der DNA-Sequenzierung erfolgte mit dem Programm SeqMan von LaserGene (GATC Biotech, Konstanz, Deutschland).

2.2.6. Allelische Expressionsanalyse mit Pyrosequencing™

Pyrosequencing™ ist eine Methode zur DNA-Sequenzierung und beruht auf der Detektion von freigesetztem Pyrophosphat (PP_i) während der DNA-Synthese (siehe Abb. 7). In einer Enzymkaskade werden Lichtphotonen erzeugt, deren Anzahl proportional zur Anzahl der eingebauten Nukleotide ist.

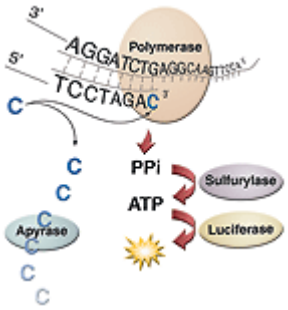

Abbildung 7. Schematische Darstellung des Pyrosequencing™-Prinzips (Quelle: www.biotage.com).

Die Zugabe der Nukleotide wird in zeitlicher Reihenfolge als Pyrogramm dargestellt (siehe Abb. 8). Dabei werden die Signalhöhen zueinander ins Verhältnis gesetzt, woraus sich verschiedene Anwendungen ergeben (Ronaghi, 2001; Uhlmann *et al.*, 2002; Ahmadian *et al.*, 2006).

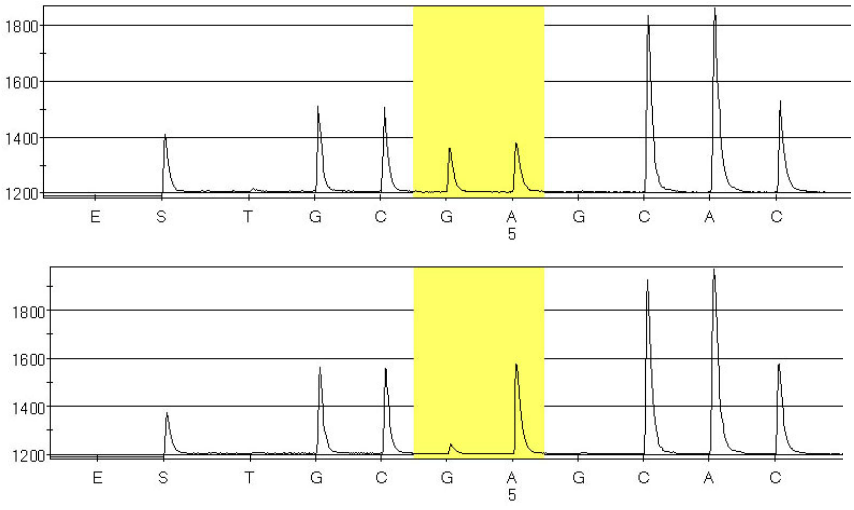

Abbildung 8. Pyrosequencing™-Pyrogramme. Auf der Abszisse sind in zeitlicher Reihenfolge die zugegebenen Nukleotide und auf der Ordinate ist die dazu gemessene Lichtintensität dargestellt. Die obere Sequenz zeigt einen G/A-Polymorphismus in heterozygoter Form mit einer 50:50-Verteilung der beiden Allele im Gen *PEG3*. Analysiert wurde ein PCR-Produkt aus genomischer DNA. *PEG3* ist ein bekanntes geprägtes Gen und ist paternal exprimiert. Die polymorphe Position ist im Pyrogramm gelb hervorgehoben. Die Sequenz kann direkt abgelesen werden und lautet GCG/ACCAAC. Die Analyse von RT-PCR-Produkten des gleichen heterozygoten Individuums weicht hingegen deutlich ab und zeigt eine allelische Verteilung von 10% zu 90% (G- zu A-Allel). Das G-Allel ist maternalen Ursprungs und das A-Allel ist paternalen Ursprungs.

Eine dieser Anwendungsmöglichkeiten ist die Untersuchung allelspezifischer Expression, die auch QUASEP (Quantification of Allele-Specific Expression by Pyrosequencing™) genannt wird (Rutsch *et al.*, 2003). Im Gegensatz zur Genotypisierung werden bei QUASEP RT-PCR-Produkte analysiert. Dabei erfolgten die Herstellung der Standardlösungen sowie die Präparation der Proben nach dem Protokoll des Herstellers. Die Sequenzierung wurde mit dem PSQ™ HS 96 (Biotage AB, Uppsala, Schweden) durchgeführt. Das Design der Primer-Sets für jeden Versuchsansatz wurde mit der Assay Design-Software des Herstellers durchgeführt.

Da die meisten autosomalen Gene biallelisch exprimiert werden, ist bei der QUASEP-Analyse von heterozygoten Individuen normalerweise auch Heterozygotie mit einer Gleichverteilung der Signale an der polymorphen Position zu erwarten. Eine Abweichung von diesem Muster kann zum Beispiel durch epigenetische Mechanismen, wie sie bei geprägten Genen zu beobachten sind, verursacht sein und zu präferentieller bzw. monoallelischer Expression von einem der beiden Allele führen (siehe Abb. 8 unten). Die Genauigkeit von QUASEP und die damit verbundene Möglichkeit, auch geringe allelische Expressionsunterschiede zuverlässig zu identifizieren, wurde durch Vorversuche belegt (siehe auch (Uhlmann *et al.*, 2002)) und wird durch die folgende Abbildung verdeutlicht.

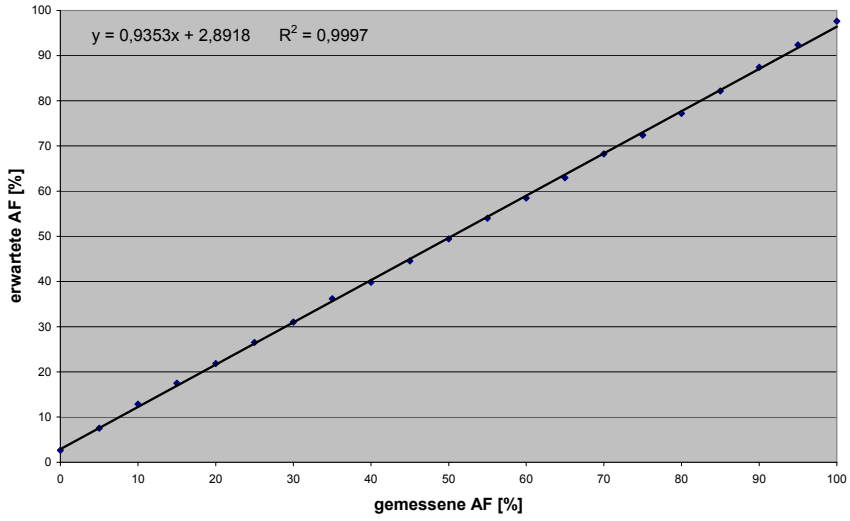

Abbildung 9. Eichgerade zur Bestimmung der tatsächlichen Allelfrequenz (AF). Die gemessene AF ist in Relation zur erwarteten AF dargestellt. Dazu wurden PCR-Produkte von verschieden homozygoten Individuen in verschiedenen Verhältnissen gemischt (von 0% bis 100% in 5%-Schritten) und mit dem Pyrosequenzer analysiert. Über den gesamten Bereich ist ein linearer Verlauf zu beobachten (R^2=0,9997) (verändert nach A. Brinckmann).

Beim Pyrosequencing™ sind zum Teil Hintergrundsignale zu beobachten, die hauptsächlich durch Sekundärstrukturen vom 3´-Ende des DNA-Einzelstranges oder die unspezifische Bindung des Sequenzierprimers verursacht werden. Beides führt zu einer Strangverlängerung durch unspezifische Startpunkte. Wenn sich derartiger Hintergrund zu

den Signalen an der SNP-Position addiert, wird die quantitative Aussage der Rohdaten verfälscht. Auffällig wird dies bei der Analyse von homozygoten Proben, bei denen im Pyrogramm für die nicht vorhandene allelische Variante ein Signal sichtbar wird (siehe Abb. 11 oben (*Zim1*) und Abb. 13 oben (AK043056)).

Um Hintergrundsignale sowie andere Faktoren, die die Signalhöhe beeinflussen, systematisch von den Messwerten zu unterscheiden, wurde für jeden Versuchsansatz eine vereinfachte Eichgerade erstellt und die gemessenen Rohdaten auf der Grundlage des in Abbildung 9 beobachteten linearen Verlaufs normalisiert. Für die vereinfachte Eichgerade wurden zunächst RT-PCR-Produkte der beiden homozygoten Elterntiere jeweils mindestens vierfach analysiert und die Mittelwerte der gemessenen Allelfrequenzen dem Erwartungswert gleichgesetzt. So wurde beispielsweise bei einem C/T-Polymorphismus in Proben mit C/C-Genotyp ein Allelverhältnis von 85% (C-Allel) zu 15% (T-Allel) ermittelt. Der Wert für das T-Allel ist bei einem C/C-Genotyp durch Hintergrundsignale bedingt. Daher wurde definiert, dass der 85%-Anteil des C-Allels dem Erwartungswert von 100% entsprechen. Die Darstellung der gemessenen Allelfrequenz gegen die erwartete Allelfrequenz ergab eine Regressionsgerade (siehe Abb. 10), aus deren Geradengleichung anschließend die tatsächliche Allelfrequenz der heterozygoten Proben berechnet wurde. Die heterozygoten Proben wurden mindestens zehnfach analysiert sowie Mittelwert und Standardabweichung berechnet. Die ermittelten Standardabweichungen lagen zwischen 0,7 und 11,6 (Durchschnitt bei 3,8; siehe Tab. 5).

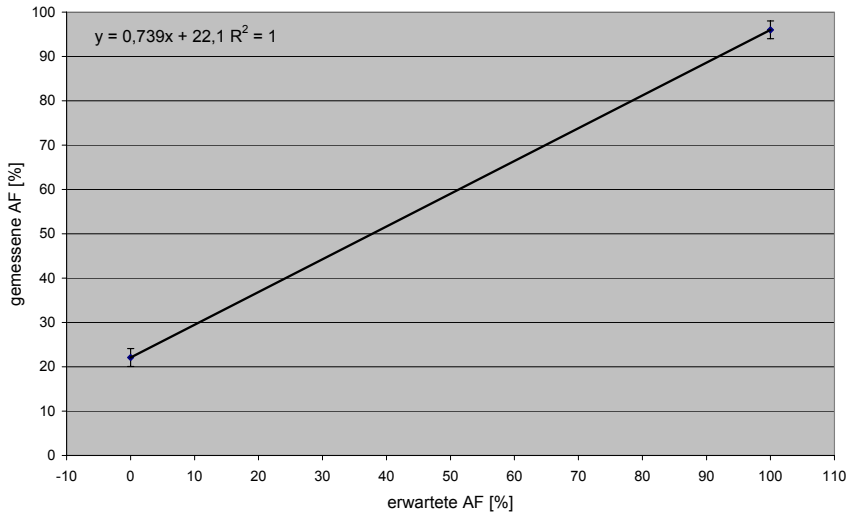

Abbildung 10. Beispiel einer verwendeten Eichgerade zur Bestimmung der tatsächlichen Allelfrequenz (AF). Die gemessene AF ist in Relation zur erwarteten AF graphisch dargestellt. RT-PCR-Produkte der beiden homozygoten Elterntiere wurden jeweils mindestens vierfach analysiert und die gemessenen Werte dem Erwartungswert gleichgesetzt. Die jeweiligen Mittelwerte sind im Diagramm dargestellt. Aus der resultierenden Regressionsgerade wurde die tatsächliche AF der heterozygoten Proben berechnet. Die Standardabweichung (SD = 2,0) ist per Fehlerbalken abgebildet.

2.2.7. Sondenpräparation zur Northern-Blot-Hybridisierung

Eine etablierte Methode, um Informationen zur Expression bzw. zum Expressionsmuster eines spezifischen Gens zu erhalten, ist die Northern-Blot-Hybridisierung ((Strachan und Read, 2005), Seite 199f). Um zusätzlich zu den Ergebnissen der RT-PCR-Experimente weitere Informationen über die Expression von AK043056 und des humanen Orthologs zu gewinnen bzw. diese zu bestätigen, wurden hierfür DNA-Sonden mittels PCR unter Verwendung der in Tabelle 3 aufgelisteten Oligonukleotide hergestellt. Die PCR-Produkte wurden anschließend aus einem zweiprozentigen Agarose-Gel ausgeschnitten und mittels Glasmilch (Easy Pure von Biozym GmbH, Oldendorf, Deutschland) aufgereinigt (Vogelstein und Gillespie, 1979).

Die Sonden wurden im nächsten Schritt mit α^{32}P-dCTP und Random-Hexamer Primern nach Standardprotokoll radioaktiv markiert (Sambrook *et al.*, 1989) und mit Microspin S200-Säulen (Amersham Bioscience, Uppsala, Schweden) aufgereinigt, um nicht gebundene Nukleotide zu entfernen. Die [^{32}P]-Einbaurate wurde vor der Hybridisierung mit einem Szintillator Typ LS6000SC (Beckmann Coulter, Krefeld, Deutschland) überprüft.

2.2.8. Hybridisierung der Northern-Blots

Die radioaktiv markierten Sonden wurden an kommerziell erhältliche Multiple Tissue Northern-Blots (MTN-Blots) hybridisiert (BD Bioscience Clontech, Palo Alto, CA, USA). Der verwendete murine MTN-Blot (Artikel-Nr.: 636808) enthielt jeweils 2 µg fixierter poly(A)-RNA aus Herz, Gehirn, Milz, Lunge, Leber, Skelettmuskel, Niere und Testis fixiert. Auf dem humanen MTN-Blot (Artikel-Nr.: 636806) waren 2 µg poly(A)-RNA aus Herz, Gehirn, Plazenta, Lunge, Leber, Skelettmuskel, Niere und Pankreas fixiert. Die Blots ließen sich nach der Hybridisierung mit sterilem Wasser (mit 0,5% SDS) bei 90 bis 100°C strippen und bis zu fünfmal wieder verwenden.

Die Hybridisierung erfolgte bei 68°C mit dem mitgelieferten ExpressHyb-Puffer für 24 bis 48 Stunden nach dem Protokoll des Herstellers. Anschließend wurden die Blots nach Standardmethode gewaschen (Sambrook *et al.*, 1989).

Vor dem Nachweis der Hybridisierungsmuster wurden die Blots für mindestens fünf Stunden auf einer Phosphor Imaging-Platte (Fuji Photo Film Co., Ltd, Tokio, Japan) exponiert. Die Hybridisierung wurde für das murine wie auch für das humane Transkript jeweils einmal wiederholt und die Effizienz des verwendeten Protokolls mit der mitgelieferten β-Aktin-Sonde kontrolliert.

2.2.9. Real-time-RCR mit TaqMan® zur Expressionsanalyse

Um die Expression eines Gens nachzuweisen oder zu quantifizieren, gibt es verschiedene Methoden. Neben der Northern-Blot-Hybridisierung, *in situ*-Hybridisierung, RNAse-Protection-Assay und cDNA-Arrays kann die Expression sehr einfach durch RT-PCR überprüft werden. Zur Quantifizierung der Genexpression ist die Real-time RT-PCR mit TaqMan® jedoch die Methode der Wahl (Lee *et al.*, 1993; Giulietti *et al.*, 2001).

Zur Ergänzung der durch RT-PCR gewonnenen Expressionsdaten von AK043056 und *Usp29* wurden für beide Gene entsprechende vom Hersteller getestete Primer und Sonden eingesetzt und mit dem ABI Prism 7700 SDS (Applied Biosystems, Foster City, USA) die Expression relativ zu einander analysiert. Als endogene Kontrolle diente dabei β-Aktin.

Die PCR wurde in 20 µL (dreifacher Ansatz) und in 12,5 µL (zweifacher Ansatz) mit cDNAs aus Gehirn, Plazenta oder Testis angesetzt. Da in diesem Experiment keine relative oder absolute Quantifizierung von AK043056 zu *Usp29* angestrebt wurde, konnte auf die Erstellung von Eichgeraden verzichtet werden. Für die PCR wurden 10 ng cDNA, der Universal PCR Master Mix sowie der Primer- und Sonden-Prämix nach dem Protokoll des Herstellers verwendet (Applied Biosystems, Foster City, USA) und die Amplifikation in 50 Zyklen à 15 Sekunden bei 92°C und 60 Sekunden bei 60°C durchgeführt.

Die Quantifizierung wird anhand eines C_T-Wertes (C_T steht für *threshold cycle*) ausgedrückt. Dieser Wert zeigt an, wann eine Amplifikationskurve eine gewählte Basislinie überschritten hat. Die Basislinie wird dabei so gewählt, dass sie die Amplifikationskurven der zu vergleichenden Proben in der exponentiellen Phase schneidet. Folglich weist ein kleinerer C_T-Wert auf eine höhere Expression hin.

2.2.10. Bestimmung der Transkriptlänge durch RACE-PCR

Mit der RACE-PCR (*rapid amplification of cDNA ends*) ist es möglich, die vollständige Transkriptlänge einer mRNA zu bestimmen (Frohman *et al.*, 1988). Dabei wird eine Zielregion zwischen einer bereits bekannten Sequenz und einer Ankersequenz amplifiziert, die am 5´- oder 3´-Ende des Transkripts liegt. Zur Überprüfung der bekannten Transkriptlänge von AK043056 und zur Bestimmung der Länge des humanen Orthologs musste eine RACE-PCR mit anschließender Klonierung und Sequenzierung durchgeführt werden.

Die so genannte 5´-add-on-Mutagenese der mRNA sowie die reverse Transkription wurde mit dem Gene Racer™-Kit (Invitrogen GmbH, Karlsruhe, Deutschland) nach dem

Protokoll des Herstellers durchgeführt. In zwei sich anschließenden Schritten wurden die gesuchten 5´ bzw. 3´-Enden mit dem Elongase Enzym-Mix (PCR und *nested*-PCR) amplifiziert sowie anschließend im Agarose-Gel aufgetrennt und ausgeschnitten. Die Aufreinigung der ausgeschnittenen Gelstücke erfolgte mit dem SNAP Gel Purification-Kit (beide Invitrogen GmbH, Karlsruhe, Deutschland).

Im nächsten Schritt wurden die aufgereinigten Amplikons mit Hilfe des TOPO TA Cloning-Kits (Invitrogen GmbH, Karlsruhe, Deutschland) kloniert. Nach erfolgter Transformation kompetenter E.coli-Bakterien erfolgte die Anzucht bis zur 3ml-Kultur nach Standardprotokoll. Zur Selektion wurde Kanamycin (50µg/ml) verwendet. Aus den 3ml-Kulturen wurden die Plasmide mittels Minipräp (Mini Spin-Kit von Qiagen GmbH, Hilden, Deutschland) nach Standardprotokoll isoliert und anschließend in einem 0,6%igen Agarose-Gel überprüft. Für die Sequenzierung wurden T3- bzw. T7-Primer verwendet.

2.2.11. Datenbankrecherchen und Sequenz-Alignments

Die Sequenz und Lokalisation der Kandidatengene aus der Microarray-Studie wurden der EICO-Datenbank entnommen (http://fantom2.gsc.riken.jp/EICODB/). Darüber hinaus stammen alle Sequenzangaben insbesondere für die Kandidatengene der Luedi-Studie aus den Datenbanken des National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) oder sind der Ensembl-Datenbank (http://www.ensembl.org/) entlehnt.

Um die Lokalisation der Kandidatengene zu überprüfen, wurde das Programm BLAT Search (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) genutzt (Kent, 2002). Für Sequenzvergleiche wurden die auf der NCBI-Homepage aufrufbaren Programme BLASTN (Altschul *et al.*, 1990) oder BLAST 2 Sequences (Tatusova und Madden, 1999) verwendet. Das Programm BLAST von Ensembl (http://www.ensembl.org/Multi/blastview) wurde benutzt, um orthologe Regionen in anderen Spezies zu identifizieren. Für Homologievergleiche von orthologen Genen wurde das Programm MegAlign (Lasergene, GATC Biotech, Konstanz, Deutschland) und der Clustal W-Algorithmus (Thompson *et al.*, 1994) verwendet.

Expressionsdaten wurden der UniGene-Datenbank von NCBI (http://www.ncbi.nlm.nih.gov/) entnommen.

3. Ergebnisse

3.1. Imprinting-Kandidatengene der Microarray-Studie

3.1.1. Auswahl der Kandidatengene

Die ungewöhnliche Konzentration geprägter Gene in diskreten Regionen lässt die Hypothese zu, dass die Lokalisation eines Imprinting-Kandidatengens in einer geprägten Region ein vielversprechender Hinweis auf tatsächliche Prägung ist. Ausgehend von dieser Hypothese wurden alle Kandidatengene aus der EICO-Datenbank ausgewählt, die in bekannten geprägten Regionen lokalisiert sind.

In der Microarray-Studie (Nikaido *et al.*, 2003), deren Ergebnisse in der EICO-Datenbank abrufbar sind, wurde im Text auf den Seiten 1402 und 1403 auf 39 Kandidatengene ohne weitere Details verwiesen, die in bekannten geprägten Regionen liegen sollen. In der Datenbank sind jedoch 68 Gene mit diesem Merkmal aufgelistet. Diese 68 Gene wurden aus der Gesamtanzahl der 2101 neuen Kandidatengene ausgewählt (siehe Tab. 5). Darunter befinden sich auch 11 Gene, die auf dem distalen Maus-Chromosom 4 (MMU4dist) lokalisiert sind. Bisher wurde auf MMU4dist weder ein geprägtes Gen identifiziert noch diese Region als Imprinting-Region definiert. Dieser chromosomale Abschnitt von 98 bis 155 Mb stellt jedoch die orthologe Region zum humanen Chromosom 1p36.33-p31.3 dar, zu dem im Imprinted Gene Catalogue (siehe http://igc.otago.ac.nz/home.html) mehrere abnormale Phänotypen aufgelistet sind, die durch parentale Effekte auf eine Imprinting-Störung schließen lassen. Daher wurden die 11 Gene auf MMU4dist in die weiteren Untersuchungen eingeschlossen.

Anschließend wurde die genomische Organisation aller 68 Kandidatengene mit Hilfe von BLAT Search analysiert. Dabei sollten vor allem die in der EICO-Datenbank aufgelistete Lokalisation bestätigt wie auch die putativen Exon-/Intronstrukturen der Kandidatengene aufgeklärt werden. Im Ergebnis dieser Untersuchung wurden die drei Kandidatengene *Zim1* (Kim *et al.*, 1999), *Ube3A* (Albrecht *et al.*, 1997) und PX00010K13 (NCBI-Eintrag: AK014392,

(Nikaido *et al.*, 2003)) als bereits bekannte geprägte Gene identifiziert. AK034097, ein viertes Transkript, ist kleinerer Bestandteil von AK038900, einem weiteren Kandidatengen, und wurde daher nicht weiter berücksichtigt. Mit den verbliebenen 64 Kandidatengenen wurde die Studie fortgesetzt.

3.1.2. Identifizierung von transkribierten SNPs

Eine allelische Diskriminierung in der mRNA lässt sich am einfachsten durch die Verwendung eines heterozygoten transkribierten SNPs (tSNPs) als Marker erreichen. Es wurde daher zunächst in den Sequenzen der verbliebenen 64 Kandidatengene und den drei Kontrollgenen (*Zim1*, PX00010K13 (AK014392) und *Slc38a4*) nach tSNPs gesucht. Dazu wurde die transkribierte Region teilweise oder vollständig sequenziert. Um eine eindeutige Bestimmung von SNPs zu gewährleisten, wurden zwei DNA-Proben von F1-Hybriden (*Mus musculus* Subspezies C57BL/6 x Cast/Ei) sowie jeweils eine DNA von den C57BL/6- und Cast/Ei-Elterntieren für die Sequenzierung eingesetzt. Dabei konnten in 60 Kandidaten- sowie in den drei Kontrollgenen tSNPs identifiziert werden, die sich für eine QUASEP-Analyse eigneten (siehe Anhang Tab. 15). In vier Kandidatengenen konnte kein tSNP identifiziert werden. Die Eignung eines tSNPs für die QUASEP-Analyse wurde mit der Assay Design-Software von Pyrosequencing™ bestimmt. Dabei wurde vor allem darauf geachtet, dass der tSNP nicht zu einem Homopolymer zählt. Ob die Kandidatengene tatsächlich geprägt sind, wurde durch eine allelische Expressionsanalyse mittels QUASEP verifiziert.

3.1.3. Allelische Expressionsanalyse durch QUASEP

Um die genomische Prägung der verbliebenen 60 Kandidatengene wie auch der drei Kontrollgene zu verifizieren, wurden RT-PCR-Produkte von F1-Maushybriden auf allelspezifische Expression untersucht. Wenn mehr als ein tSNP pro Gen identifiziert werden konnte, wurde der am besten geeignete Polymorphismus anhand des höchsten von der Assay Design-Software ermittelten Gesamtwertes gewählt.

In der allelischen Expressionsanalyse der Kontrollgene zeigen *Zim1* und PX00010K13 (AK014392) die erwartete maternale bzw. paternale monoallelische Expression (siehe Abb. 11). Dies bestätigt die Eignung von QUASEP für die Analyse genomischer Prägung.

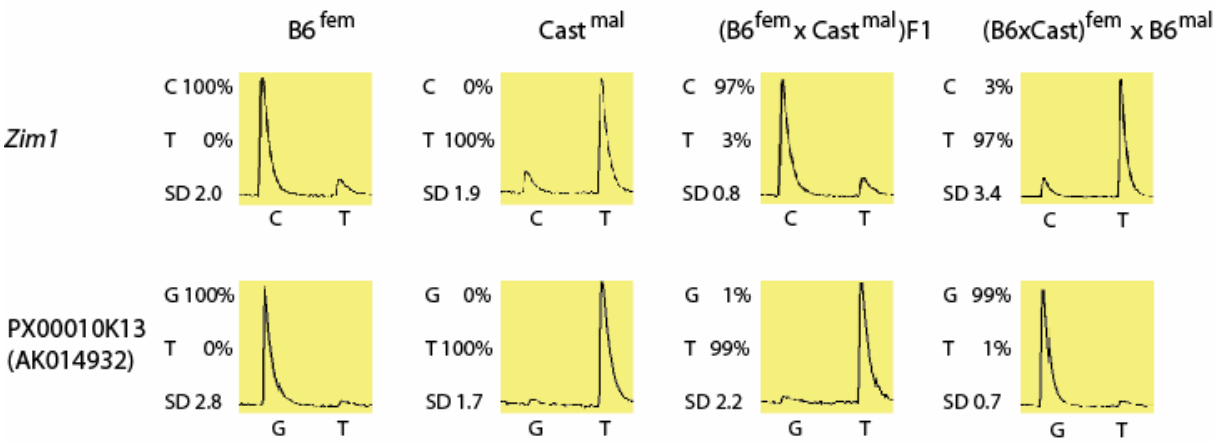

Abbildung 11. QUASEP-Ergebnisse, die die Eignung für die Imprinting-Analyse verdeutlichen. Die Ausschnitte aus den Pyrogrammen zeigen die homozygoten Elterntiere ($B6^{fem}$ für das C57BL/6-Muttertier und $Cast^{mal}$ für das Cast/Ei-Vatertier). Die F1-Hybriden ($B6^{fem}$x$Cast/Ei^{mal}$)F1 (E11.5) weisen wie erwartet maternale Expression für *Zim1* sowie paternale Expression für AK014392 auf. Die Analyse von informativen reziproken Hybriden ($(B6xCast)^{fem}xB6^{mal}$) ergab reziproke Ergebnisse. Das Verhältnis der beiden Allele wie auch die Standardabweichung (SD) sind jeweils links angegeben.

In einer weiteren Eignungsprüfung der QUASEP-Methode wurde die bekannte gewebespezifische Prägung von *Slc38a4* (auch als *Ata3* bezeichnet) bestätigt. Wie in Abbildung 12 dargestellt, stimmen die Ergebnisse mit den bisher publizierten Daten überein, die biallelische Expression in Leber und Testis sowie paternale Expression in Skelettmuskel, Herz, Milz und Niere belegen (Mizuno *et al.*, 2002; Smith *et al.*, 2003).

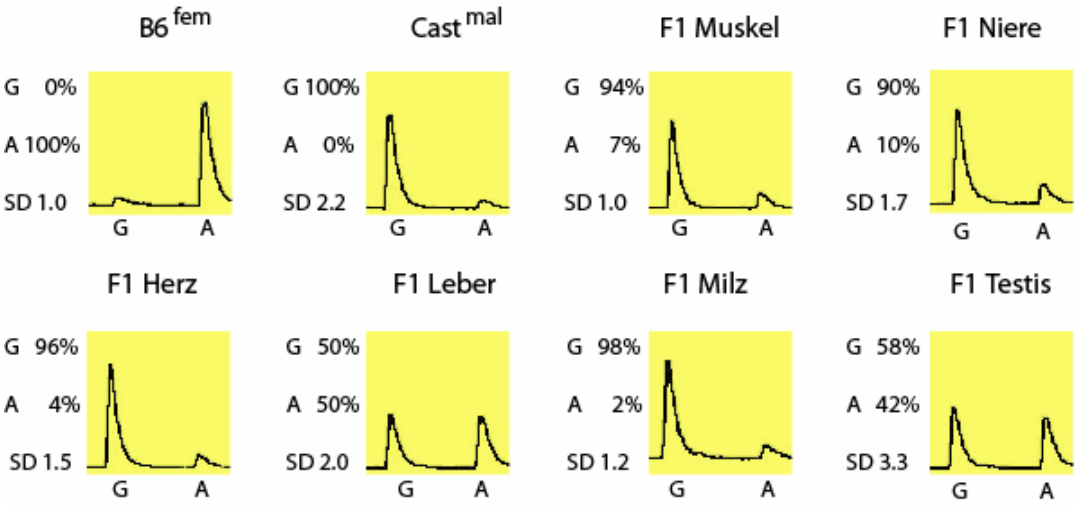

Abbildung 12. Gewebespezifische Prägung von *Slc38a4*. Die QUASEP-Analyse von RT-PCR-Produkten aus verschiedenen Geweben von F1-Maushybriden bestätigt die biallelische Expression in Leber und Testis sowie die paternale Expression in Skelettmuskel, Herz, Milz und Niere. Das Verhältnis der beiden Allele wie auch die Standardabweichung (SD) sind jeweils links angegeben.

Nach der Untersuchung aller 60 Kandidatengene in F1-Maushybriden (E11.5) konnten acht potentiell geprägte Transkripte identifiziert werden, wobei vier Transkripte monoallelische Expression (AK043056, AK039482, AK018142 und AK039514) und vier Transkripte präferentiell allelspezifische Expression (AK038367, AK043873, AK035722 und AK014850) aufweisen. Die elternspezifische Expression stimmte bis auf AK043873 mit den Vorhersagen der Microarray-Studie überein. Von den verbliebenen 52 Kandidatengenen weisen 47 Transkripte biallelische Expression auf, während fünf Transkripte nicht in E11.5-Embryonen exprimiert sind (siehe Tab. 5).

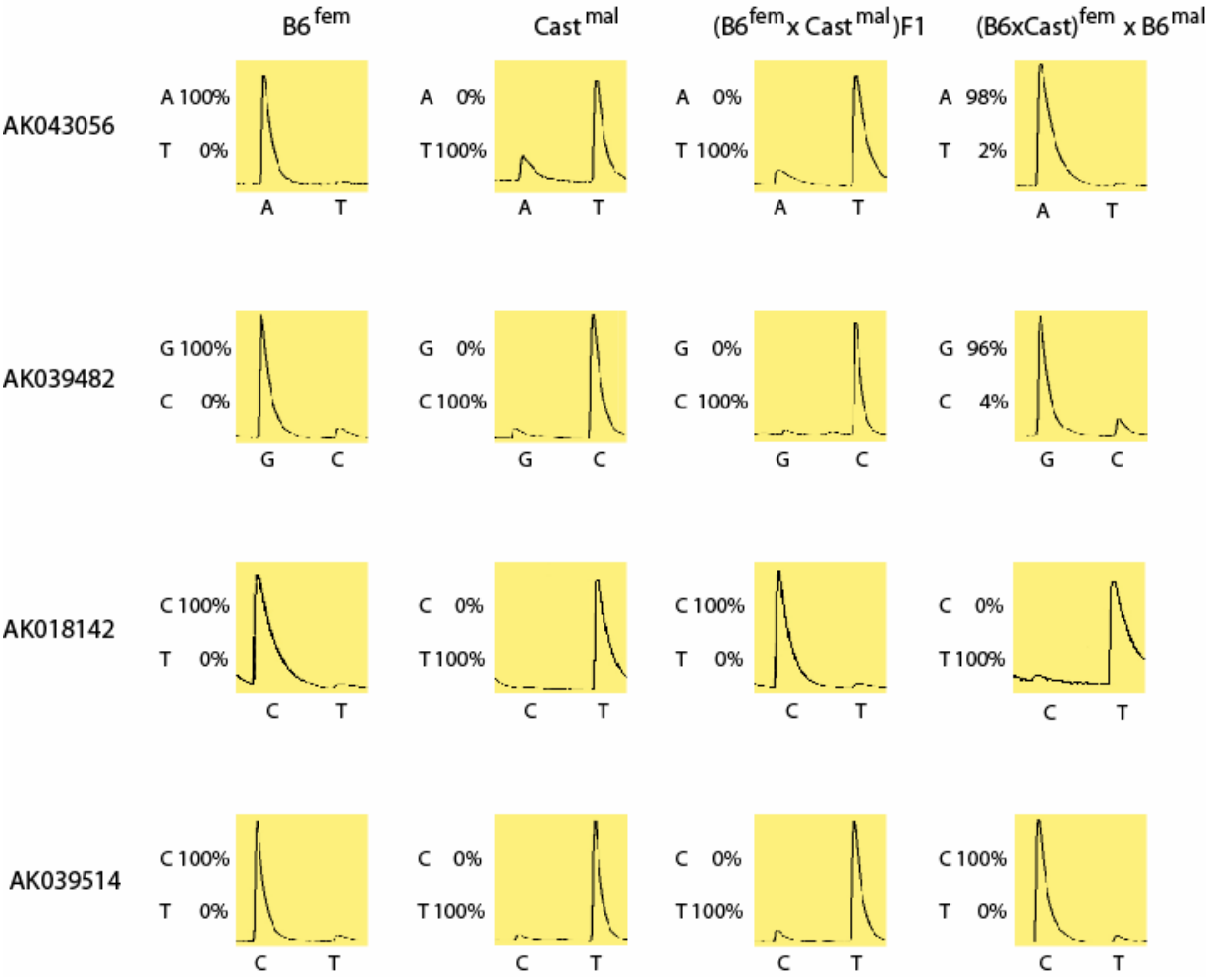

Abbildung 13. Identifizierung von vier geprägten Transkripten. Für jedes Gen sind Ausschnitte aus den Pyrogrammen von homozygoten Elterntieren ($B6^{fem}$ für das C57BL/6-Muttertier und $Cast^{mal}$ für das Cast/Ei-Vatertier) sowie von F1-Hybriden ($B6^{fem}$xCast/Ei^{mal})F1 (E11.5) und von informativen reziproken Hybriden ($(B6xCast)^{fem}$x$B6^{mal}$) dargestellt. Das Verhältnis der beiden Allele ist jeweils links in Prozent angegeben.

Um die potentielle Prägung dieser acht Transkripte zu bestätigen, wurden informative Mausembryonen einer reziproken Kreuzung mit QUASEP untersucht. In diesen Experimenten zeigten die vier Gene mit monoallelischer Expression ein reziprokes Ergebnis und bestätigen damit genomische Prägung (siehe Abb. 13). Diese Ergebnisse konnten durch

cDNA-Sequenzierung anhand von tSNPs bestätigt werden. Von diesen vier Transkripten weisen AK043056 und AK039482 paternale Expression auf und kartieren in die Imprinting-Regionen auf dem proximalen bzw. distalen Maus-Chromosom 7. Das maternal exprimierte Transkript AK018142 kartiert in die Imprinting-Region auf dem distalen Maus-Chromosom 12. Das Transkript AK039514 ist paternal exprimiert und in der Imprinting-Region auf dem proximalen Maus-Chromosom 17 lokalisiert.

Die QUASEP-Analyse der vier Transkripte mit präferentiell allelspezifischer Expression in informativen Mausembryonen der reziproken Kreuzung ergab das gleiche Ergebnis wie die Untersuchung der F1-Hybriden, so dass es sich hier um stammspezifische Effekte handelt. Bei den Transkripten AK038367, AK043873 und AK014850 wird das Cast/Ei-Allel, bei dem Transkript AK035722 hingegen das B6-Allel präferentiell exprimiert (siehe Abb. 14). Für eine Einteilung als präferentiell allelspezifische Expression wurde ein Grenzwert von mindestens 65% zu 35% festgelegt. Bei einem Verhältnis von kleiner als 65% zu 35% wurde die Expression als biallelisch bewertet.

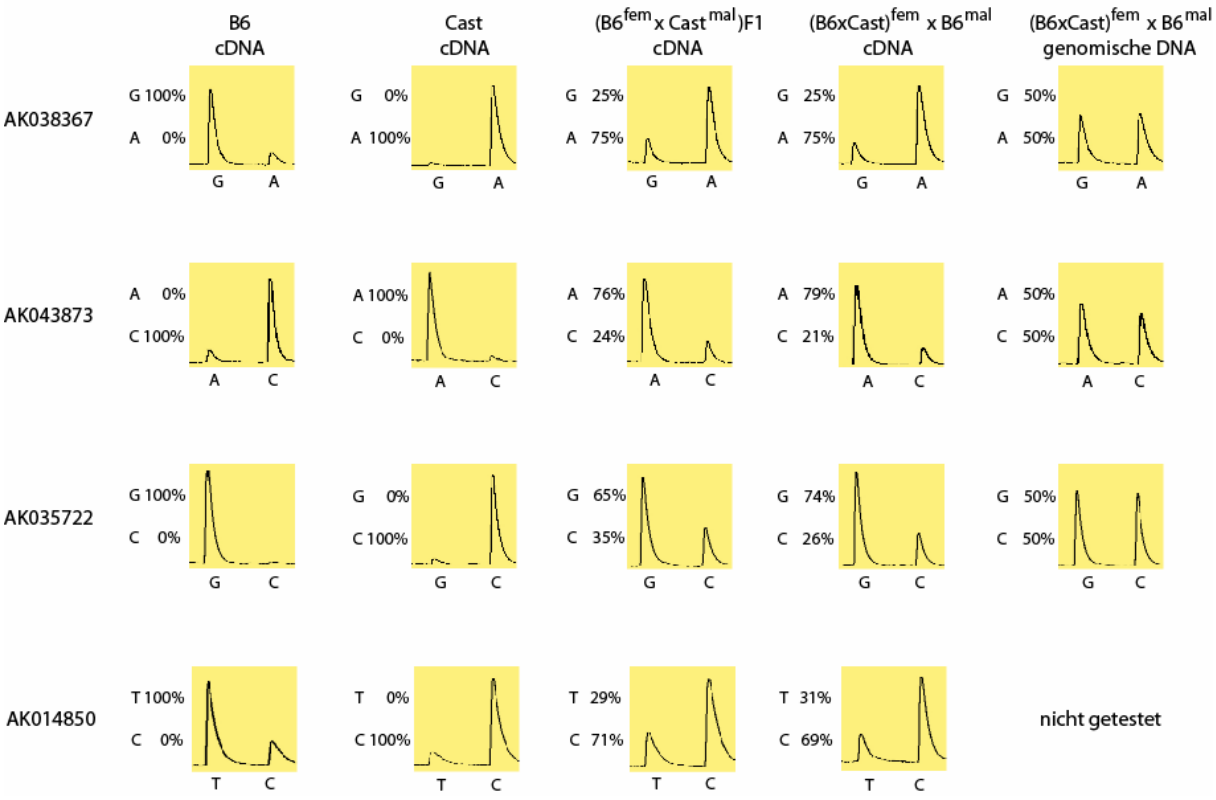

Abbildung 14. Identifizierung von vier stammspezifisch exprimierten Transkripten. Für jedes Transkript sind Ausschnitte aus den Pyrogrammen von homozygoten Elterntieren ($B6^{fem}$ für das C57BL/6-Muttertier und $Cast^{mal}$ für das Cast/Ei-Vatertier) sowie von F1-Hybriden ($B6^{fem}$xCast/Ei^{mal})F1 (E11.5) und von informativen reziproken Hybriden ((B6xCast)^{fem}x$B6^{mal}$) dargestellt. Das Verhältnis der beiden Allele ist jeweils links in Prozent angegeben. Bei den Transkripten AK038367, AK043873 und AK014850 wird das Cast/Ei-Allel, bei dem Transkript AK035722 hingegen das B6-Allel präferentiell exprimiert.

Tabelle 5. Ergebnisse der allelischen Expressionsanalyse der untersuchten Imprinting-Kandidatengene. Die vier geprägten sowie die vier präferentiell stammspezifisch exprimierten Transkripte sind fett hervorgehoben.

Genname (EICO)	Klon-ID (EICO)	GenBank-Eintrag	Chromosomale Position (bp)	Allelische Expression C57BL/6 versus Cast/Ei SD* = Standardabweichung
unclassifiable	B130039H18	AK045137	prox2 (4208431-4218375)	biallelisch (56% : 44%; SD 2,6)
HPRP18 (PRE-MRNA SPLICING FACTOR SIMILAR TO S, CEREVISIAE PRP18) homolog [Homo sapiens]	9430083N14	AK035073	prox2 (4541300-4569385)	biallelisch (50% : 50%; SD 3,2)
f-box only protein 3	4921535A12	AK019550	cen2 (103732615-103766155)	biallelisch (47% : 53%; SD 1,2)
integrin beta 4 binding protein	2700008E01	AK012218	dist2 (155276776-155283625)	biallelisch (48% : 52%; SD 2,2)
unknown EST	A730085E23	AK080554	dist2 (155629834-155631061)	biallelisch (46% : 54%; SD 1,0)
unclassifiable	**A130099D17**	**AK038367**	**dist2 (158157925-158161148)**	**präferentiell stammspezifisch Cast/Ei (25% : 75%; SD 7,6)**
MELANOCORTIN-3 RECEPTOR (MC3-R)	A230054D12	AK038677	dist2 (171757226-171759848)	biallelisch (41% : 59%; SD 11,6)
bone morphogenetic protein 7	9430012M17	AK020411	dist2 (172376586-172400560)	biallelisch (57% : 43%; SD 3,8)
RAE1 RNA export 1 homolog (S, pombe)	9630054N05	AK079383	dist2 (172460605-172476200)	biallelisch (46% : 54%; SD 6,4)
unknown EST	A530024C12	AK079954	dist2 (172571303-172592738)	biallelisch (45% : 55%; SD 5,6)
hypothetical Cysteine-rich region containing protein	4932434E15	AK016545	dist2 (172756001-172759452)	keine Expression
unclassifiable	A130023E15	AK037521	dist2 (173139921-173142243)	biallelisch (62% : 38%; SD 3,8)
unknown EST	9030418K01	AK018518	dist2 (178139527-178141383)	biallelisch (55% : 45%; SD 1,2)
unknown EST	A130029E23	AK037602	dist2 (179248156-179251604)	biallelisch (49% : 51%; SD 3,2)
nuclear receptor binding factor 1	B230313J17	AK045834	dist4 (130640364-130641437)	biallelisch (57% : 43%; SD 2,8)
unknown EST	A630096I18	AK042496	dist4 (130676252-130677382)	biallelisch (42% : 58%; SD 0,9)
unknown EST	9430023J17	AK034678	dist4 (130720315-130722049)	biallelisch (51% : 49%; SD 1,6)
DELTA-TYPE OPIOID RECEPTOR	**A830044L10**	**AK043873**	**dist4 (130907612-130941372)**	**präferentiell stammspezifisch Cast/Ei (24% : 76%; SD 3,9)**
unknown EST	**9530096D07**	**AK035722**	**dist4 (131097326-131100107)**	**präferentiell stammspezifisch C57BL/6 (65% : 35%; SD 8,9)**
hypothetical Type-1 copper (blue) domain/Leucine-rich region containing protein	4931440N07	AK029920	dist4 (131685804-131693446)	biallelisch (54% : 46%; SD 2,1)
similar to cDNA FLJ13976 FIS, CLONE Y79AA1001594, WEAKLY SIMILAR TO HYALURONAN-MEDIATED MOTILITY RECEPTOR [Homo sapiens]	5730495A12	AK030742	dist4 (133090759-133147950)	biallelisch (53% : 47%; SD 11,3)
similar to cDNA FLJ10747 FIS, CLONE NT2RP3001799 [Homo sapiens]	A230080D12	AK038972	dist4 (133789307-133814142)	kein tSNP
unclassifiable	9530059D22	AK035517	dist4 (134376144-134378373)	kein tSNP
inferred: Mus musculus von Willebrand factor A-related protein (Warp) mRNA, complete cds	9530003H10	AK035229	dist4 (154262332-154268524)	biallelisch (59% : 41%; SD 0,8)
stromal cell derived factor 4	3830404L10	AK014418	dist4 (154485753-154502681)	biallelisch (44% : 56%; SD 3,0)
unknown EST	A230094E16	AK039087	prox6 (5117371-5118547)	biallelisch (52% : 48%; SD 0,8)
pyruvat dehydrogenase 4	9930014N18	AK036818	prox6 (5456108-5469066)	biallelisch (49% : 51%; SD 3,3)
unknown EST	A130084J17	AK038180	prox6 (29847086-29847770)	biallelisch (49% : 51%; SD 2,9)
hypothetical protein	9430029A11	AK034724	prox6 (32531158-32534738)	biallelisch (63% : 37%; SD 6,6)
unclassifiable	A230072E17	AK038886	prox7 (5540239-5543861)	in Cast/Ei nicht exprimiert
ZFP77P (FRAGMENT) homolog [Mus musculus]	A330079D15	AK039658	prox7 (5559289-5575671)	biallelisch (50% : 50%; SD 5,0)
zinc finger protein 28	2700045L05	AK028202	prox7 (5579109-5590975)	biallelisch (54% : 46%; SD 2,3)
unclassifiable	**A730051H19**	**AK043056**	**prox7 (5938638-5941248)**	**monoallelisch paternal (0% : 100%; SD 4,2)**
unclassifiable	9530064F02	AK079260	cen7 (52998444-53001134)	biallelisch (44% : 56%; SD 7,2)

Fortsetzung Tabelle 5

Genname (EICO)	Klon-ID (EICO)	Accession (GenBank)	Chromosomale Position (bp)	Allelische Expression C57BL/6 versus Cast/Ei SD* = Standardabweichung
unclassifiable	A130033P09	AK037653	cen7 (53480636-53482047)	biallelisch (52% : 48%; SD 10,3)
unknown EST	A230073K19	AK038900	cen7 (53564185-53567299)	kein tSNP
unknown EST	9330162G02	AK034189	cen7 (53524491-53527482)	biallelisch (45% : 55%; SD 1,4)
unclassifiable	A230006K03	AK038418	cen7 (54865859-54869149)	keine Expression
phosphatidylserine synthase 2	9630061G10	AK036368	dist7 (135533274-135558142)	biallelisch (58% : 42%; SD 2,3)
hypothetical Leucine-rich repeat containing protein	5730427C23	AK017603	dist7 (135596145-135612043)	biallelisch (45% : 55%; SD 1,0)
GL008 homolog [Homo sapiens]	9830166G06	AK036717	dist7 (135896130-135942579)	biallelisch (60% : 40%; SD 2,2)
similar to MUCIN (FRAGMENT) [Homo sapiens]	A130042M24	AK037730	dist7 (136251868-136270545)	keine Expression
cathepsin D	F730002E02	AK089294	dist7 (136788942-136800852)	biallelisch (57% : 43%; SD 1,6)
unclassifiable	9430050L03	AK034864	dist7 (137446304-137466697)	biallelisch (55% : 45%; SD 5,0)
weakly similar to ORF 2 PROTEIN (FRAGMENT) [Rattus norvegicus]	**A330049H05**	**AK039482**	**dist7 (137651381-137653568)**	**monoallelisch paternal (0% : 100%; SD 2,2)**
fibroblast growth factor 15	5730551C10	AK017829	dist7 (139308455-139312873)	biallelisch (48% : 52%; SD 2,9)
unclassifiable	9430062P05	AK034922	cen9 (90043752-90047093)	biallelisch (48% : 52%; SD 4,0)
unknown EST	A130002D10	AK037266	prox10 (11082191-11084486)	biallelisch (41% : 59%; SD 5,2)
hypothetical protein	A130098G01	AK038360	prox10 (12657455-12660035)	biallelisch (43% : 57%; SD 2,2)
hypothetical protein	4921519G13	AK029538	prox11 (23459752-23528340)	biallelisch (46% : 54%; SD 1,5)
unclassifiable	A330045P15	AK039463	dist12 (104441659-104442287)	biallelisch (50% : 50%; SD 1,6)
unknown EST	6330404N04	AK018121	dist12 (105038425-105039910)	kein tSNP
unknown EST	**6330408G06**	**AK018142**	**dist12 (105040258-105041469)**	**monoallelisch maternal (100% : 0%; SD 4,6)**
unknown EST	A430099D21	AK040459	dist12 (106056529-106058296)	biallelisch (54% : 46%; SD 0,7)
similar to X-RAY REPAIR CROSS-COMPLEMENTING PROTEIN 3 (X-RAY REPAIR COMPLEMENTING DEFECTIVE REPAIR IN CHINESE HAMSTER CELLS 3) [Homo sapiens]	4432412E01	AK014491	dist12 (107284534-107293690)	biallelisch (54% : 46%; SD 2,0)
unknown EST	A130078K24	AK038100	cen14 (67051563-67055101)	biallelisch (56% : 44%; SD 7,0)
unclassifiable	A130030G06	AK037626	cen14 (67697954-67700748)	biallelisch (41% : 59%; SD 3,6)
hypothetical Leucine-rich repeat/bZIP (Basic-leucine zipper) transcription factor family/Leucine-rich repeat, typical subtype containing protein	**4921509B22**	**AK014850**	**cen14 (69428154-69470413)**	**präferentiell stammspezifisch Cast/Ei (29% : 71%; SD 5,2)**
unknown EST	A230009B12	AK038428	prox17 (9096279-9113813)	keine Expression
unclassifiable	**A330053J22**	**AK039514**	**prox17 (11416402-11417667)**	**monoallelisch paternal (4% : 96%; SD 2,3)**
unknown EST	A830092C01	AK044118	prox17 (11820831-11822518)	biallelisch (50% : 50%; SD 2,2)
unknown EST	A430047I21	AK040032	prox17 (12969213-12971928)	biallelisch (59% : 41%; SD 7,8)
programmed cell death 2	B230330P13	AK045985	prox17 (13525238-13530954)	biallelisch (42% : 58%; SD 3,0)
Niemann Pick type C1	A430089E03	AK040367	cen18 (12407137-12423960)	biallelisch (45% : 55%; SD 4,9)

Um PCR-bedingte Abweichungen auszuschließen, wurde die QUASEP-Analyse auch mit PCR-Produkten von genomischer DNA wiederholt. Dabei konnten für AK038367, AK043873 und AK035722 eine eindeutige 50:50-Verteilung der beiden Allele nachgewiesen werden und somit ein durch PCR und/oder QUASEP bedingtes falsch positives Ergebnis ausgeschlossen werden (siehe Abb. 14 rechts). Da für AK014850 der Versuch mit genomischer DNA aufgrund der Produktgröße nicht möglich war, wurde ein zweiter QUASEP-Versuch mit einem weiteren tSNP und anderen Primern durchgeführt, der das

Ergebnis des ersten Versuchs bestätigte (Daten nicht gezeigt). Die vier Transkripte mit stammspezifischer Expression sind in den bekannten geprägten Regionen auf dem distalen Maus-Chromosom 2 (AK038367), dem distalen Maus-Chromosom 4 (AK043873 und AK35722) sowie dem zentralen Maus-Chromosom 14 lokalisiert.

3.1.4. Erhalt der Prägung im adulten Stadium

Des Weiteren wurde die gewebespezifische Expression und genomische Prägung von AK043056, AK039482, AK018142 und AK039514 im adulten Stadium durch eine QUASEP-Analyse von RT-PCR-Produkten aus Gehirn, Herz, Niere, Leber, Lunge, Skelettmuskel und Milz untersucht. Dabei wurde vorausgesetzt, dass sich die Expression mittels RT-PCR in den jeweiligen Proben stabil reproduzieren lässt.

Das Transkript AK039482 wies dabei eine Relaxation der genomischen Prägung, jedoch weiterhin präferentiell paternale Expression in allen untersuchten Geweben auf. Die übrigen drei Transkripte offenbarten weiterhin strikt monoallelische Expression in den untersuchten Geweben (siehe Tab. 6). Alle Ergebnisse konnten in Gewebeproben der reziproken Kreuzung bestätigt werden.

Tabelle 6. Geprägte Expression in verschiedenen adulten Geweben von F1-Maushybriden.

GenBank-Eintrag	Gewebe	Verhältnis allelischer Expression (C57BL/6 versus Cast/Ei) SD = Standardabweichung	
AK043056	Gehirn	monoallelisch paternal	(0% : 100%; SD 1,0)
	Skelettmuskel	monoallelisch paternal	(0% : 100%; SD 1,7)
	Milz	monoallelisch paternal	(0% : 100%; SD 0,6)
AK039482	Gehirn	präferentiell paternal	(19% : 81%; SD 1,5)
	Herz	präferentiell paternal	(26% : 74%; SD 4,7)
	Niere	präferentiell paternal	(30% : 70%; SD 3,4)
	Leber	präferentiell paternal	(20% : 80%; SD 4,0)
	Lunge	präferentiell paternal	(37% : 63%; SD 4,2)
	Skelettmuskel	präferentiell paternal	(23% : 77%; SD 1,0)
	Milz	präferentiell paternal	(29% : 71%; SD 2,5)
AK018142	Gehirn	monoallelisch maternal	(100% : 0%; SD 0,6)
	Herz	monoallelisch maternal	(100% : 0%; SD 8,3)
	Niere	monoallelisch maternal	(100% : 0%; SD 0,2)
	Leber	monoallelisch maternal	(100% : 0%; SD 6,8)
	Lunge	monoallelisch maternal	(100% : 0%; SD 6,1)
	Skelettmuskel	monoallelisch maternal	(100% : 0%; SD 1,6)
	Milz	monoallelisch maternal	(100% : 0%; SD 0,4)
AK039512	Gehirn	monoallelisch paternal	(12% : 88%; SD 2,8)
	Herz	monoallelisch paternal	(10% : 90%; SD 1,3)
	Niere	monoallelisch paternal	(10% : 90%; SD 2,4)
	Leber	monoallelisch paternal	(9% : 91%; SD 4,9)
	Lunge	monoallelisch paternal	(5% : 95%; SD 1,5)
	Skelettmuskel	monoallelisch paternal	(10% : 90%; SD 3,6)
	Milz	monoallelisch paternal	(1% : 99%; SD 2,7)

3.1.5. Genomische Organisation der geprägten Transkripte

Im nächsten Schritt wurde die genomische Position bestimmt und die benachbarten Regionen der vier neuen geprägten Transkripte analysiert. In der EICO-Datenbank sind alle vier Transkripte als vollständige cDNA-Klone angegeben. Der Sequenzvergleich gegen das Mausgenom mit dem Programm BLAT Search ergab, dass alle vier Transkripte aus nur einem Exon bestehen.

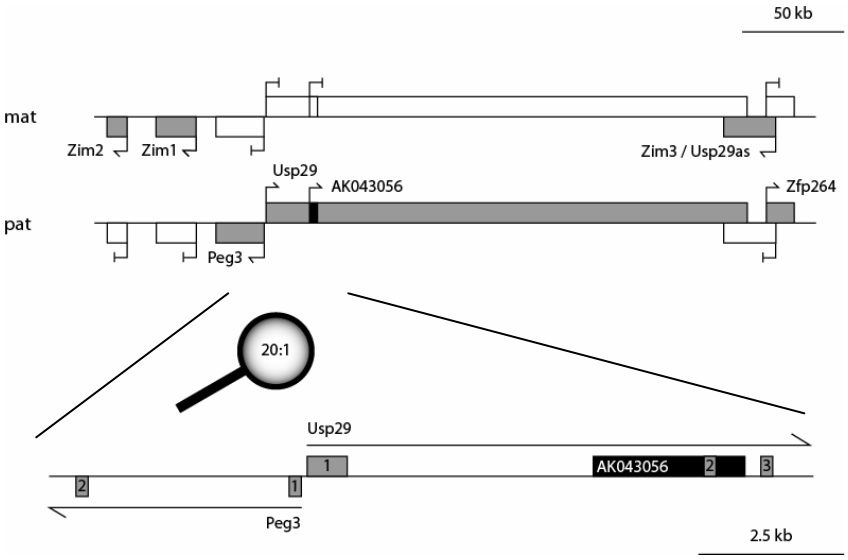

Abbildung 15. Genomische Karte vom proximalen Maus-Chromosom 7 mit den bekannten geprägten Genen. Die untere Vergrößerung (20fach) zeigt die genomische Struktur von AK043056, *Usp29, Peg3* unter Berücksichtigung der Exons (nummerierte graue Rechtecke) und Introns von *Usp29* und *Peg3.* Die Pfeile kennzeichnen die Transkriptionsrichtung jedes Gens. Das aktive Allel von AK043056 wird durch ein schwarzes Rechteck dargestellt. Die grauen Rechtecke stellen die aktiven Allele der bekannten geprägten Gene dar. Inaktive Allele werden durch weiße Rechtecke dargestellt. Mat steht für das maternale und pat für das paternale Chromosom.

Das paternal exprimierte Transkript AK043056 kartiert in die Region auf MMU7prox und liegt innerhalb der Transkriptionseinheit von *Usp29* (siehe Abb. 15). Neben dem ebenfalls paternal exprimierten Gen *Usp29* (Kim *et al.*, 2000b) sind in diesem Bereich noch fünf weitere geprägte Gene bekannt. Dazu zählen die paternal exprimierten Gene *Peg3* (Kuroiwa *et al.*, 1996) und *Zfp264* (Kim *et al.*, 2001) sowie die maternal exprimierten Gene *Zim1* (Kim *et al.*, 1999), *Zim2* (Kim *et al.*, 2004b) und *Zim3* (Kim *et al.*, 2001). Die Sequenz von AK043056 stimmt vollständig mit einem Bereich beginnend am 3´-Ende des Intron 1 über das Exon 2 bis in das 5´-Ende des Intron 2 von *Usp29* überein (siehe Abb. 15). Darüber hinaus besteht keine weitere signifikante Ähnlichkeit mit anderen Exons von *Usp29.*

Der Vergleich der mRNA-Sequenz von *Usp29* (NCBI-Eintrag: NM_021323.1) mit der chromosomalen Karte von Celera ergibt, dass neun Exons über 230 kb verteilt vorliegen

(siehe Tab. 7). Die Anzahl der Exons steht damit im Widerspruch zu den Daten von Kim *et al.* (2000b) mit fünf Exons.

Tabelle 7. Position und Größe der Exons und Intron von *Usp29* (RefSeq: NM_021323.1).

Exon	Position	Länge (bp)	Intron	Länge (bp)
1	1-720	720	1	6111
2	721-790	70	2	1158
3	791-943	153	3	169853
4	944-1035	92	4	19802
5	1036-1105	70	5	777
6	1106-1189	84	6	19554
7	1190-1265	76	7	7711
8	1266-1452	187	8	944
9	1453-7550	6098		

Das ebenfalls paternal exprimierte AK039482 ist in der Imprinting-Region auf MMU7dist lokalisiert. Dieser Bereich wie auch die orthologe Region auf dem humanen Chromosom 11p15.5 wurden bereits intensiv erforscht und beherbergen 12 maternal exprimierte Gene (*H19*, *Ascl2*, *Cd81*, *Tssc4*, *Kcnq1*, *Cdkn1c*, *Msuit1*, *Slc22a1l*, *Phlda2*, *Nap1l4*, *Tnfrsf23* und *Osbpl5*) sowie drei paternal exprimierte Gene (*Kcnq1ot1(Lit1)*, *Ins2* und *Igf2*) (Onyango *et al.*, 2000; Paulsen *et al.*, 2000). AK039482 kartiert ca. 16,5 kb stromabwärts von *Lit1* im Intron 10 von *Kcnq1* (siehe Abb. 16). Das paternal exprimierte *Lit1*-Transkript ist eine aus nur einem Exon bestehende nicht-kodierende RNA und wird aktuell mit einer Länge von ca. 84 kb angegeben (NCBI: NR_001461). In einer Studie wurde jedoch eine Verlängerung des Transkripts über die Position von AK39482 hinaus belegt (Yatsuki *et al.*, 2000). Diese Verlängerung wird von einer Vielzahl sich überlappender ESTs (*expressed sequence tags*) validiert. Daher ist es sehr wahrscheinlich, dass AK039482 zu *Lit1* gehört.

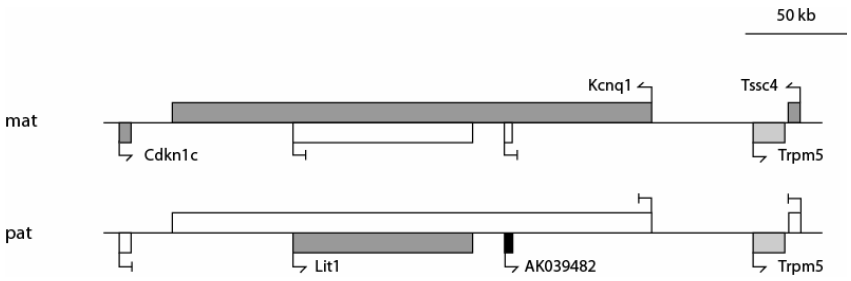

Abbildung 16. Genomische Karte vom distalen Maus-Chromosom 7 mit AK039482 und vier benachbarten geprägten Genen. Die Pfeile kennzeichnen die Transkriptionsrichtung jedes Gens. Das aktive Allel von AK039482 wird durch ein schwarzes Rechteck dargestellt. Die grauen Rechtecke stellen die aktiven Allele der bekannten geprägten Gene dar. Inaktive Allele werden durch weiße Rechtecke dargestellt. Hellgraue Rechtecke markieren biallelisch exprimierte Gene. Mat steht für das maternale und pat für das paternale Chromosom.

Das maternal exprimierte AK018142 kartiert 340 bp stromabwärts von *Gtl2* in der geprägten Region auf MMU12dist (siehe Abb. 17). In diesem Bereich sind außerdem die paternal exprimierten Gene *Dlk1, Rtl1, Dio3* sowie die maternal exprimierten Gene *Gtl2,*

Rian, Anti-Rtl1, Mirg und sechs C/D snoRNA kodierende Gene lokalisiert (Miyoshi *et al.*, 2000; Schmidt *et al.*, 2000; Takada *et al.*, 2000). Ein ca. 10 kb großer Bereich am 3´-Ende von *Gtl2* mit einer Vielzahl sich überlappender ESTs weist darauf hin, dass das Transkript von *Gtl2* wahrscheinlich länger ist als die bisher angegebenen 4326 bp und somit AK018142 lediglich zur 3´-UTR von *Gtl2* gehören könnte.

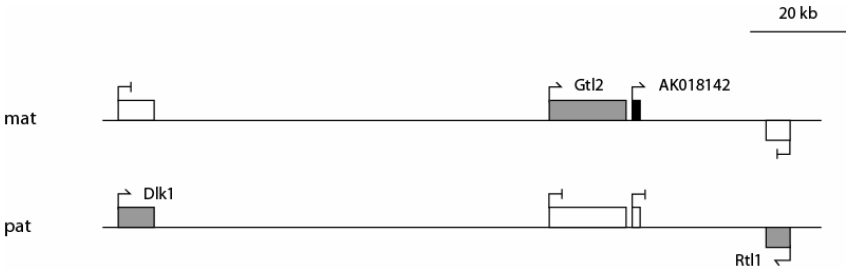

Abbildung 17. Genomische Karte vom distalen Maus-Chromosom 12 mit AK018142 sowie den bekannten geprägten Genen *Dlk1*, *Gtl2* und *Rtl1*. Die Pfeile kennzeichnen die Transkriptionsrichtung jedes Gens. Das aktive Allel von AK039482 wird durch ein schwarzes Rechteck dargestellt. Die grauen Rechtecke stellen die aktiven Allele der bekannten geprägten Gene dar. Inaktive Allele werden durch weiße Rechtecke dargestellt. Mat steht für das maternale und pat für das paternale Chromosom.

Das auf MMU17prox kartierte und ebenfalls paternal exprimierte Transkript AK039514 wurde durch einen Sequenzvergleich als Teil der Transkriptionseinheit des geprägten Gens *Air* identifiziert (Genbank: AJ249895.1 (Lyle *et al.*, 2000)). Wie in Abbildung 18 dargestellt, überlappt *Air* das Gen *Igf2r* in gegensätzlicher Richtung und wirkt als bidirektionaler Silencer für eine ca. 400 kb große Imprinting-Region mit den zwei weiteren geprägten Genen *Slc22a2* und *Slc22a3* (Sleutels *et al.*, 2002; Sleutels *et al.*, 2003).

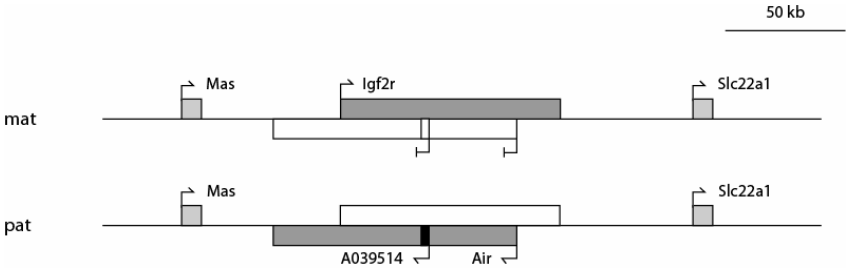

Abbildung 18. Genomische Karte vom proximalen Maus-Chromosom 17 mit AK039514 als Teil von *Air* und dem bekannten geprägten Gen *Igf2r*. Die Pfeile kennzeichnen die Transkriptionsrichtung jedes Gens. Das aktive Allel von AK039482 wird durch ein schwarzes Rechteck dargestellt. Die grauen Rechtecke stellen die aktiven Allele der bekannten geprägten Gene dar. Inaktive Allele werden durch weiße Rechtecke dargestellt. Hellgraue Rechtecke markieren biallelisch exprimierte Gene. Mat steht für das maternale und pat für das paternale Chromosom.

3.1.6. Weitere Charakterisierung von AK043056

Nachdem die genomische Lage der vier identifizierten geprägten Transkripte festgestellt wurde, galt es experimentell zu klären, ob AK043056, bedingt durch die besondere Position, ein eigenständiges Transkript ist und somit möglicherweise eine Rolle bei der Imprinting-Regulation auf MMU7prox spielen könnte. Zu diesem Zweck wurde zunächst die Expression mittels RT-PCR in neun verschiedenen Geweben geprüft (siehe Abb. 19). Jeder Ansatz wurde dabei fünfmal wiederholt und ergab, dass *Usp29* in Gehirn und Testis exprimiert ist. Diese Daten stimmen mit den Angaben in der Literatur überein (Kim *et al.*, 2000b). Darüber hinaus ergab die RT-PCR aus Plazenta eine reproduzierbare Population unterschiedlich großer PCR-Produkte, deren Größen jedoch nicht mit der erwarteten Produktgröße (373 bp, Primernummer 277) übereinstimmten.

Für AK043056 konnte eine deutliche Expression in Gehirn, Testis und Plazenta sowie schwache aber nicht kontinuierlich reproduzierbare Expression in Muskel, Niere, Herz und Milz festgestellt werden (118 bp, Primernummer 215).

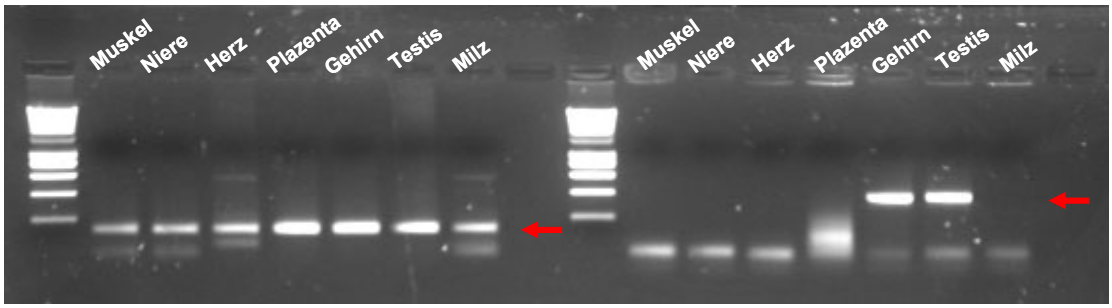

Abbildung 19. Expressionsvergleich von AK043056 (links) und *Usp29* (rechts). Der rote Pfeil markiert das jeweilige RT-PCR-Produkt (118 bp für AK043056 und 373 bp für *Usp29*) dargestellt in einem Agarose-Gel.

Da die unterschiedliche Größe der RT-PCR-Produkte einerseits ein falsch positives Ergebnis bedingen kann und andererseits durch die RT-PCR keine eindeutigen Expressionsunterschiede oder -gemeinsamkeiten belegt werden konnten, sollte im nächsten Schritt per Northern-Blot-Hybridisierung die gewebespezifische Expression von AK043056 untersucht werden. Dieses Experiment war jedoch nicht erfolgreich. Deshalb wurde in weiteren Versuchen mittels TaqMan® Gene Expression nach Unterschieden bzw. Gemeinsamkeiten zwischen AK043056 und *Usp29* gesucht.

Diese Experimente ergaben zum einen eine Expression von *Usp29* und AK043056 in Plazenta, Gehirn und Testis sowie eine tendenziell ähnliche Expressionshöhe, die auf eine

gemeinsame Regulation hinweisen. Zum anderen zeigte sich, dass *Usp29* in Plazenta niedriger exprimiert ist als AK043056, während *Usp29* in Gehirn und Testis höher exprimiert ist als AK043056 (siehe Tab. 8).

Tabelle 8. C_T-Werte der Untersuchung mit TaqMan® Gene Expression. SD steht für Standardabweichung.

Gewebe	C_T-Wert AK043056 (SD)	C_T-Wert *Usp29* (SD)	C_T-Wert *β-Actin* (SD)
Plazenta	31,6 (0,11)	32,6 (0,53)	19,0 (0,12)
Gehirn	27,5 (0,09)	25,3 (0,25)	21,8 (0,15)
Testis	34,6 (0,44)	33,0 (0,41)	19,7 (0,15)

In nachfolgenden RACE-Experimenten konnte die ursprünglich bekannte Länge von AK043056 von 2612 bp auf 2979 bp in 5´-Richtung verlängert werden. Das 3´-Ende bleibt unbekannt. Dieser Bereich ist besonders durch repetitive Sequenzen gekennzeichnet, und mehrere Versuche einer 3´-RACE erbrachten lediglich unspezifische Produkte. Die Bezeichnung AK043056 ist für das mittlerweile auf 2979 bp verlängerte paternal exprimierte Transkript nicht eindeutig, daher wird dieses Gen im Folgenden entsprechend der üblichen Nomenklatur als *Peg14* (für *Paternally expressed gene 14*) benannt.

3.1.7. Identifizierung des humanen Orthologs von *Peg14*

Vergleichende Kartierungsdaten belegen, dass die meisten menschlichen Orthologe der geprägten Gene auf MMU7prox in der homologen Region auf dem humanen Chromosom 19q13.4 lokalisiert sind. Darüber hinaus ist bereits die paternale Expression von *PEG3* und *ZIM2* belegt (Kim *et al.*, 1997; Kim *et al.*, 2000a). Eine Ensembl-BLAST-Suche zum Vergleich der Sequenz von AK043056 mit dem Humangenom ergab eine 513 bp große Region mit einer Übereinstimmung von 66% auf HSA19q13.4 (Sequenzvergleich siehe Anhang Kapitel 7.5). Durch weitere Vergleiche gegen das Genom von Ratte, Rind, Huhn, Fruchtfliege und Kugelfisch konnte die orthologe Region zwischen diesen Spezies auf 127 bp eingegrenzt werden. Mit einem Primerpaar (Primernummer 239), welches innerhalb dieser Region bindet, konnte in RT-PCR-Experimenten ein Transkript in RNA-Proben aus fetalem Gehirn nachgewiesen werden. Durch überlappende RT-PCR und RACE ergab sich eine Transkriptlänge von 3940 bp für das humane Ortholog.

Des Weiteren wurde die Expression des humanen Orthologs von *Peg14* mittels RT-PCR in fetalem Gehirn, fetalem und adultem Testis, Lymphoblastoiden, Amniozyten sowie fetalen und adulten Fibroblasten untersucht und eine Expression in Gehirn und Testis nachgewiesen. Der Nachweis des humanen Orthologs per Northern-Blot-Hybridisierung blieb trotz mehrfacher Versuche ohne Erfolg.

3.1.8. Monoallelische Expression des humanen Orthologs

Das humane Ortholog sollte in einem weiteren Schritt durch einen QUASEP-Versuch mit informativen Individuen auf allelische Expression überprüft werden. Zunächst wurden in elf Proben aus fetalem Gehirn, sechs Proben aus adultem Testis und in zehn Proben aus fetalem Testis nach geeigneten heterozygoten tSNPs mittels Sequenzierung gesucht. In den sequenzierten Fragmenten wurden ein A/C-Polymorphismus in heterozygoter Form bei 7 Individuen und ein Insertions-/Deletions-Polymorphismus (indelA) bei 8 Individuen (jeweils Gehirn und Testis zusammen) identifiziert. Für beide Polymorphismen konnte ein QUASEP-Versuch erfolgreich etabliert werden. Die darauf folgende Untersuchung ergab für die heterozygoten RT-PCR-Produkte (A/C-SNP) monoallelische Expression in fetalem Gehirn und biallelische Expression in fetalem und adultem Testis (siehe Abb. 20). Da sich bei den beiden Allelen der elterliche Ursprung aufgrund nicht vorliegender elterlicher DNA nicht ermitteln lässt, kann lediglich monoallelische Expression, jedoch keine Prägung im Sinne von paternaler oder maternaler Expression, festgestellt werden. Das Ergebnis des QUASEP-Versuchs mit dem indelA-Polymorphismus bestätigte die monoallelische Expression im Gehirn und biallelische Expression in Testis (Daten nicht gezeigt). Unter der Annahme einer paternalen Expression des humanen Orthologs von *Peg14* wird dieses Gen im Folgenden als *PEG14* benannt.

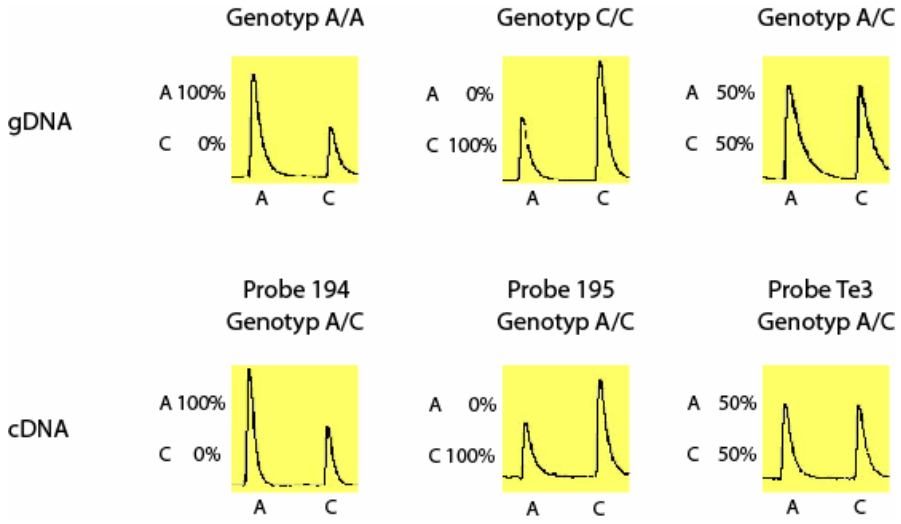

Abbildung 20. Allelische Expressionsanalyse für das humane *PEG14* in Gehirn und Testis. Oben dargestellt sind drei Ausschnitte aus den Pyrogrammen mit den drei möglichen Genotypen für den A/C-SNP. Analysiert wurden PCR-Produkte von genomischer DNA (gDNA). Unten dargestellt sind die monoallelische Expression in fetalem Gehirn anhand der Individuen 194 und 195 sowie biallelische Expression in Testis anhand des Individuums Te3. Wie die Sequenz GTTCA/CATATA zeigt, befinden sich unmittelbar links bzw. rechts neben dem A/C-SNP ein Cytosin bzw. ein Adenin. Da sich Signale von mehreren gleichen benachbarten Nukleotiden im Pyrogramm addieren und jeweils ein Signal ergeben, setzen sich hier die Signale aus den Nukleotiden an der polymorphen Position und den beiden benachbarten Nukleotiden zusammen. Über die Signalhöhe ist eine eindeutige Bestimmung des Genotyps möglich. Das Verhältnis der beiden Allele ist links in Prozent angegeben.

3.1.9. Vergleich der orthologen Transkripte

Mit der bekannten Transkriptlänge des humanen *PEG14* und dem bekannten 5´-Ende des murinen *Peg14* lässt sich ein genaueres Bild von der genomischen Lokalisation gewinnen sowie auch eine erste Bewertung hinsichtlich eines Genprodukts vornehmen.

Beide Transkripte befinden sich in einer Kopf-an-Kopf-Orientierung am 5´-Ende von *PEG3* bzw. *Peg3* und bestehen wahrscheinlich nur aus einem Exon (siehe Abb. 21). Das 5´-Ende von *Peg14* liegt 4756 bp von *Peg3* entfernt. Das humane *PEG14* beginnt 3920 bp von *PEG3* und enthält im Gegensatz zum Maus-Transkript keine *USP29*-Sequenz, da beim humanen *USP29* im Vergleich zur Maus die ersten fünf nicht-kodierenden Exons fehlen. Der längste Leserahmen von *Peg14* ist 255 bp lang (entspricht 85 Aminosäuren) und zeigt keine signifikante Homologie (6% mit Clustal W-Algorithmus) zum längsten Leserahmen des humanen Orthologs mit 339 bp (entspricht 113 Aminosäuren). Ein Vergleich der Nukleotidsequenzen mit dem Clustal W-Algorithmus ergibt eine Übereinstimmung von 34%.

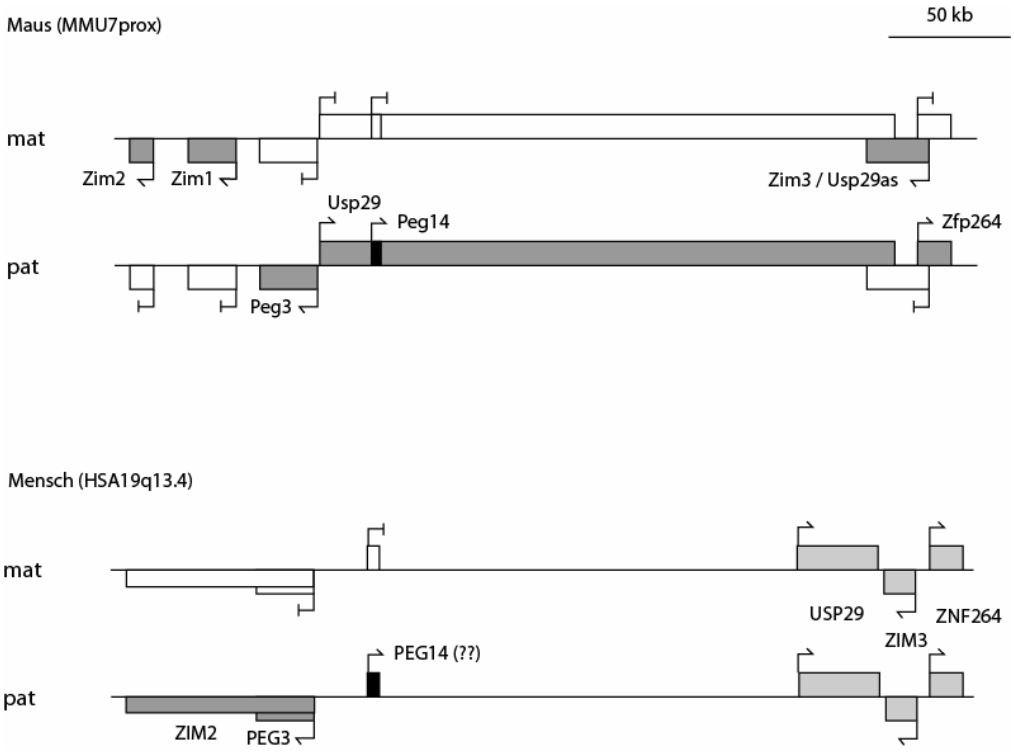

Abbildung 21. Genomische Karten von MMU7prox (oben) und HSA19q13.4 (unten) mit *Peg14* und den benachbarten Genen. Die Pfeile kennzeichnen die Transkriptionsrichtung jedes Gens. Das aktive Allel von *Peg14* wird durch ein schwarzes Rechteck dargestellt. Es wird vermutet, dass das humane Transkript paternal exprimiert wird. Eventuelle allelische Expressionsunterschiede von *USP29*, *ZIM3* und *ZNF264* sind noch nicht bekannt. Die grauen Rechtecke stellen die aktiven Allele der bekannten geprägten Gene dar. Inaktive Allele werden durch weiße Rechtecke dargestellt. Hellgraue Rechtecke markieren bisher nicht auf Prägung untersuchte Gene. Mat steht für das maternale und pat für das paternale Chromosom.

Da weder die DNA- noch die Aminosäuresequenz der putativen Genprodukte signifikant homolog sind, könnte eine Analyse der beiden Sequenzen hinsichtlich repetitiver Elemente näheren Aufschluss über die Gemeinsamkeiten der beiden Transkripte geben. Dies wurde mit dem Programm Repeatmasker des Instituts für Systembiologie in Seattle, WA, USA (Version 3.1.5, siehe http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) durchgeführt und ergab die in Tabelle 9 aufgelisteten repetitiven Elemente.

Tabelle 9. Art und Zusammensetzung repetitiver Elemente in *Peg14* sowie dem humanen Ortholog.

	murines *Peg14*	humanes *PEG14*
Länge	2979 bp	3940 bp
GC-Anteil	34,34%	34,82%
Repeats:		
SINE	1 B1-Element (155 bp)	2 Alu-Elemente (377 bp)
SINE	1 B2-B4-Element (171 bp)	-
LINE	2 LINE1-Elemente (537 bp)	2 LINE1-Elemente (1468 bp)
DNA-Elemente	-	1 MER1-Element (266 bp)
einfache Repeats	3 Elemente (215 bp)	-
Gesamtanteil im Transkript	36,19%	58,58%

Neben einem fast identischen GC-Gehalt (beide 34%) beinhalten Maus- und Mensch-Transkript SINE- und LINE-Transposons. Im ca. 1 kb längeren humanen *PEG14* befindet sich darüber hinaus ein MER1-Element, während im murinen *Peg14* drei einfache Repeat-Motive zu finden sind. Der Gesamtgehalt repetitiver Elemente ist im humanen Transkript mit etwa 59% deutlich größer als im Maus-Transkript mit 36%.

3.1.10. Allelische Expressionsanalyse von *USP29*, *ZIM3* und *ZNF264* auf HSA19q13.4

Während die genomische Prägung der Gene auf MMU7prox bereits gut untersucht ist, fehlt diese Analyse auf dem orthologen humanen Chromosom für *USP29, ZIM3* und *ZNF264* bisher. Die chromosomale Karte dieses Bereiches weist mehrere Unterschiede zwischen Maus und Mensch auf. So sind bei den humanen *ZIM2* und *PEG3* die ersten sieben Exons identisch und beide Gene paternal exprimiert, während die murinen Gene räumlich getrennt sind und eine reziproke Prägung entwickelt haben (Kim *et al.*, 2000a). Für das murine *Zim1* wurde bisher kein humanes Ortholog identifiziert. *Zim1* wird als nachträglich in MMU7prox inseriertes Gen diskutiert (Kim *et al.*, 2004b). Darüber gibt es noch weitere Unterschiede von *USP29, ZIM3* und *ZNF264* zu den murinen Orthologen. So fehlen bei *USP29* die ersten fünf nicht-kodierenden Exons (NCBI-Eintrag: NM_020903), was im Gegensatz zur orthologen Maus-Region, wo eine gemeinsame Imprinting-Regulation von *Peg3* und *Usp29* vermutet wird (Kim *et al.*, 2003), auf eine von *PEG3* getrennte Regulation

schließen lässt. Darüber hinaus weisen *USP29, ZIM3* und *ZNF264* im Gegensatz zu den murinen Orthologen keine einander überlappenden Sequenzen auf (siehe Abb. 21 unten).

Für die allelische Expressionsanalyse wurden aus der SNP-Datenbank von NCBI (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=snp) bekannte tSNPs für *USP29* (rs3764574), *ZIM3* (rs4801433) und *ZNF264* (rs917340) ausgewählt und die betreffenden genomischen Bereiche anschließend zur Bestimmung der Genotypen sequenziert. In der folgenden QUASEP-Analyse konnten bei informativen Individuen für alle drei Gene biallelische Expression in den untersuchten Geweben (fetales Gehirn, fetaler und adulter Testis) festgestellt werden. Als positive Kontrolle wurde die bekannte geprägte Expression von *PEG3* in informativen Individuen analysiert (siehe Abb. 22).

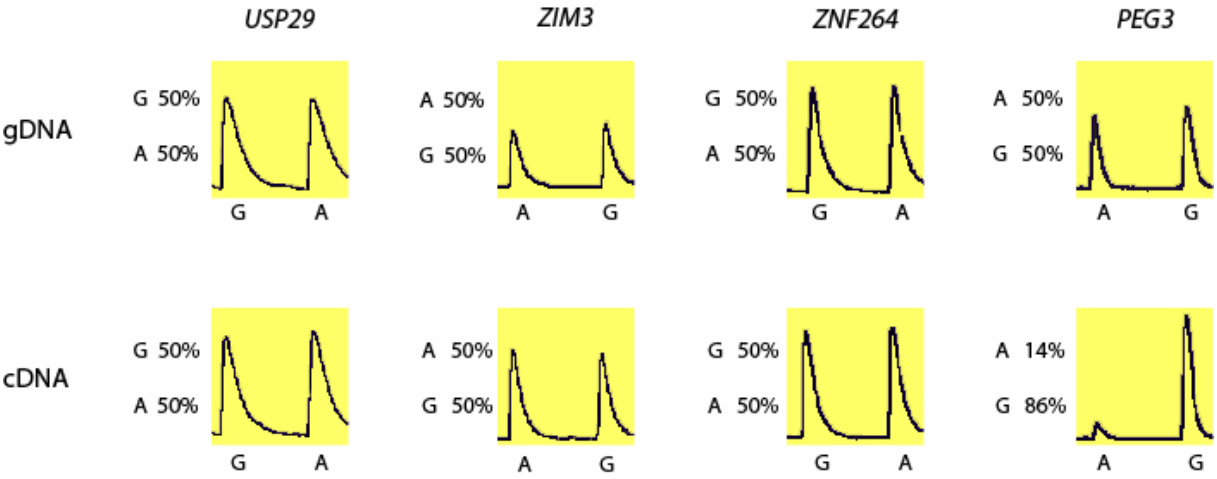

Abbildung 22. Biallelische Expression von *USP29*, *ZIM3* und *ZNF264* sowie präferentiell allelische Expression von *PEG3* auf dem humanen Chromosom 19q13.4. Dargestellt sind Ausschnitte aus den Pyrogrammen von untersuchten informativen Individuen. Oben sind jeweils die Pyrogramme dargestellt, die durch die Analyse genomischer DNA (gDNA) gewonnen wurden. Die unten abgebildeten Ausschnitte zeigen die Ergebnisse von RT-PCR-Produkten (cDNA) des jeweils gleichen Individuums. Im *USP29*-Transkript (links) wurde ein G/A-SNP (rs3764574) analysiert. Daneben ist das Ergebnis von einem A/G-SNP (rs4801433) im *ZIM3*-Transkript abgebildet. Die Analyse des *ZNF264*-Transkripts (drittes von links) basiert auf der Untersuchung eines G/A-SNPs (rs1055359). Als positive Kontrolle diente die bekannte geprägte Expression von *PEG3*. Wie rechts abgebildet, weist die Analyse von RT-PCR-Produkten eine präferentielle Expression des G-Allels auf. Das Verhältnis der beiden Allele ist links in Prozent angegeben.

3.2. Imprinting-Kandidatengene der sequenzbasierten Studie

3.2.1. Auswahl und Analyse der Kandidatengene

Nach der Vorhersage von 600 neuen Imprinting-Kandidatengenen im Mausgenom (Luedi *et al.*, 2005) wurden in einer Analyse für das menschliche Genom ebenfalls neue Imprinting-Kandidatengene vorhergesagt (Luedi *et al.*, noch unveröffentlicht). Unter Berücksichtigung der Ergebnisse beim menschlichen Modell wurden die benannten 600 neuen Kandidatengene im Mausgenom durch Philippe Luedi (Duke University, Durham, NC, USA) hinsichtlich des Algorithmuswertes im Mausmodell und übereinstimmender Vorhersagen beim menschlichen Modell priorisiert und auf 18 Gene eingeschränkt. Je größer die dabei ermittelten Werte für Priorität oder Klassifikationsalgorithmus sind, desto vielversprechender ist das jeweilige Kandidatengen. In den Wert für Priorität sind darüber hinaus auch weitere Parameter wie in Beziehung stehende Phänotypen und/oder weitere benachbarte Kandidatengene eingeflossen (siehe Tab. 10).

Im nächsten Schritt wurden durch Sequenzierung bei 16 der 18 Kandidatengene geeignete tSNPs identifiziert (siehe Anhang Tab. 16) und anschließend eine QUASEP-Analyse durchgeführt.

Tabelle 10. Allelische Expressionsanalyse der untersuchten Imprinting-Kandidatengene (Luedi-Studie). Die beiden Gene mit monoallelischer bzw. präferentiell stammspezifischer Expression sind fett hervorgehoben.

Genname	Priorität (Algorithmus-Wert)	Vorhersagewert im menschlichen Modell	chromosomale Position (bp)	Verhältnis allelischer Expression (C57BL/6 versus Cast/Ei) SD = Standardabweichung
Foxg1	9 (0,32)	17	12cen (47396718-47400567)	biallelisch (49% : 51%; SD 0,6)
Kcnk9	**9 (0,53)**	**17**	**15dist (72591862-72626022)**	**monoallelisch maternal (99% : 1%; SD 1,6)**
Prdm16	9 (0,01)	15	4dist (153309834-153626302)	biallelisch (52% : 48%; SD 0,9)
Nppc	8 (0,01)	10	1cen (86566967-86571248)	biallelisch (55% : 45%; SD 2,8)
Wnt7b	8 (0,21)	9	15dist (85617762-85663903)	biallelisch (53% : 47%; SD 3,4)
Dok1	8 (0,21)	8	6cen (83056368-83058899)	biallelisch (54% : 46%; SD 1,8)
Nkx6-2	7 (0,4)	11	7dist (135993086-135996580)	biallelisch (43% : 57%; SD 2,2)
Hes5	7 (0,44)	8	4dist (153950465-153951913)	kein tSNP
Cdk6	7 (0,07)	8	5prox (3350318-3528231)	biallelisch (47% : 53%; SD 1,4)
Rarres1	**7 (0,34)**	**4**	**3cen (67579972-67616619)**	**präferentiell stammspezifisch Cast/Ei (33% : 67%, SD 9,3)**
Camk2b	6 (0,38)	8	11prox (5921675-6017751)	biallelisch (51% : 49%, SD 3,9)
Ntng2	6 (0,39)	5	2prox (29202132-29255351)	biallelisch (51% : 49%, SD 3,2)
Disc1	6 (0,21)	3	8dist (123641593-123849582)	kein tSNP
Irx4	5 (0,38)	15	13cen (70192290-70201087)	biallelisch (50% : 50%; SD 1,6)
Stk32c	5 (0,008)	8	7dist (135523291-135602224)	biallelisch (52% : 48%, SD 4,1)
Ly6d	5 (0,17)	8	15cen (74842606-74844117)	biallelisch (59% : 41%; SD 4,2)
Gad2	5 (0,06)	7	2prox (22607267-22675939)	biallelisch (50% : 50%; SD 1,5)
Gdnf	5 (0,08)	6	15prox (7637824-7664388)	biallelisch (47% : 53%; SD 2,4)

Dabei wies *Kcnk9* eine monoallelische maternale Expression bei F1-Maushybriden (E11.5) auf, die durch die QUASEP-Analyse von informativen Mausembryonen einer reziproken Kreuzung bestätigt wurde (siehe Abb. 23). Bei *Rarres1,* einem weiteren Kandidatengen, ist das Cast/Ei-Allel präferentiell exprimiert. Die übrigen 14 Kandidatengene zeigten bei der QUASEP-Analyse von RT-PCR-Produkten aus F1-Hybriden (E11.5) biallelische Expression (siehe Tab. 10).

Anschließend wurde die Expression von *Kcnk9* mittels RT-PCR in adulten Geweben (Muskel, Niere, Leber, Lunge, Herz, Testis, Gehirn und Milz) untersucht, konnte jedoch lediglich im Gehirn nachgewiesen werden. Die nachfolgende QUASEP-Analyse ergab auch im Gehirn eine monoallelische maternale Expression (siehe Abb. 23).

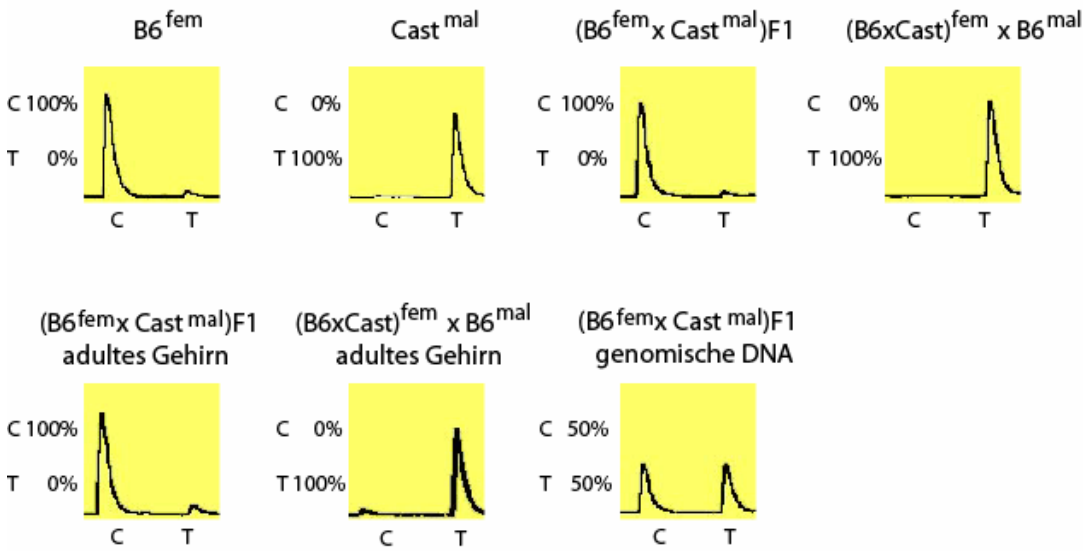

Abbildung 23. Allelische Expressionsanalyse für das murine *Kcnk9*-Gen. Für jedes Transkript sind Ausschnitte aus den Pyrogrammen von den homozygoten Elterntieren ($B6^{fem}$ für das C57BL/6-Muttertier und $Cast^{mal}$ für das Cast/Ei-Vatertier) sowie von den F1-Hybriden ($B6^{fem}$ x $Cast/Ei^{mal}$)F1 (E11.5) und informativen reziproken Hybriden ($(B6xCast)^{fem}$ x $B6^{mal}$) dargestellt. Wie in der unteren Reihe der Abbildung gezeigt, bleibt die maternale Expression von *Kcnk9* im Gehirn auch im adulten Stadium erhalten. Die Analyse von PCR-Produkten aus genomischer DNA weist die erwartete Verteilung auf. Das Verhältnis der beiden Allele ist jeweils links in Prozent angegeben.

Kcnk9 gehört zur heterogenen Gruppe der Kaliumkanalgene (Patel und Lazdunski, 2004). Die mRNA ist 1209 bp lang und besteht aus zwei Exons (NCBI-Eintrag: NM_001033876). *Kcnk9* ist auf MMU15dist in der zytologischen Bande D3 lokalisiert und damit ca. 260 kb vom paternal exprimierten Gen *Peg13* entfernt (Smith *et al.*, 2003). MMU15dist ist laut Homologiekarte des Jackson Laboratory syntän zu HSA8q24.3 (http://www.jax.org/resources/index.html). Für *Peg13* konnte bisher kein humanes Ortholog identifiziert werden. Ein Vergleich der Sequenz von *Peg13* mit dem Humangenom mittels Ensembl-BLAST ergab keine Homologie. Für das murine *Kcnk9*-Gen ist hingegen ein humanes Ortholog bekannt, welches auf HSA8q24.3 kartiert.

3.2.2. Allelische Expressionsanalyse des humanen *KCNK9*

Die Mehrzahl der geprägten Gene im Mausgenom ist auch beim Menschen geprägt (Morison *et al.*, 2001). Darüber hinaus wurde im menschlichen Vorhersagemodell für das humane *KCNK9* die höchste Wahrscheinlichkeit einer genomischen Prägung berechnet (siehe Tab. 10). Aufgrund dieser beiden Indizien wurde auch das humane *KCNK9* hinsichtlich allelspezifischer Expression in den entsprechenden Geweben untersucht. *KCNK9* ist im Cerebellum am stärksten exprimiert, kann aber auch in verschiedenen anderen Regionen des Gehirns wie auch in verschiedenen Teilen des Verdauungstraktes nachgewiesen werden (Chapman *et al.*, 2000; Medhurst *et al.*, 2001; Kovacs *et al.*, 2005).

Für den bekannten T/C-SNP im Exon 2 von *KCNK9* (rs2615374) konnten zwei heterozygote Individuen (DNA88 und DNA261) identifiziert werden, von denen fetale Gehirn-RNA vorlag. Die QUASEP-Analyse ergab bei beiden Individuen monoallelische Expression des T-Allels (siehe Abb. 24).

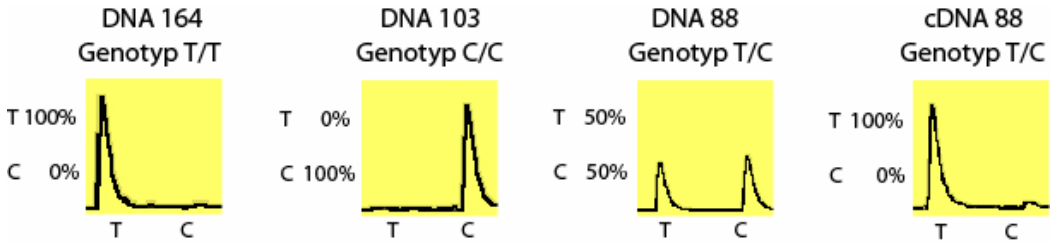

Abbildung 24. Allelische Expressionsanalyse für das humane *KCNK9*-Gen im fetalen Gehirn. Hier dargestellt sind die Ausschnitte aus den Pyrogrammen mit den drei verschiedenen Genotypen für den T/C-Polymorphismus (rs2615374) im Exon 2 von *KCNK9*, die in PCR-Produkten aus genomischer DNA analysiert wurden. Ganz rechts ist die monoallelische Expression in fetalem Gehirn am Beispiel der Probe 88 dargestellt. Das Verhältnis der beiden Allele ist links in Prozent angegeben.

Die Ergebnisse der daraufhin an der Duke University durchgeführten Sequenzierung von RT-PCR-Produkten aus Gehirnproben von informativen Individuen, bei denen der elterliche Ursprung der Allele bekannt war, belegten die monoallelische maternale Expression von *KCNK9* und bestätigten damit die QUASEP-Ergebnisse sowie die bioinformatisch vorhergesagte maternale Expression (siehe Supplement von (Luedi *et al.*, 2005)).

3.2.3. CpG-Inseln in der *Kcnk9*-Promotorregion

In der flankierenden Region liegende CpG-Inseln sind ein wichtiges Merkmal geprägter Gene, welches für die sequenzbasierte Vorhersage der neuen Imprinting-Kandidatengene von Luedi *et al.* (2005) berücksichtigt wurde.

Die *in silico*-Suche nach CpG-Inseln mit dem Programm Emboss-CpG-Plot (siehe http://www.ebi.ac.uk/emboss/cpgplot/) und der genomischen Sequenz beider Orthologe beginnend 5 kb vor Exon 1 bis 2 kb nach dem letzten Exon ergab drei bzw. sechs Sequenzbereiche (siehe Tab. 11), die typische Eigenschaften von CpG-Inseln aufweisen. Dazu zählen nach den Programmparametern des Emboss-CpG-Plots ein GC-Anteil von über 50% und eine Länge von mindestens 200 bp. In der Literatur sind unterschiedliche Angaben zur Mindestlänge von CpG-Inseln von 500 bis 1000 bp zu finden ((Reik und Walter, 2001; Robertson, 2005; Strachan und Read, 2005) Seite 239).

Tabelle 11. CpG-Inseln in der flankierenden Region des murinen und humanen *Kcnk9*-Gens.

	genomischer Bereich (kb)	Größe (bp)	relative Position zum ATG (bp)
murines *Kcnk9*	41	317	-1326 bis -1009
		726	-905 bis -179
		302	-149 bis +153
humanes *KCNK9*	97	544	-3168 bis -2624
		213	-1498 bis -1285
		619	-1264 bis -645
		1274	-632 bis +640
		214	+83906 bis +84120
		275	+84327 bis +84602

Eine allelspezifische Methylierung innerhalb der CpG-Inseln, insbesondere in der Promotorregion, könnte ursächlich für die monoallelische maternale Expression von *Kcnk9* verantwortlich sein. Daher wird in zukünftigen Experimenten eine Methylierungsanalyse der in Tabelle 11 aufgelisteten CpG-Inseln durchzuführen sein.

4. Diskussion

Wie viele geprägte Loci bzw. Gene es gibt, ist bisher nicht bekannt. Daher sind systematische Ansätze zur genomweiten Identifizierung neuer geprägter Gene von großer Bedeutung. In einer Microarray-Studie wurden im Mausgenom 2101 neue Imprinting-Kandidatengene identifiziert (Nikaido *et al.*, 2003). In einer weiteren Untersuchung konnten durch eine vergleichende Sequenzanalyse bekannter geprägter Gene *in silico* insgesamt 600 neue Imprinting-Kandidatengene im Mausgenom vorhergesagt werden (Luedi *et al.*, 2005). Da die methodischen Ansätze beider Studien keine zuverlässige Aussage hinsichtlich tatsächlicher Prägung erlauben, ist es folglich eine wichtige Aufgabe, die ermittelten Kandidatengene in einer allelspezifischen Expressionsanalyse zu validieren.

In dieser Arbeit wurden ausgewählte Kandidatengene aus beiden Studien in (C57BL/6 x Cast/Ei)F1-Hybriden (E11.5) validiert. Dazu wurde eine neue Methode (QUASEP) etabliert, die auf der Pyrosequencing™-Technologie basiert. Von den 55 getesteten Kandidatengenen der Microarray-Studie wiesen 47 Transkripte im Embryonalstadium eine biallelische Expression, vier Transkripte eine präferentiell stammspezifische Expression und vier Transkripte eine geprägte Expression auf (AK043056 (*Peg14*), AK018142, AK039514 und AK039482; siehe Tab. 5). In verschiedenen adulten Geweben konnte die Aufrechterhaltung der genomischen Prägung dieser vier Transkripte während der Ontogenese nachgewiesen werden (siehe Tab. 6). Aufgrund der besonderen Lage von AK043056 innerhalb der Transkriptionseinheit von *Usp29* wurde dieses Transkript in weiteren Experimenten näher charakterisiert. In einem Expressionsvergleich zwischen AK043056 und *Usp29* konnte eine tendenziell ähnliche, aber dennoch unterschiedliche Expression in Gehirn, Testis und Plazenta gezeigt werden (siehe Tab. 8). Darüber hinaus wurde in der syntänen Region auf dem humanen Chromosom 19q13.4 ein monoallelisch exprimiertes Transkript identifiziert. Die Länge dieses humanen Orthologs von AK043056 wurde mit 3940 bp bestimmt. Das murine Transkript konnte mittels 5´-RACE von 2612 bp auf 2979 bp verlängert werden. Entsprechend der üblichen Nomenklatur wurden beide Orthologe als *Peg14* (für *Paternally expressed gene* 14) benannt. Die QUASEP-Analyse der benachbarten Gene *USP29*, *ZIM3* und *ZNF264*, die in der Maus geprägt sind, jedoch beim

Menschen bisher nicht untersucht worden waren, ergab biallelische Expression in humanem fetalen Gehirn sowie humanem fetalen und adulten Testis. Von den 16 getesteten Imprinting-Kandidatengenen der Luedi-Studie zeigten 14 Transkripte biallelische Expression, das Gen *Rarres1* präferentielle Expression des Cast/Ei-Allels und das Gen *Kcnk9* monoallelische maternale Expression. Die QUASEP-Analyse verschiedener Gewebe zeigte den Erhalt der Prägung von *Kcnk9* im adulten Stadium. Die Expressionsanalyse des humanen *KCNK9*-Gens im fetalen Gehirn von informativen Individuen ergab ebenfalls monoallelische maternale Expression.

4.1. Microarray-Studie

Von den ursprünglich ausgewählten 68 Kandidatengenen der Microarray-Studie von (Nikaido *et al.*, 2003) konnte bei 55 Kandidatengenen die vorhergesagte geprägte Expression experimentell überprüft werden. Bei den vier Transkripten AK039514, AK018142, AK043056 und AK039482 wurde eine geprägte Expression nachgewiesen.

4.1.1. Transkript AK039514 auf MMU17prox

Das Transkript AK039514 wurde, wie bereits beschrieben, als Teil des bekannten geprägten Gens *Air* identifiziert. Im derzeitigen Eintrag der Referenzsequenz-Datenbank von NCBI (RefSeq: NR_002853.1) wird die Länge von *Air* mit 3699 bp angegeben, obwohl die experimentell ermittelte Transkriptlänge ca. 108 kb beträgt (GenBank: AJ249895.1 (Lyle *et al.*, 2000)). AK039514 kartiert ca. 25 kb stromabwärts von NR_002853.1 und wurde daher anfangs bei den Sequenzvergleichen mittels BLAT-Search und NCBI-BLAST als eigenständiges Transkript identifiziert. AK039514 wird nachträglich als weitere Kontrolle betrachtet, durch die zusätzlich die Eignung von QUASEP zur allelspezifischen Expressionsanalyse überprüft wurde.

4.1.2. Transkript AK018142 auf MMU12dist

Bei dem maternal exprimierten AK018142-Transkript besteht der begründete Verdacht, dass dieses zum ebenfalls maternal exprimierten Gen *Gtl2* gehört. AK018142 liegt nur 340 bp stromabwärts hinter dem 3´-Ende von *Gtl2* (RefSeq: NM_144513.2). *Gtl2* wurde von Schuster-Gossler *et al.* (1998) als vermutlich nicht-kodierende RNA mit vielen

verschiedenen Spleißvarianten beschrieben. Die Autoren der Arbeit konnten mittels Northern-Blot-Hybridisierung und Sequenzierung zahlreicher cDNA-Klone eine verlangsamte Prozessierung der mRNA nachweisen, wodurch sich die Anzahl der verschiedenen Isoformen erhöht. Die Länge der reifen mRNA wird unter 2 kb vermutet (Schuster-Gossler *et al.*, 1998). In einem Ansatz zur Sequenzierung einiger tausend cDNA-Klone wurde die Transkriptlänge von *Gtl2* mit 4326 bp einschließlich dem poly(A)-Ende identifiziert (Strausberg *et al.*, 2002). Die Arbeit von Schuster-Gossler *et al.* (1998) beinhaltet eine wesentlich genauere Analyse der verschiedenen *Gtl2*-Transkripte, die durch eine weitere Studie bestätigt wurde (Miyoshi *et al.*, 2000). Eine Analyse des 3´-Endes bzw. eine Suche nach weiteren Isoformen mit alternativen 3´-Enden wurde in beiden Studien jedoch nicht unternommen. Die Darstellung im ContigView von Ensembl weist eine Vielzahl sich überlappender ESTs für einen ca. 10 kb großen Bereich hinter dem angegebenen 3`-Ende von *Gtl2* aus. Es deutet somit vieles darauf hin, dass AK018142 Teil einer *Gtl2*-Isoform mit einem alternativen 3´-Ende ist. Durch die nachgewiesene verlangsamte Prozessierung der mRNA ist es denkbar, dass AK018142 zu einer heteronukleären *Gtl2*-Isoform gehört. *Gtl2* ist gemäß UniGene in zahlreichen Geweben jedoch nicht im Skelettmuskel exprimiert. Die für die vorliegende Arbeit durchgeführten Expressions- und Imprinting-Analysen von AK018142 in verschiedenen adulten Geweben der Maus belegen jedoch auch eine monoallelisch maternale Expression im Skelettmuskel (siehe Tab. 6). Unter der Annahme, dass AK018142 Teil einer *Gtl2*-Isoform ist, kann abgeleitet werden, dass diese putative Isoform im Muskel exprimiert ist und dort wahrscheinlich eine spezifische Funktion erfüllt.

An dieser Stelle muss auf den Callipyge-Phänotyp verwiesen werden. Dabei handelt es sich um eine Muskelhypertrophie beim Schaf, welche durch eine heterozygote Transition (A>G) auf dem paternalen Allel hervorgerufen wird. Die Mutation liegt im Bereich zwischen *DLK1* und *GTL2* und ist mit einer Hypomethylierung sowie einer erhöhten bidirektionalen Expression korreliert (Takeda *et al.*, 2006). Es ist denkbar, dass eine muskelspezifische *Gtl2*-Isoform auch beim Schaf existiert und bei einer durch den Verlust der genomischen Prägung verursachten, erhöhten Expression zum Callipyge-Phänotyp beiträgt.

4.1.3. Transkript AK039482 auf MMU7dist

Das paternal exprimierte AK039482 ist sehr wahrscheinlich Teil des bekannten geprägten Gens *Lit1*. *Lit1* ist eine nicht-kodierende RNA und besteht nur aus einem Exon. Die RefSeq-Datenbank von NCBI gibt die Transkriptlänge von *Lit1* mit 84 kb an (NR_001461.1). Durch Yatsuki *et al.* (2000) wurde jedoch experimentell nachgewiesen, dass die *Lit1*-RNA mindestens 20 kb länger und daher insgesamt über 104 kb lang ist. Aufgrund dieser Länge ist es sehr schwierig das 3`-Ende von *Lit1* korrekt zu identifizieren (Yatsuki *et*

al., 2000). AK039482 liegt ca. 16,5 kb stromabwärts vom derzeit angegebenen 3´-Ende von *Lit1*. Durch die Verlängerung von *Lit1* um 20 kb ist AK039482 somit als Teil von *Lit1* verifiziert.

Die in der vorliegenden Arbeit nachgewiesene Relaxation der Prägung von AK039482 (*Lit1*) ist sehr beachtenswert (siehe Tab. 6). Während AK039482 im Embryo strikt monoallelisch paternal exprimiert ist, kann im adulten Gewebe eine präferentiell paternale Expression nachgewiesen werden. Bislang wurde in keiner Studie diese Relaxation der Prägung von *Lit1* identifiziert. Die geprägte Expression von *Lit1* wurde von Yatsuki *et al.* (2000) mittels RFLP-Analyse beschrieben. Mit QUASEP lassen sich im Vergleich dazu auch geringe allelische Expressionsunterschiede detektieren. Somit kann eine monoallelische Expression von einer präferentiell allelspezifischen Expression unterschieden werden. Die in dieser Arbeit nachgewiesene Relaxation der *Lit1*-Prägung im adulten Stadium sollte in einer Folgearbeit durch zusätzliche QUASEP-Experimente mit einem weiteren *Lit1*-Polymorphismus bestätigt werden. Da *Lit1* als bidirektionaler Repressor für die transkriptionelle Inaktivierung der paternalen Allele benachbarter Gene verantwortlich ist (Mancini-Dinardo *et al.*, 2006), sollte darüber hinaus durch eine systematische QUASEP-Analyse der benachbarten Gene ergründet werden, ob und welche Auswirkungen eine Relaxation der *Lit1*-Prägung auf deren geprägte Expression hat.

4.1.4. *Peg14* (AK043056) auf MMU7prox

Die besondere genomische Lage von AK043056 innerhalb der Transkriptionseinheit von *Usp29* (siehe Abb. 14), der Expressionsvergleich von AK043056 und *Usp29* sowie der Nachweis eines monoallelisch exprimierten Transkripts in der orthologen Region auf dem humanen Chromosom 19q13.4, welches unabhängig von *USP29* transkribiert wird, belegen, dass AK043056 ein neues und eigenständiges geprägtes Gen ist. Darüber hinaus sprechen noch folgende Hinweise für die Eigenständigkeit von AK043056. Zum einen wurden bei der cDNA-Synthese oligo(dT)-Primer verwendet, so dass hauptsächlich poly(A)-RNA in cDNA umgeschrieben wurde. Die Distanz zwischen dem 3´-Ende von AK043056 und dem 3´-Ende einer heteronukleären *Usp29*-RNA, das mit dem poly(A)-Ende als Startpunkt der reversen Transkription dient, beträgt ca. 226 kb. Weitere näher liegende poly(A)-Motive, die als Bindungsstellen der oligo(dT)-Primer dienen können, sind in der heteronukleären RNA von *Usp29* nicht zu finden. Es ist daher anzunehmen, dass die analysierten RT-PCR-Produkte nicht von heteronukleärer *Usp29*-RNA, sondern von einer eigenständigen RNA amplifiziert wurden und stabil amplifizierbar sind. Zum anderen beweist die reproduzierbare Verlängerung am 5´-Ende von AK043056 mittels RACE die Eigenständigkeit dieses Transkripts. Das neu identifizierte 5`-Ende von AK043056 weist keine Ähnlichkeit mit dem 5´-

Ende von *Usp29* auf. Die Benennung als *Peg14* erfolgte gemäß der üblichen Nomenklatur für paternal exprimierte Gene.

4.1.5. Präferentiell stammspezifisch exprimierte Transkripte

Bei den vier Transkripten AK038367, AK043873, AK035722 und AK014850 wurde präferentiell stammspezifische Expression nachgewiesen. Während bei AK035722 das C57BL/6-Allel präferentiell exprimiert ist, zeigen die übrigen drei Transkripte eine bevorzugte Expression des Cast/Ei-Allels. Die stammspezifische allelische Expression wird vermutlich durch *cis*-regulatorische Faktoren wie DNA-Polymorphismen und/oder DNA-Methylierung in der flankierenden Sequenz hervorgerufen. In einer Studie über *cis*-regulatorische Polymorphismen wurden 69 Gene in den Kreuzungen von vier Mausstämmen untersucht. Dabei wurden vier Gene mit allelischen Expressionsunterschieden von über 50% identifiziert. In der Promotorregion dieser vier Gene wurden anschließend mehrere Polymorphismen identifiziert, die die Expression beeinflussen könnten (Cowles *et al.*, 2002). Man vermutet, dass diese regulatorischen Polymorphismen zur phänotypischen Diversität und zur Entstehung zahlreicher Erkrankungen beitragen (Pastinen und Hudson, 2004).

4.1.6. Biallelisch exprimierte Kandidatengene

Von den mittels QUASEP analysierten 55 Kandidatengenen weisen 47 Transkripte (entspricht 85%) biallelische Expression im Embryo der Maus (E11.5) auf. In der bereits erwähnten Microarray-Studie von Mizuno *et al.* (2002) waren unter den 25 ausgewählten Imprinting-Kandidatengenen nur 11 biallelisch exprimierte Gene zu finden (entspricht 44%). Dieser signifikante Unterschied in der Anzahl falsch positiver Kandidatengene lässt sich damit begründen, dass diese nach verschiedenen Kriterien ausgewählt wurden. Während Mizuno *et al.* (2002) ein analysiertes Gen erst bei einem 75fachen allelischen Expressionsunterschied zwischen androgenetischen (AG) und parthenogenetischen (PG) Mausembryonen und unabhängig von der chromosomalen Lokalisation als Kandidatengen definierten, wurden bei Nikaido *et al.* (2003) die Transkripte bereits bei einem zehnfachen allelischen Expressionsunterschied als Kandidatengene definiert, die dann in der vorliegenden Arbeit aufgrund ihrer Lokalisation in bekannten geprägten Regionen ausgewählt wurden.

Die hohe Anzahl falsch positiver Kandidatengene weist darauf hin, dass eine Vielzahl nicht geprägter Gene in AG- und PG-Embryonen unterschiedlich exprimiert ist. In diesem Zusammenhang ist in Betracht zu ziehen, dass die von Surani et al. (1986) beschriebenen

phänotypischen Unterschiede zwischen uniparentalen Embryonen an sich bereits eine Erklärung für die zahlreichen falsch positiven Kandidatengene liefert. So wird vermutet, dass sich die regulatorischen Netzwerke zwischen normalen und uniparentalen Embryonen deutlich unterscheiden, was zu unterschiedlicher Expression nicht geprägter Gene führt (Mizuno *et al.*, 2002). Des Weiteren ist denkbar, dass geprägte Gene als Bestandteil des globalen Regulationsnetzwerkes die Transkription nicht geprägter Gene induzieren.

Letztlich kann auch nicht ausgeschlossen werden, dass einige der im Embryo (E11.5) getesteten Transkripte gewebespezifisch geprägt sind oder ihre Prägung nach dem neunten Tag der Embryonalentwicklung (E9.5), wie bei Nikaido *et al.* (2003) getestet, verlieren. In beiden Fällen würde man nur eine sehr kleine Anzahl an Genen erwarten, die durch die hier beschriebene systematische Imprinting-Analyse nicht hätten identifiziert werden können.

4.1.7. Zusammenfassende Bewertung

Durch die systematische Imprinting-Analyse aller Kandidatengene der Microarray-Studie von Nikaido *et al.* (2003), die in bekannten geprägten Regionen lokalisiert sind, wurden drei neue geprägte Transkripte identifiziert. Darüber hinaus wurde bei vier Transkripten präferentiell stammspezifische Expression nachgewiesen. In vier Transkripten konnte kein geeigneter tSNP identifiziert werden und bei fünf weiteren Kandidatengenen konnte in F1-Embryonen (E11.5) keine Expression detektiert werden. Die Expression eines Kandidatengens in den untersuchten E11.5-Embryonen wie auch geeignete tSNPs sind Voraussetzung für eine QUASEP-basierte Imprinting-Analyse und müssen als Einschränkung dieses experimentellen Ansatzes bewertet werden.

47 Kandidatengene weisen biallelische Expression auf und zeigen damit zum einen, dass die Lokalisation eines Kandidatengens kein geeigneter Prädiktor für eine Priorisierung ist, und zum anderen, dass der Vergleich von Expressionsprofilen von uniparentalen Mausembryonen mittels cDNA-Microarrays kein effizienter Ansatz zur Identifizierung neuer geprägter Gene ist.

4.2. Genomische Prägung auf MMU7prox und HSA19q13.4

Peg14 ist ein neues geprägtes Gen und kartiert innerhalb der *Peg3*-Domäne auf MMU7prox bzw. in der orthologen Region auf HSA19q13.4. Welche biologische Funktion *Peg14* erfüllt, ist noch unklar. Dies betrifft insbesondere auch eine mögliche Rolle bei der Imprinting-Regulation der benachbarten Gene. Die Regulation der geprägten Gene in der

Peg3-Domäne ist ursächlich bisher nicht vollständig aufgeklärt. Auch das Wissen um die gewebespezifische Expression ist unvollständig. Eine genaue laborexperimentelle Untersuchung und Analyse der geprägten Gene hätte den Rahmen dieser Arbeit jedoch gesprengt, so dass im Folgenden nur der derzeitige Stand des Wissens reflektiert werden kann.

4.2.1. *Peg14* – eine nicht-kodierende RNA

Neben der im Kapitel 3.1.7 beschriebenen 513 bp großen Region, die eine Homologie von 66% aufweist, ergibt ein Vergleich der Nukleotidsequenzen des murinen und humanen *PEG14* mit insgesamt 34% keine signifikante Übereinstimmung (siehe auch Kapitel 3.1.9). Trotz der geringen Homologie sind beide Transkripte an vergleichbarer Position in den zueinander orthologen und konservierten Regionen auf MMU7prox und HSA19q13.4 lokalisiert. Auch die Aminosäuresequenzen, die aus dem jeweils längsten offenen Leserahmen (ORF) resultieren, weisen mit 6% keine signifikante Ähnlichkeit auf. Die Aminosäurezusammensetzung des längsten humanen *PEG14*-ORF hat darüber hinaus keine Ähnlichkeit mit typischen Säugerproteinen, da kein Glutamat sowie zu viel Prolin und Serin enthalten sind[3]. Das gilt auch für zwei weitere ORFs, die weiter am 3´-Ende im Transkript zu finden sind.

Eine BLAST-Suche, bei der die putative Aminosäuresequenz des humanen *PEG14* mit der Proteindatenbank von NCBI verglichen wird, weist zahlreiche Alu-Sequenzen mit signifikanter Homologie über 70% aus. Eine dieser Sequenzen ist P39192 ALU5_HUMAN (Alu subfamiliy SC sequence contamination entry 585AA) mit 72%iger Homologie. Man kann daher vermuten, dass es sich bei *PEG14* um ein noch aktives Transposon der Alu-Familie und nicht um ein exprimiertes Protein handelt. Die RT-PCR von *PEG14* mit oligo(dT)-Primern basierte jedoch auf der Existenz einer komplementären poly(A)-Sequenz, wie sie bei der Transkription durch RNA-Polymerase II entsteht ((Strachan und Read, 2005) Seite 25). Da Alu-Transposons jedoch nur sehr selten und dann lediglich durch RNA-Polymerase III transkribiert werden (Grover *et al.*, 2004), ist die obige Vermutung eher auszuschließen.

Eine BLAST-Suche mit der Sequenz des putativen murinen *Peg14*-ORF ergab keine Übereinstimmung mit bekannten Sequenzmotiven in der NCBI-Datenbank. Neben den LINE1-Elementen beinhaltet das murine *Peg14* mit dem B1-Element eine analoge Sequenz zum primatenspezifischen Alu-Element. Eine funktionelle Rolle als Teil von *Peg14* ist jedoch nicht anzunehmen, da sich Alu- bzw. B1-Elemente im Laufe der Evolution im humanen wie im murinen Genom immer weiter ausgebreitet haben (Kass *et al.*, 2000; Polak und Domany,

[3] persönliche Kommunikation mit Prof. Dr. Wolfgang Höhne, Humboldt-Universität zu Berlin

2006), was dafür spricht, dass beide Orthologe erst nachträglich mit Alu- bzw. B1-Elementen durchsetzt wurden.

Aus den aufgezählten Sequenzmerkmalen lässt sich keine Schlussfolgerung in Bezug auf eine Funktion von *Peg14* ziehen. Wahrscheinlich handelt es sich bei beiden Orthologen um nicht-kodierende RNAs. Die Funktionen beider Gene sind bislang unbekannt.

4.2.2. Nicht-kodierende RNAs auf MMU7prox

Obwohl die beiden orthologen Regionen ein hohes Maß an Übereinstimmung aufweisen (siehe Abb. 21), gibt es zahlreiche funktionelle Unterschiede zwischen den einzelnen orthologen Genen. In Tabelle 12 ist unter Berücksichtigung von *Peg14* der aktuelle Stand des Wissens zusammengefasst. In der Maus weisen die Gene bis auf *Zim2*, *Zim1* und *Zfp264* in allen bisher untersuchten Geweben geprägte Expression auf. *Zim2* und *Zfp264* sind in Testis biallelisch exprimiert und bei *Zim1* gibt es Hinweise auf biallelische Expression im Gehirn (Kuroiwa *et al.*, 1996; Kim *et al.*, 1999; Kim *et al.*, 2000b; Kim *et al.*, 2001; Murphy *et al.*, 2001; Kim *et al.*, 2004b).

Tabelle 12. Allelische Expression und Funktion der geprägten Gene auf MMU7prox / HSA19q13.4.

	Murines Ortholog		Humanes Ortholog	
Gen	Expression	kodiert für	Expression	kodiert für
Zim2/ZIM2	maternal	RNA, kein ORF	paternal	Zinkfinger
Zim1	maternal	Zinkfinger	kein humanes Ortholog von *Zim1* identifiziert	
Peg3/PEG3	paternal	Zinkfinger	paternal	Zinkfinger
Peg14/PEG14	paternal	RNA, vermutlich kein ORF	monoallelisch	RNA, vermutlich kein ORF
Usp29/USP29	paternal	Peptidase	biallelisch	Peptidase
Zim3/ZIM3	maternal	RNA, kein ORF	biallelisch	Zinkfinger
Zfp264/ZNF264	paternal	RNA, kein ORF	biallelisch	Zinkfinger

Bemerkenswerterweise kodieren die Maus-Gene *Zim2*, *Zim3* und *Zfp264*, deren humane Orthologe biallelische Expression oder reziproke Prägung aufweisen, im Gegensatz zu den humanen Genen nicht für Proteine oder weisen wie bei *Usp29* nur eine geringe Homologie des ORF auf (65%, siehe (Kim *et al.*, 2000b)). Welche biologische Funktion diese nicht-kodierenden RNA-Gene auf MMU7prox erfüllen bzw. welche Rolle die geprägte Expression dabei spielt, ist noch unklar. Nach der Evolutionstheorie lässt die allelspezifische Expression dieser Gene, die mit der transkriptionellen Inaktivierung eines Allels verbunden ist, einen Selektionsvorteil gegenüber biallelischer Expression erwarten (Wilkins, 2005).

4.2.3. *Peg14* und die Imprinting-Regulation

Nicht-kodierende RNAs spielen eine wichtige Rolle bei der Regulation geprägter Gene. Bis heute wurden zwei generelle mechanistische Prinzipien entwickelt, die aus der systematischen Erforschung der *H19*- und *Lit1*-Subdomänen auf MMU7dist resultieren (Reik *et al.*, 2004). Mit dem ersten Prinzip, dem so genannten *insulator loop model*, kann die Regulation der *H19*-Subdomäne (siehe Kapitel 1.2), aber auch der *Dlk1/Gtl2*-Domäne auf MMU12dist erklärt werden. Das zweite Prinzip, welches als *noncoding RNA model* bezeichnet wird, basiert auf der Wirkung einer nicht-kodierenden RNA, die als *cis*-aktiver Repressor agiert. Mit diesem zweiten Prinzip lässt sich die Regulation der *Lit1*-Subdomäne erklären, bei der das paternal exprimierte *Lit1* als bidirektionaler Repressor die Expression der paternalen Allele der benachbarten Gene inhibiert (Mancini-Dinardo *et al.*, 2006). Während man bisher davon ausgegangen ist, dass die Transkription von *Lit1 per se* der entscheidende Schritt für die Wirkung als Repressor ist, konnte in zwei aktuellen Studien gezeigt werden, dass die Transkriptlänge einen großen Einfluss auf die Effizienz der Repression hat (Kanduri *et al.*, 2006; Mancini-Dinardo *et al.*, 2006). Das *noncoding RNA model* ist auch für *Igf2r/Air*, die Inaktivierung des X-Chromosoms durch *Xist* sowie andere maternal geprägte Loci anwendbar (Reik *et al.*, 2004). Der genaue Mechanismus ist jedoch nicht bekannt. Man weiß jedoch, dass das Chromatin im Bereich des Promotors eines Gens bei Anwesenheit eines Antisense-Gens zahlreiche Modifikationen (z.B. Hypermethylierung) aufweist, die charakteristisch für inaktives Chromatin sind (Kanduri *et al.*, 2006).

Trotz des Vorhandenseins dieser vier geprägten, nicht-kodierenden RNAs auf MMU7prox gibt es bisher keine Anhaltspunkte, dass diese bei der Imprinting-Regulation, wie im *noncoding RNA model* vorgeschlagen, eine wichtige Rolle spielen. Alle geprägten Gene auf MMU7prox sind mit CpG-Inseln assoziiert. Nur eine ca. 5 kb große CpG-Insel, die den Promotor von *Usp29* und *Peg3* umfasst und sich bis ins erste Intron von *Peg3* erstreckt, weist jedoch eine Hypermethylierung des maternalen Allels auf. Diese Region wird als *Peg3*-DMR bezeichnet (Li *et al.*, 2000). Weitere differentiell methylierte Regionen konnten nicht identifiziert werden (Choo *et al.*, 2006). In einer Folgestudie wurde eine gehäufte Bindung von macroH2A1 am methylierten Allel der *Peg3*-DMR nachgewiesen (Choo *et al.*, 2006). Dabei handelt es sich um eine seltene Variante des H2-Kernhistons. Bei Hefen agiert die macro-Domäne von macroH2A1 agiert durch die Bindung an DNA als transkriptioneller Repressor (Perche *et al.*, 2000). Eine vergleichbare Wirkung ist bei Säugern anzunehmen. Die Assoziation von macroH2A1 zum jeweils methylierten Allel der DMR in den *Xist*-, *H19*/*Igf2*-, *Gtl2*/*Dlk1*- und *Gnas*-Domänen bestätigt, dass die Interaktion von DNA-Methylierung und macroH2A1 eine Schlüsselrolle bei der transkriptionellen Inaktivierung

geprägter Gene spielt. Wie die epigenetische Regulation der geprägten Gene im *Peg3*-Locus mechanistisch genau funktioniert, bleibt weiterhin unklar. Mit der Bindung von macroH2A1 an die *Peg3*-DMR kann zwar die transkriptionelle Inaktivierung des maternalen Allels von *Peg3*, *Peg14* und *Usp29* bei der Maus sowie von *PEG3*, *ZIM2* und *PEG14* beim Menschen erklärt werden, aber um die transkriptionelle Inaktivierung der paternalen Allele der murinen Gene *Zim1*, *Zim2*, *Zim3* sowie des maternalen Allels von *Zfp264* zu erklären, fehlt bisher noch ein wichtiges Puzzleteil. Es ist denkbar, dass das murine *Peg14* dieses fehlende Puzzleteil ist und eine regulatorische bzw. vermittelnde Rolle spielt. Es könnten aber auch andere Faktoren eine regulatorische Funktion ausüben. Wie bereits erwähnt, ist anzunehmen, dass die geprägte Expression von *Zim1*, *Zim2*, *Zim3* und *Zfp264* eine biologische Funktion erfüllt. Vermutlich spielen diese nicht-kodierenden RNAs wiederum eine Rolle bei der Regulation anderer Gene.

Eine weitere Hypothese lässt sich aus der vermutlich nachträglichen Insertion von *Zim1* ableiten (Kim *et al.*, 1999). So ist es vorstellbar, dass die genomische Prägung von *Zim2*, *Zim3* und *Zfp264* erst durch die Insertion von *Zim1* und die damit verbundene chromosomale Umstrukturierung entstanden sein könnte. Auf dem humanen Chromosom fehlt das orthologe Gen, was die biallelische Expression von *ZIM3* und *ZNF264* zur Folge haben könnte. Bei *USP29* lässt sich die biallelische Expression dadurch erklären, dass das murine und humane Gen verschiedene Promotoren haben. Während der murine *Usp29*-Promotor innerhalb der *Peg3*-DMR liegt, ist eine Interaktion zwischen der *PEG3*-DMR und dem *USP29*-Promotor auf dem humanen Chromosom durch die Distanz von über 200 kb eher auszuschließen. Darüber hinaus lässt sich aus der geringen Homologie von ca. 65% zwischen dem *Usp29*-ORF und seinem humanen Ortholog schlussfolgern, dass sich beide Proteine funktionell unterscheiden.

4.2.4. Zusammenfassende Bewertung

Mit *Peg14* konnte das siebte geprägte Gen in der Imprinting-Region auf MMU7prox identifiziert werden. Der genaue Mechanismus der geprägten Expression aller sieben Gene ist bisher nicht bekannt. Durch die Interaktion zwischen dem methylierten Allel der *Peg3*-DMR und macroH2A1 (Choo *et al.*, 2006) kann zumindest die transkriptionelle Inaktivierung der maternalen Allele von *Peg3*, *Peg14* und *Usp29* bei der Maus sowie von *PEG3*, *ZIM2* und *PEG14* beim Menschen erklärt werden. Die orthologe Region beim Menschen auf HSA19q13.4 weist neben zahlreichen Gemeinsamkeiten wie der Konservierung der Gene mit Ausnahme von *Zim1* auch Unterschiede in der genomischen Organisation auf, die wiederum mit Unterschieden in der allelischen Expression korrelieren. Ein oder mehrere entscheidende Faktoren, die auf MMU7prox für die geprägte Expression von *Usp29*, *Zim3*

und *Zfp264* verantwortlich sind, fehlen demnach auf HSA19q13.4 oder haben im Laufe der Evolution ihre regulatorische Funktion wieder verloren. Im ersten Fall könnte das Gen *Zim1* zumindest ein vermittelnder Faktor sein, welcher auf dem humanen Chromosom fehlt. Im zweiten Fall hingegen kommt für das Gen *Peg14* eine regulatorische Funktion in Frage.

Mit beiden oben genannten Modellen (Reik *et al.*, 2004) lässt sich die geprägte Expression weder in der murinen noch in der humanen *Peg3*-Domäne vollständig erklären. Dennoch konnte auch für die Imprinting-Region mit der Interaktion zwischen DNA-Methylierung und Bestandteilen des Chromatins ein epigenetisches Netzwerk zumindest teilweise identifiziert werden.

4.3. Sequenzbasierte Vorhersage neuer geprägter Gene

In vergleichenden Sequenzanalysen konnte gezeigt werden, dass sich der genomische Kontext geprägter Gene signifikant von nicht geprägten Genen unterscheidet und dass die spezifischen Sequenzmerkmale von Mensch und Maus konserviert sind (Onyango *et al.*, 2000; Greally, 2002; Ke *et al.*, 2002; Allen *et al.*, 2003; Smith *et al.*, 2004). In der Studie von Luedi *et al.* (2005) wurden diese Sequenzmerkmale erstmalig benutzt, um neue geprägte Gene im Genom der Maus vorherzusagen. Die Anzahl der ursprünglich 600 ermittelten neuen Kandidatengene wurde durch eine Priorisierung auf 18 eingeschränkt und 16 davon mittels QUASEP analysiert. Dabei wurden zum einen die maternale Expression von *Kcnk9* und zum anderen die präferentielle Expression des Cast/Ei-Allels von *Rarres1* identifiziert. In nachfolgenden Experimenten konnte auch beim humanen *KCNK9*-Gen monoallelisch maternale Expression nachgewiesen werden. Die stammspezifische Expression von *Rarres1* wird vermutlich durch *cis*-regulatorische Faktoren wie DNA-Polymorphismen und/oder DNA-Methylierung in der flankierenden Sequenz hervorgerufen (siehe auch Punkt 4.1.5). Die geprägte Expression von *Kcnk9* und dessen humanem Ortholog wirft unterdessen zahlreiche Fragen auf, da *Kcnk9* eine wichtige Rolle bei Apoptose und Tumorgenese spielt ((Mu *et al.*, 2003; Pei *et al.*, 2003; Patel und Lazdunski, 2004), siehe auch Kapitel 4.3.2).

4.3.1. Die Effizienz von sequenzbasierten Vorhersagen

Mehrere Studien belegen die spezifischen Sequenzmerkmale geprägter Gene. So wurde in einer vergleichenden Sequenzanalyse von 24 humanen und murinen geprägten Genen 16 Consensus-Sequenzen identifiziert, die bei der Suche nach neuen geprägten

Genen hilfreich sein könnten (Wang *et al.*, 2004). In einer weiteren genomweiten Studie wurden die Methylierungsmuster von CpG-Inseln untersucht, um neue differentiell methylierte Regionen zu identifizieren (Strichman-Almashanu *et al.*, 2002). Von insgesamt 62 untersuchten Regionen wurde bei 12 CpG-Inseln allelspezifische Methylierung nachgewiesen. Beide Analysen bestätigen zwar die Signifikanz der spezifischen Sequenzmerkmale geprägter Regionen, verweisen jedoch nicht auf Kandidatengene für eine nachfolgende Validierung. Im Gegensatz dazu wurden in der Studie von Luedi *et al.* (2005) erstmalig Imprinting-Kandidatengene im Genom der Maus vorhergesagt. Unter den vielversprechendsten 16 untersuchten Kandidatengenen wurde mit *Kcnk9* ein neues geprägtes Gen identifiziert. Dies entspricht einem prozentualen Anteil von ca. 6%. Im Vergleich dieser 16 Kandidatengene weist *Kcnk9* die höchsten Werte hinsichtlich Priorität und Algorithmus auf (siehe Tab. 10) und damit auch die höchste Wahrscheinlichkeit hinsichtlich Prägung. Unter der Voraussetzung, dass die berechneten Werte von *Kcnk9* und die tatsächliche Prägung nicht zufällig zusammen auftreten, ist der Ansatz von Luedi *et al.* (2005) grundsätzlich als richtig zu bewerten. Die übrigen untersuchten 15 Kandidatengene zeigten in den F1-Hybriden (E11.5) keine geprägte Expression und weisen in den Berechnungen auch niedrigere Werte auf.

Um die Effizienz der Studie von Luedi *et al.* (2005) bewerten zu können, muss jedoch berücksichtigt werden, dass die getesteten Kandidatengene nicht zufällig, sondern durch eine Priorisierung ausgewählt wurden. Da die nicht getesteten 584 Kandidatengene eine geringere Wahrscheinlichkeit aufweisen, tatsächlich geprägt zu sein, kann angenommen werden, dass diese ebenfalls biallelisch exprimiert sind. Dadurch würde der Anteil der tatsächlich geprägten Gene von 6% auf 0,167% sinken. Der gewählte Ansatz wäre somit ineffizient. Dies gilt insbesondere, wenn man berücksichtigt, dass die bisher identifizierten geprägten Gene ca. 0,3% der Gesamtzahl der Gene ausmachen.

Ein kritischer Punkt aller bisherigen sequenzbasierten Studien ist die geringe Zahl an bekannten geprägten Genen, die zur Identifizierung der spezifischen Sequenzmuster verwendet wurde. So beschränkten Onyango *et al.* (2000) ihre Analyse auf die Beckwith-Wiedemann-Syndrom-Region, während Greally (2002) und Ke *et al.* (2002) jeweils 31 bekannte geprägte Gene für ihre Untersuchung einsetzten. Auch bei Allen *et al.* (2003) wurden mit 39 und bei Luedi *et al.* (2005) mit 44 nicht alle zum damaligen Zeitpunkt bekannten geprägte Gene für die Analyse eingesetzt. Es sollte zwar grundsätzlich möglich sein, mit einer unvollständigen Anzahl geprägter Gene eine Identifizierung der spezifischen Sequenzmuster durchzuführen, ob daraus folgend jedoch die signifikanten Schwellenwerte für einzelne Sequenzmerkmale richtig ermittelt werden können, ist aufgrund der Ergebnisse der vorliegenden Arbeit zu bezweifeln.

4.3.2. Genomische Prägung von *Kcnk9*

Das hier erstmalig als geprägt beschriebene *Kcnk9*-Gen gehört zur Familie der Kaliumkanalgene mit zwei tandemartig angeordneten transmembranösen Poren (K2P) und kodiert für das Protein TASK-3. Die biochemischen Eigenschaften von TASK-3 sind bereits beschrieben, eine detaillierte funktionelle Charakterisierung fehlt jedoch. Die Gene der K2P-Familie sind scheinbar an der Ausbildung des Ruhepotentials von verschiedenen neuronalen Zellen beteiligt, wobei auch die Funktionen der anderen Gene dieser Familie nicht genau bekannt sind (Patel und Lazdunski, 2004). Die starke Expression von TASK-3 in zahlreichen Regionen des Gehirns weist auf eine Funktion im zentralen Nervensystem hin (Chapman *et al.*, 2000; Medhurst *et al.*, 2001). Vor diesem Hintergrund ist die genomische Prägung von *Kcnk9* ein wichtiger Befund, der zahlreiche Fragen hinsichtlich der Bedeutung der geprägten Expression von *Kcnk9* bei der Entwicklung des Gehirns aufwirft. Neben der Bedeutung geprägter Gene in der Embryonalentwicklung, beim Wachstum der Plazenta oder bei postnatalen Anpassungsprozessen (Constancia *et al.*, 2004) gibt es auch Hinweise, dass geprägte Gene bei der Entwicklung des Gehirns eine entscheidende Rolle spielen. Eine spezifische Funktion bei der Gehirnentwicklung, die sich eindeutig von anderen Entwicklungsprozessen unterscheiden lässt, konnte jedoch bislang noch für kein geprägtes Gen nachgewiesen werden (Davies *et al.*, 2005). Mit *Kcnk9* könnte bei einer ausschließlichen Expression im murinen Gehirn die Auswirkung eines geprägten Gens auf die Entwicklung des Gehirns erstmalig untersucht werden.

Neben den obigen Fragen, die durch die geprägte *Kcnk9*-Expression ergeben, gibt es auch zahlreiche Hinweise, dass *Kcnk9* an der Entstehung von Krankheiten beteiligt ist. So wurde bei einem Rattenstamm, der ein etabliertes Tiermodell für die Absence-Epilepsie darstellt, eine pathogene Mutation in *Kcnk9* nachgewiesen (Holter *et al.*, 2005). Des Weiteren wird *Kcnk9* eine Rolle bei Apoptose und Tumorgenese zugeschrieben (Patel und Lazdunski, 2004). Das humane *KCNK9*-Gen weist eine genomische Amplifikation und Überexpression bei einer signifikanten Anzahl von Mammakarzinomen auf (Mu *et al.*, 2003; Pei *et al.*, 2003). Möglicherweise ist die Überexpression von *KCNK9* auch an der Entstehung von Darmkrebs beteiligt (Kim *et al.*, 2004a). Die zellulären Mechanismen sind zwar bisher unbekannt, aber es wird vermutet, dass die Aktivität des Kaliumkanals an der Zellzyklus-Kontrolle beteiligt ist. Bei verschiedenen Zellarten konnte bereits nachgewiesen werden, dass eine Aktivierung der Kaliumkanäle zur Proliferation führt. Es ist möglich, dass aktive Kaliumkanäle notwendig sind, um eine bestimmte Phase des Zellzyklus zu durchlaufen. Eventuell beeinflusst der Ausstrom von Kalium-Ionen auch intrazelluläre Calcium-Signalwege. Die Veränderung des intrazellulären Calciumspiegels hat eine signifikante

Auswirkung auf Differenzierung und Proliferation einer Zelle. Es ist denkbar, dass die Überexpression von TASK-3 zu einem gesteigerten Ausstrom vom Kalium-Ionen führt, der wiederum einen erhöhten Einstrom von Calcium-Ionen bewirkt und dadurch das Zellvolumen und die Proliferationsrate steigert (Patel und Lazdunski, 2004).

Ein Vergleich der Expressionsmuster des murinen und humanen *Kcnk9*-Gens im adulten Gehirn zeigt signifikante Unterschiede. Während die Expression des humanen *KCNK9* im Cerebellum am stärksten ist und weitere Bereiche des Gehirns lediglich schwache Expression zeigen (Medhurst *et al.*, 2001), lässt sich beim murinen *Kcnk9* die stärkste Expression in Medulla oblongata, Rückenmark, Hirnstamm, Cerebellum und Hypothalamus nachweisen (Talley *et al.*, 2001). Ein Vergleich der Aminosäuresequenzen mittels BLAST 2 Sequences ergibt nur 75%ige Homologie zwischen den orthologen Proteinen (humanes KCNK9: NP_057685.1 und murines Kcnk9: NP_001029048.1). Trotz übereinstimmender maternaler Expression der orthologen Gene lassen die unterschiedlichen Expressionsmuster und die geringe Homologie der Proteinsequenzen auf funktionelle Unterschiede schließen.

In einer Methylierungsanalyse wird zu zeigen sein, ob die Unterschiede im Expressionsmuster auf unterschiedliche regulatorische Mechanismen zurückzuführen sind. So ist beim murinen *Kcnk9*-Gen denkbar, dass die von Smith *et al.* (2003) identifizierte Methylierung des maternalen Allels, die mit dem benachbarten *Peg13*-Gen assoziiert ist, auch in die Imprinting-Regulation von *Kcnk9* involviert ist. *Peg13* ist wie auch *Kcnk9* im Gehirn stark exprimiert (Davies *et al.*, 2004), was zusätzlich auf eine koordinierte Regulation schließen lässt. Im Gegensatz dazu konnte bisher kein humanes Ortholog von *Peg13* identifiziert werden. Es erscheint daher notwendig, beide Regionen nach weiteren geprägten Genen und regulatorischen Elementen zu untersuchen, um den Mechanismus der geprägten Expression aufzuklären.

4.3.3. Zusammenfassende Bewertung

Die Ergebnisse zahlreicher Studien belegen spezifische Sequenzmerkmale geprägter Gene, die sich von nicht geprägten Genen signifikant unterscheiden. Trotzdem konnte bisher keine Methode etabliert werden, um effizient neue Imprinting-Kandidatengene zu identifizieren.

Durch die Überprüfung von 16 der ursprünglich 600 Kandidatengene von (Luedi *et al.*, 2005) konnte die maternale Expression von *Kcnk9* bei Maus und Mensch nachgewiesen werden. Die Effizienz der Methode ist jedoch als gering einzuschätzen. Da *Kcnk9* von den ausgewählten Kandidatengenen als am wahrscheinlichsten geprägt galt, kann vermutet werden, dass durch strengere Signifikanzkriterien die Anzahl der vorhergesagten

Kandidatengene niedriger gewesen und damit die Effektivität der Methode erhöht worden wäre. Des Weiteren wäre bei zukünftigen Studien wichtig, eine wesentlich höhere Anzahl bekannter geprägter Gene für die vergleichende Sequenzanalyse einzusetzen, um die Signifikanz einzelner Sequenzmerkmale besser abschätzen zu können. Dadurch sollte vor allem die Anzahl falsch positiver Kandidatengene verringert werden können.

Insgesamt ist der methodische Ansatz als vielversprechend zu bewerten, da ohne hohen laborexperimentellen Aufwand genomweit neue Imprinting-Kandidatengene identifiziert werden können. Die sequenzbasierte Vorhersage stellt somit eine positive Ergänzung zu den zuvor erläuterten Ansätzen dar.

4.4. Vergleich beider Ansätze

In der vorliegenden Arbeit wurden zwei Gruppen von Imprinting-Kandidatengenen mittels QUASEP hinsichtlich tatsächlicher Prägung validiert. In beiden Gruppen wurde mit *Peg14* und *Kcnk9* jeweils ein neues geprägtes Gen identifiziert. Die Methoden zur experimentellen Identifizierung von potentiell geprägten Genen sind im Vergleich zu den sequenzbasierten Ansätzen wesentlich weiter entwickelt und waren bisher auch erfolgreicher. Mit *Kcnk9* wurde hier erstmalig die geprägte Expression eines Kandidatengens aus einer sequenzbasierten Studie bestätigt. In zukünftigen Untersuchungen gilt es zu bestätigen, ob geprägte Gene durch *in silico*-Analysen identifiziert werden können, oder ob die Vorhersage von Luedi *et al.* (2005) nur zufällig zur Identifizierung eines neuen geprägten Gens geführt hat. Neben den im Vergleich zu experimentellen Verfahren geringeren finanziellen Kosten kann mit diesem methodischen Ansatz genomweit in verschiedenen Spezies nach neuen Kandidatengenen gesucht werden. Momentan erscheint der sequenzbasierte Ansatz zwar vielversprechend, seine Tauglichkeit und Effizienz kann jedoch noch nicht beurteilt werden.

Der größte Teil der Kandidatengene von Nikaido *et al.* (2003) ist in F1-Hybriden (E11.5) biallelisch exprimiert. Ein sehr hoher Anteil der von Nikaido definierten Kandidatengene basiert somit nicht auf genomischer Prägung, sondern eher auf generellen Expressionsunterschieden zwischen parthenogenetischen und androgenetischen Mausembryonen, die durch die großen phänotypischen Unterschiede verursacht werden. Daher ist Verwendung von uniparentalen Embryonen in Kombination mit cDNA-Microarrays zur Identifizierung neuer geprägter Gene als ineffiziente Methode zu bewerten.

Mit den gewählten Versuchsbedingungen ist es weder Nikaido *et al.* (2003) noch Luedi *et al.* (2005) gelungen, neue geprägte Gene effizient zu identifizieren. Dennoch

erscheinen beide Methoden grundsätzlich geeignet und könnten als Kombination vermutlich einen Ansatz mit höherer Effizienz darstellen.

4.5. Eignung von QUASEP zur Imprinting-Analyse

Um die geprägte Expression der Imprinting-Kandidatengene einschließlich der drei Kontrollgene (*Zim1*, *Slc38a4* und AK014392) mittels QUASEP zu validieren, wurden 115 verschiedene Primer-Sets verwendet. Die dabei erzielten Ergebnisse belegen die Eignung von QUASEP zur allelspezifischen Expressionsanalyse. Mit QUASEP ist es möglich, selbst geringe allelische Expressionsunterschiede zu detektieren. Die Genauigkeit und Zuverlässigkeit der Methode wird durch die durchschnittlichen Standardabweichungen (3,7 bei den Nikaido-Kandidaten und 2,8 bei den Luedi-Kandidaten) demonstriert. Die höchste Standardabweichung wurde mit 11,6 und die niedrigste mit 0,7 berechnet. Diese Zahlen korrelieren mit der beobachteten Abhängigkeit von der umgebenden DNA-Sequenz. Je genauer sich die Allele unterscheiden lassen und je geringer die Hintergrundsignale ausfallen, umso kleiner fällt die Streuung der Allelfrequenzen und damit die Standardabweichung aus. Um dies zu erreichen, sollten drei Voraussetzungen erfüllt sein. Erstens sollte der untersuchte SNP nicht zu einem Homopolymer gehören, weil dadurch eine genaue Quantifizierung erschwert oder verhindert wird. Zweitens sollte die PCR sorgfältig optimiert werden, um ein spezifisches PCR-Produkt mit hoher Ausbeute zu gewinnen. Drittens sollten keine SNPs verwendet werden, die mit repetitiven Regionen assoziiert sind, da sich sonst Sekundärstrukturen oder eine unspezifische Hybridisierung des Sequenzierprimers und damit Hintergrundsignale ergeben können.

Unter Berücksichtigung der oben genannten Empfehlungen bietet QUASEP im Vergleich zu anderen experimentellen Methoden zahlreiche Vorteile bei der Imprinting-Analyse. So lassen sich monoallelische und präferentielle allelische Expression genau voneinander unterscheiden und quantifizieren. Damit kann beispielsweise eine Relaxation der Prägung wie bei AK039482 (*Lit1*) im adulten Stadium identifiziert werden, was mit bisher beschriebenen Methoden nicht oder nur sehr ungenau möglich war. Unter der Bedingung, dass ein geeigneter tSNP identifiziert wurde, stellt QUASEP insbesondere aufgrund seiner Genauigkeit bei der Untersuchung geprägter Expression die Methode der Wahl dar.

Neben dem Einsatz zur Imprinting-Analyse ist QUASEP generell geeignet, um allelische Imbalancen zu untersuchen. So belegen zahlreiche aktuelle Studien, dass es genauer und robuster Methoden zur allelischen Quantifizierung bedarf (Cowles *et al.*, 2002; Bray *et al.*, 2003; Lo *et al.*, 2003). In der Studie von Lo *et al.* (2003) konnten mittels Affymetrix HuSNP arrays in Kombination mit quantitativer Real-time-PCR nur zweifache

Expressionsunterschiede detektiert werden. Auch Cowles *et al.* (2002) waren mit dem gewählten quantitativen Genotypisierungsansatz nicht in der Lage, allelische Expressionsunterschiede unter 50% zu analysieren.

4.6. Ausblick

Da die Anzahl geprägter Gene nach wie vor eine offene Frage ist, wird die Suche nach neuen geprägten Genen weiterhin das Ziel zahlreicher Studien sein. Die Ergebnisse der vorliegenden Arbeit sind für diese zukünftigen Studien, die mit vergleichbarer Methodik wie Nikaido *et al.* (2003) oder Luedi *et al.* (2005) neue Imprinting-Kandidatengene identifizieren wollen, von großer Bedeutung. Das bedeutet zum einen, dass, wie von Oakey *et al.* (2002) vorgeschlagen, eine effektivere Suche nach neuen geprägten Genen möglich sein könnte, wenn Kandidatengene aus bisher publizierten Studien verglichen werden würden. Zum anderen stellt ein verbesserter sequenzbasierter Ansatz, der auf einer neuen Bewertung der spezifischen Sequenzmerkmale beruht, eine positive Erweiterung des methodischen Spektrums dar und könnte in einer Kombination mit laborexperimentellen Ansätzen eine Möglichkeit bieten, um neue geprägte Gene effektiv zu identifizieren.

Für allelspezifische Expressionsanalysen stellt die hier beschriebene QUASEP-Methode eine wesentlich genauere und gleichzeitig robuste Methode dar, mit der die geprägte Expression exakt analysiert werden kann. So könnte in einer Folgestudie mittels QUASEP eine umfassende Imprinting-Analyse von geprägten Regionen, wie beispielsweise der *Peg3*-Domäne auf MMU7prox, durchgeführt werden, um die gewebespezifische Prägung der einzelnen Gene zu untersuchen und dadurch ein differenziertes Bild der koordinierten Regulation in einer geprägten Domäne zu erhalten.

Die in dieser Arbeit validierte genomische Prägung von *Kcnk9* und *Peg14* lässt viele Fragen offen. Eine weitere Charakterisierung auf funktioneller Ebene ist notwendig, und die daraus zu gewinnenden Erkenntnisse können einen wichtigen Beitrag leisten, um ein besseres Verständnis der geprägten Genexpression zu erreichen. Es gilt dabei aufzuklären, welche Rolle die geprägte Expression der beiden Gene bei Embryonalentwicklung, Wachstum der Plazenta, postnatalen Anpassungsprozessen und Gehirnentwicklung spielt. Von besonderem Interesse ist die Frage, inwieweit die geprägte Expression von *Peg14* und *Kcnk9* in den Kontext bestehender Theorien zur Evolution der genomischen Prägung passt.

5. Zusammenfassung

Bislang ist die Anzahl der geprägten Gene nicht bekannt, daher sind systematische Analysen zur genomweiten Identifizierung neuer geprägter Gene notwendig. In zwei aktuellen Studien wurden mit zwei verschiedenen methodischen Ansätzen 2101 bzw. 600 neue geprägte Kandidatengene im Mausgenom vorhergesagt, die noch durch eine allelspezifische Expressionsanalyse validiert werden mussten. Das primäre Ziel der vorliegenden Arbeit war die Identifizierung neuer geprägter Gene durch die systematische Analyse vielversprechender Kandidatengene aus beiden Studien.

Zunächst wurde mit QUASEP eine neue Methode zur allelspezifischen Expressionsanalyse etabliert, mit der es möglich ist, selbst geringe allelische Expressionsunterschiede genau und reproduzierbar zu analysieren. Damit stellt QUASEP im Vergleich zu anderen Anwendungen die Methode der Wahl dar.

Durch die systematische QUASEP-Analyse von ausgewählten Kandidatengenen wurde erstmalig die paternale Expression von *Peg14* auf dem proximalen Maus-Chromosom 7 sowie die maternale Expression von *Kcnk9* auf dem distalen Maus-Chromosom 15 nachgewiesen. Da die große Mehrzahl der untersuchten Gene jedoch biallelisch exprimiert ist, muss die Effizienz der methodischen Ansätze beider Studien als gering eingeschätzt werden. Dies sollte in zukünftigen Studien berücksichtigt werden, bei denen mit vergleichbarer Methodik neue geprägte Kandidatengene identifiziert werden sollen.

Neben den zwei neuen geprägten murinen Genen konnte in der humanen orthologen Region von *Peg14* auf dem Chromosom 19q13.4 ein bisher unbekanntes Transkript mit einer Länge von 3940 bp und monoallelischer Expression im humanen fetalen Gehirn identifiziert werden. Des Weiteren wurde durch die QUASEP-Analyse des humanen *KCNK9*-Gens monoallelische Expression in fetalem Gehirn nachgewiesen. Die Prägung von *Peg14* und *Kcnk9* und deren menschlichen Orthologen lässt unterdessen zahlreiche Fragen offen. Während bei dem vermutlich nicht-kodierenden *Peg14* eine Funktion bei der Imprinting-Regulation der benachbarten geprägten Gene denkbar ist, wird *Kcnk9* eine Funktion bei Apoptose und Tumorgenese zugeschrieben. Welche Rolle die genomische Prägung von *Kcnk9* bei beiden Prozessen spielt, muss in zukünftigen Studien geklärt werden.

6. Literaturverzeichnis

Ahmadian, A., Ehn, M. und Hober, S., Pyrosequencing: history, biochemistry and future, Clin Chim Acta 363 (2006) 83-94.

Albrecht, U., Sutcliffe, J.S., Cattanach, B.M., Beechey, C.V., Armstrong, D., Eichele, G. und Beaudet, A.L., Imprinted expression of the murine Angelman syndrome gene, Ube3a, in hippocampal and Purkinje neurons, Nat Genet 17 (1997) 75-8.

Allen, E., Horvath, S., Tong, F., Kraft, P., Spiteri, E., Riggs, A.D. und Marahrens, Y., High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes, Proc Natl Acad Sci U S A 100 (2003) 9940-5.

Allen, N.D., Logan, K., Lally, G., Drage, D.J., Norris, M.L. und Keverne, E.B., Distribution of parthenogenetic cells in the mouse brain and their influence on brain development and behavior, Proc Natl Acad Sci U S A 92 (1995) 10782-6.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. und Lipman, D.J., Basic local alignment search tool, J Mol Biol 215 (1990) 403-10.

Barlow, D.P., Methylation and imprinting: from host defense to gene regulation? Science 260 (1993) 309-10.

Barlow, D.P., Gametic imprinting in mammals, Science 270 (1995) 1610-3.

Bartolomei, M.S., Zemel, S. und Tilghman, S.M., Parental imprinting of the mouse H19 gene, Nature 351 (1991) 153-5.

Barton, S.C., Surani, M.A. und Norris, M.L., Role of paternal and maternal genomes in mouse development, Nature 311 (1984) 374-6.

Bird, A., DNA methylation patterns and epigenetic memory, Genes Dev 16 (2002) 6-21.

Brambilla, R., Gnesutta, N., Minichiello, L., White, G., Roylance, A.J., Herron, C.E., Ramsey, M., Wolfer, D.P., Cestari, V., Rossi-Arnaud, C., Grant, S.G., Chapman, P.F., Lipp, H.P., Sturani, E. und Klein, R., A role for the Ras signalling pathway in synaptic transmission and long-term memory, Nature 390 (1997) 281-6.

Bray, N.J., Buckland, P.R., Owen, M.J. und O'Donovan, M.C., Cis-acting variation in the expression of a high proportion of genes in human brain, Hum. Genet. 113 (2003) 149-153.

Caspary, T., Cleary, M.A., Baker, C.C., Guan, X.J. und Tilghman, S.M., Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster, Mol Cell Biol 18 (1998) 3466-74.

Chapman, C.G., Meadows, H.J., Godden, R.J., Campbell, D.A., Duckworth, M., Kelsell, R.E., Murdock, P.R., Randall, A.D., Rennie, G.I. und Gloger, I.S., Cloning, localisation and functional expression of a novel human, cerebellum specific, two pore domain potassium channel, Brain Res Mol Brain Res 82 (2000) 74-83.

Charalambous, M., Smith, F.M., Bennett, W.R., Crew, T.E., Mackenzie, F. und Ward, A., Disruption of the imprinted Grb10 gene leads to disproportionate overgrowth by an Igf2-independent mechanism, Proc Natl Acad Sci U S A 100 (2003) 8292-7.

Chen, M., Gavrilova, O., Liu, J., Xie, T., Deng, C., Nguyen, A.T., Nackers, L.M., Lorenzo, J., Shen, L. und Weinstein, L.S., Alternative Gnas gene products have opposite effects on glucose and lipid metabolism, Proc Natl Acad Sci U S A 102 (2005) 7386-91.

Choo, J.H., Kim, J.D., Chung, J.H., Stubbs, L. und Kim, J., Allele-specific deposition of macroH2A1 in imprinting control regions, Hum Mol Genet 15 (2006) 717-24.

Constancia, M., Kelsey, G. und Reik, W., Resourceful imprinting, Nature 432 (2004) 53-7.

Constancia, M., Angiolini, E., Sandovici, I., Smith, P., Smith, R., Kelsey, G., Dean, W., Ferguson-Smith, A., Sibley, C.P., Reik, W. und Fowden, A., Adaptation of nutrient supply to fetal demand in the mouse involves interaction between the Igf2 gene and placental transporter systems, Proc Natl Acad Sci U S A 102 (2005) 19219-24.

Cowles, C.R., Hirschhorn, J.N., Altshuler, D. und Lander, E.S., Detection of regulatory variation in mouse genes, Nat. Genet. 32 (2002) 432-437.

Curley, J.P., Barton, S., Surani, A. und Keverne, E.B., Coadaptation in mother and infant regulated by a paternally expressed imprinted gene, Proc Biol Sci 271 (2004) 1303-9.

Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E. und Iozzo, R.V., Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility, J Cell Biol 136 (1997) 729-43.

Davies, W., Smith, R.J., Kelsey, G. und Wilkinson, L.S., Expression patterns of the novel imprinted genes Nap1l5 and Peg13 and their non-imprinted host genes in the adult mouse brain, Gene Expr Patterns 4 (2004) 741-7.

Davies, W., Isles, A.R. und Wilkinson, L.S., Imprinted gene expression in the brain, Neurosci Biobehav Rev 29 (2005) 421-30.

DeChiara, T.M., Robertson, E.J. und Efstratiadis, A., Parental imprinting of the mouse insulin-like growth factor II gene, Cell 64 (1991) 849-59.

Doerfler, W., On the biological significance of DNA methylation, Biochemistry (Mosc) 70 (2005) 505-24.

Duvillie, B., Cordonnier, N., Deltour, L., Dandoy-Dron, F., Itier, J.M., Monthioux, E., Jami, J., Joshi, R.L. und Bucchini, D., Phenotypic alterations in insulin-deficient mutant mice, Proc Natl Acad Sci U S A 94 (1997) 5137-40.

Farrall, M. und Morris, A.P., Gearing up for genome-wide gene-association studies, Hum Mol Genet 14 Spec No. 2 (2005) R157-62.

Fitzpatrick, G.V., Soloway, P.D. und Higgins, M.J., Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1, Nat. Genet. 32 (2002) 426-431.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y.Z., Plass, C. und Esteller, M., Epigenetic differences arise during the lifetime of monozygotic twins, Proc Natl Acad Sci U S A 102 (2005) 10604-9.

Frohman, M.A., Dush, M.K. und Martin, G.R., Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer, Proc Natl Acad Sci U S A 85 (1988) 8998-9002.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P. und Kouzarides, T., The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation, J Biol Chem 278 (2003) 4035-40.

Giulietti, A., Overbergh, L., Valckx, D., Decallonne, B., Bouillon, R. und Mathieu, C., An overview of real-time quantitative PCR: applications to quantify cytokine gene expression, Methods 25 (2001) 386-401.

Greally, J.M., Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome, Proc Natl Acad Sci U S A 99 (2002) 327-32.

Gregory, S.G., et al., The DNA sequence and biological annotation of human chromosome 1, Nature 441 (2006) 315-21.

Grover, D., Mukerji, M., Bhatnagar, P., Kannan, K. und Brahmachari, S.K., Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition, Bioinformatics 20 (2004) 813-7.

Guillemot, F., Caspary, T., Tilghman, S.M., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Anderson, D.J., Joyner, A.L., Rossant, J. und Nagy, A., Genomic imprinting of Mash2, a mouse gene required for trophoblast development, Nat Genet 9 (1995) 235-42.

Hagiwara, Y., Hirai, M., Nishiyama, K., Kanazawa, I., Ueda, T., Sakaki, Y. und Ito, T., Screening for imprinted genes by allelic message display: identification of a paternally expressed gene impact on mouse chromosome 18, Proc Natl Acad Sci U S A 94 (1997) 9249-54.

Haig, D., Genomic imprinting and kinship: how good is the evidence? Annu Rev Genet 38 (2004) 553-85.

Hashimshony, T., Zhang, J., Keshet, I., Bustin, M. und Cedar, H., The role of DNA methylation in setting up chromatin structure during development, Nat Genet 34 (2003) 187-92.

Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M. und et al., Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method, Nat Genet 6 (1994) 33-40.

Hirschhorn, J.N. und Daly, M.J., Genome-wide association studies for common diseases and complex traits, Nat Rev Genet 6 (2005) 95-108.

Holter, J., Carter, D., Leresche, N., Crunelli, V. und Vincent, P., A TASK3 channel (KCNK9) mutation in a genetic model of absence epilepsy, J Mol Neurosci 25 (2005) 37-51.

Horsthemke, B., Was ist Epigenetik? Medizinische Genetik 17 (2005) 251-253.

Horsthemke, B. und Buiting, K., Imprinting defects on human chromosome 15, Cytogenet Genome Res 113 (2006) 292-9.

Jenuwein, T. und Allis, C.D., Translating the histone code, Science 293 (2001) 1074-80.

Jiang, Y.H., Armstrong, D., Albrecht, U., Atkins, C.M., Noebels, J.L., Eichele, G., Sweatt, J.D. und Beaudet, A.L., Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation, Neuron 21 (1998) 799-811.

Jones, B.K., Levorse, J.M. und Tilghman, S.M., Igf2 imprinting does not require its own DNA methylation or H19 RNA, Genes Dev 12 (1998) 2200-7.

Jonkers, J., van Amerongen, R., van der Valk, M., Robanus-Maandag, E., Molenaar, M., Destree, O. und Berns, A., In vivo analysis of Frat1 deficiency suggests compensatory activity of Frat3, Mech Dev 88 (1999) 183-94.

Kanduri, C., Thakur, N. und Pandey, R.R., The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing, Embo J 25 (2006) 2096-106.

Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S.C., Ishino, F. und Surani, M.A., Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization, Nat. Genet. 11 (1995) 52-59.

Kass, D.H., Raynor, M.E. und Williams, T.M., Evolutionary history of B1 retroposons in the genus Mus, J Mol Evol 51 (2000) 256-64.

Ke, X., Thomas, N.S., Robinson, D.O. und Collins, A., The distinguishing sequence characteristics of mouse imprinted genes, Mamm Genome 13 (2002) 639-45.

Kelsey, G., Bodle, D., Miller, H.J., Beechey, C.V., Coombes, C., Peters, J. und Williamson, C.M., Identification of imprinted loci by methylation-sensitive representational difference analysis: application to mouse distal chromosome 2, Genomics 62 (1999) 129-38.

Kent, W.J., BLAT--the BLAST-like alignment tool, Genome Res 12 (2002) 656-64.

Keverne, E.B., Genomic imprinting in the brain, Curr Opin Neurobiol 7 (1997) 463-8.

Kile, B.T., Metcalf, D., Mifsud, S., DiRago, L., Nicola, N.A., Hilton, D.J. und Alexander, W.S., Functional analysis of Asb-1 using genetic modification in mice, Mol Cell Biol 21 (2001) 6189-97.

Kim, C.J., Cho, Y.G., Jeong, S.W., Kim, Y.S., Kim, S.Y., Nam, S.W., Lee, S.H., Yoo, N.J., Lee, J.Y. und Park, W.S., Altered expression of KCNK9 in colorectal cancers, Apmis 112 (2004a) 588-94.

Kim, J., Ashworth, L., Branscomb, E. und Stubbs, L., The human homolog of a mouse-imprinted gene, Peg3, maps to a zinc finger gene-rich region of human chromosome 19q13.4, Genome Res. 7 (1997) 532-540.

Kim, J., Lu, X. und Stubbs, L., Zim1, a maternally expressed mouse Kruppel-type zinc-finger gene located in proximal chromosome 7, Hum. Mol. Genet. 8 (1999) 847-854.

Kim, J., Bergmann, A. und Stubbs, L., Exon sharing of a novel human zinc-finger gene, ZIM2, and paternally expressed gene 3 (PEG3), Genomics 64 (2000a) 114-8.

Kim, J., Noskov, V.N., Lu, X., Bergmann, A., Ren, X., Warth, T., Richardson, P., Kouprina, N. und Stubbs, L., Discovery of a novel, paternally expressed ubiquitin-specific processing protease gene through comparative analysis of an imprinted region of mouse chromosome 7 and human chromosome 19q13.4, Genome Res. 10 (2000b) 1138-1147.

Kim, J., Bergmann, A., Wehri, E., Lu, X. und Stubbs, L., Imprinting and evolution of two Kruppel-type zinc-finger genes, ZIM3 and ZNF264, located in the PEG3/USP29 imprinted domain, Genomics 77 (2001) 91-98.

Kim, J., Kollhoff, A., Bergmann, A. und Stubbs, L., Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3, Hum. Mol. Genet. 12 (2003) 233-245.

Kim, J., Bergmann, A., Lucas, S., Stone, R. und Stubbs, L., Lineage-specific imprinting and evolution of the zinc-finger gene ZIM2, Genomics 84 (2004b) 47-58.

Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D.J., Nicholls, R.D. und Cedar, H., Allele-specific replication timing of imprinted gene regions, Nature 364 (1993) 459-63.

Knoll, J.H., Cheng, S.D. und Lalande, M., Allele specificity of DNA replication timing in the Angelman/Prader-Willi syndrome imprinted chromosomal region, Nat Genet 6 (1994) 41-6.

Kobayashi, S., Wagatsuma, H., Ono, R., Ichikawa, H., Yamazaki, M., Tashiro, H., Aisaka, K., Miyoshi, N., Kohda, T., Ogura, A., Ohki, M., Kaneko-Ishino, T. und Ishino, F., Mouse Peg9/Dlk1 and human PEG9/DLK1 are paternally expressed imprinted genes closely located to the maternally expressed imprinted genes: mouse Meg3/Gtl2 and human MEG3, Genes Cells 5 (2000) 1029-37.

Kovacs, I., Pocsai, K., Czifra, G., Sarkadi, L., Szucs, G., Nemes, Z. und Rusznak, Z., TASK-3 immunoreactivity shows differential distribution in the human gastrointestinal tract, Virchows Arch 446 (2005) 402-10.

Kurella, M., Hsiao, L.L., Yoshida, T., Randall, J.D., Chow, G., Sarang, S.S., Jensen, R.V. und Gullans, S.R., DNA microarray analysis of complex biologic processes, J Am Soc Nephrol 12 (2001) 1072-8.

Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L.L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S.C., Ishino, F. und Surani, M.A., Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein, Nat. Genet. 12 (1996) 186-190.

La Salle, S. und Trasler, J.M., Dynamic expression of DNMT3a and DNMT3b isoforms during male germ cell development in the mouse, Dev Biol (2006)

Lau, M.M., Stewart, C.E., Liu, Z., Bhatt, H., Rotwein, P. und Stewart, C.L., Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality, Genes Dev 8 (1994) 2953-63.

Lee, L.G., Connell, C.R. und Bloch, W., Allelic discrimination by nick-translation PCR with fluorogenic probes, Nucleic Acids Res 21 (1993) 3761-6.

Lee, M.P., Ravenel, J.D., Hu, R.J., Lustig, L.R., Tomaselli, G., Berger, R.D., Brandenburg, S.A., Litzi, T.J., Bunton, T.E., Limb, C., Francis, H., Gorelikow, M., Gu, H., Washington, K., Argani, P., Goldenring, J.R., Coffey, R.J. und Feinberg, A.P., Targeted disruption of the Kvlqt1 gene causes deafness and gastric hyperplasia in mice, J Clin Invest 106 (2000) 1447-55.

Lefebvre, L., Viville, S., Barton, S.C., Ishino, F., Keverne, E.B. und Surani, M.A., Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest, Nat Genet 20 (1998) 163-9.

Li, E., Beard, C. und Jaenisch, R., Role for DNA methylation in genomic imprinting, Nature 366 (1993) 362-5.

Li, E., Chromatin modification and epigenetic reprogramming in mammalian development, Nat. Rev. Genet. 3 (2002) 662-673.

Li, L., Keverne, E.B., Aparicio, S.A., Ishino, F., Barton, S.C. und Surani, M.A., Regulation of maternal behavior and offspring growth by paternally expressed Peg3, Science 284 (1999) 330-3.

Li, L.L., Szeto, I.Y., Cattanach, B.M., Ishino, F. und Surani, M.A., Organization and parent-of-origin-specific methylation of imprinted Peg3 gene on mouse proximal chromosome 7, Genomics 63 (2000) 333-40.

Liang, P. und Pardee, A.B., Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction, Science 257 (1992) 967-71.

Lo, H.S., Wang, Z., Hu, Y., Yang, H.H., Gere, S., Buetow, K.H. und Lee, M.P., Allelic variation in gene expression is common in the human genome, Genome Res. 13 (2003) 1855-1862.

Lottspeich, F. und Zorbas, H. (1998). "Bioanalytik," Spektrum Akademischer Verlag.

Luedi, P.P., Hartemink, A.J. und Jirtle, R.L., Genome-wide prediction of imprinted murine genes, Genome Res 15 (2005) 875-84.

Lyle, R., Watanabe, D., te Vruchte, D., Lerchner, W., Smrzka, O.W., Wutz, A., Schageman, J., Hahner, L., Davies, C. und Barlow, D.P., The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1, Nat Genet 25 (2000) 19-21.

Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S. und Tilghman, S.M., Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes, Genes Dev 20 (2006) 1268-82.

Mathew, C.G. und Lewis, C.M., Genetics of inflammatory bowel disease: progress and prospects, Hum Mol Genet 13 Spec No 1 (2004) R161-8.

Maxam, A.M. und Gilbert, W., Sequencing end-labeled DNA with base-specific chemical cleavages, Methods Enzymol 65 (1980) 499-560.

McGrath, J. und Solter, D., Completion of mouse embryogenesis requires both the maternal and paternal genomes, Cell 37 (1984) 179-83.

Medhurst, A.D., Rennie, G., Chapman, C.G., Meadows, H., Duckworth, M.D., Kelsell, R.E., Gloger, II und Pangalos, M.N., Distribution analysis of human two pore domain potassium channels in tissues of the central nervous system and periphery, Brain Res Mol Brain Res 86 (2001) 101-14.

Mertineit, C., Yoder, J.A., Taketo, T., Laird, D.W., Trasler, J.M. und Bestor, T.H., Sex-specific exons control DNA methyltransferase in mammalian germ cells, Development 125 (1998) 889-97.

Miyoshi, N., Wagatsuma, H., Wakana, S., Shiroishi, T., Nomura, M., Aisaka, K., Kohda, T., Surani, M.A., Kaneko-Ishino, T. und Ishino, F., Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q, Genes Cells 5 (2000) 211-220.

Mizuno, Y., Sotomaru, Y., Katsuzawa, Y., Kono, T., Meguro, M., Oshimura, M., Kawai, J., Tomaru, Y., Kiyosawa, H., Nikaido, I., Amanuma, H., Hayashizaki, Y. und Okazaki, Y., Asb4, Ata3, and Dcn are novel imprinted genes identified by high-throughput screening using RIKEN cDNA microarray, Biochem. Biophys. Res. Commun. 290 (2002) 1499-1505.

Moon, Y.S., Smas, C.M., Lee, K., Villena, J.A., Kim, K.H., Yun, E.J. und Sul, H.S., Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity, Mol Cell Biol 22 (2002) 5585-92.

Moore, T. und Haig, D., Genomic imprinting in mammalian development: a parental tug-of-war, Trends Genet 7 (1991) 45-9.

Morgan, H.D., Santos, F., Green, K., Dean, W. und Reik, W., Epigenetic reprogramming in mammals, Hum Mol Genet 14 Spec No 1 (2005) R47-58.

Morison, I.M., Paton, C.J. und Cleverley, S.D., The imprinted gene and parent-of-origin effect database, Nucleic. Acids Res. 29 (2001) 275-276.

Morison, I.M., Ramsay, J.P. und Spencer, H.G., A census of mammalian imprinting, Trends Genet 21 (2005) 457-65.

Mu, D., Chen, L., Zhang, X., See, L.H., Koch, C.M., Yen, C., Tong, J.J., Spiegel, L., Nguyen, K.C., Servoss, A., Peng, Y., Pei, L., Marks, J.R., Lowe, S., Hoey, T., Jan, L.Y., McCombie, W.R., Wigler, M.H. und Powers, S., Genomic amplification and oncogenic properties of the KCNK9 potassium channel gene, Cancer Cell 3 (2003) 297-302.

Murphy, S.K., Wylie, A.A. und Jirtle, R.L., Imprinting of PEG3, the human homologue of a mouse gene involved in nurturing behavior, Genomics 71 (2001) 110-117.

Murrell, A., Heeson, S. und Reik, W., Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops, Nat Genet 36 (2004) 889-93.

Murrell, A., Rakyan, V.K. und Beck, S., From genome to epigenome, Hum Mol Genet 14 Spec No 1 (2005) R3-R10.

Nikaido, I., Saito, C., Mizuno, Y., Meguro, M., Bono, H., Kadomura, M., Kono, T., Morris, G.A., Lyons, P.A., Oshimura, M., Hayashizaki, Y. und Okazaki, Y., Discovery of imprinted transcripts in the mouse transcriptome using large-scale expression profiling, Genome Res. 13 (2003) 1402-1409.

Nikaido, I., Saito, C., Wakamoto, A., Tomaru, Y., Arakawa, T., Hayashizaki, Y. und Okazaki, Y., EICO (Expression-based Imprint Candidate Organizer): finding disease-related imprinted genes, Nucleic. Acids Res. 32 (2004) 548-551.

Oakey, R.J. und Beechey, C.V., Imprinted genes: identification by chromosome rearrangements and post-genomic strategies, Trends Genet. 18 (2002) 359-366.

Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., Hino, T., Suzuki-Migishima, R., Ogonuki, N., Miki, H., Kohda, T., Ogura, A., Yokoyama, M., Kaneko-Ishino, T. und Ishino, F., Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality, Nat Genet 38 (2006) 101-6.

Onyango, P., Miller, W., Lehoczky, J., Leung, C.T., Birren, B., Wheelan, S., Dewar, K. und Feinberg, A.P., Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain, Genome Res. 10 (2000) 1697-1710.

Paldi, A., Gyapay, G. und Jami, J., Imprinted chromosomal regions of the human genome display sex-specific meiotic recombination frequencies, Curr Biol 5 (1995) 1030-5.

Pastinen, T. und Hudson, T.J., Cis-acting regulatory variation in the human genome, Science 306 (2004) 647-650.

Patel, A.J. und Lazdunski, M., The 2P-domain K+ channels: role in apoptosis and tumorigenesis, Pflugers Arch 448 (2004) 261-73.

Paulsen, M., El-Maarri, O., Engemann, S., Strodicke, M., Franck, O., Davies, K., Reinhardt, R., Reik, W. und Walter, J., Sequence conservation and variability of imprinting in the Beckwith-Wiedemann syndrome gene cluster in human and mouse, Hum. Mol. Genet. 9 (2000) 1829-1841.

Paulsen, M. und Ferguson-Smith, A.C., DNA methylation in genomic imprinting, development, and disease, J Pathol 195 (2001) 97-110.

Pei, L., Wiser, O., Slavin, A., Mu, D., Powers, S., Jan, L.Y. und Hoey, T., Oncogenic potential of TASK3 (Kcnk9) depends on K+ channel function, Proc Natl Acad Sci U S A 100 (2003) 7803-7.

Perche, P.Y., Vourc'h, C., Konecny, L., Souchier, C., Robert-Nicoud, M., Dimitrov, S. und Khochbin, S., Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density, Curr Biol 10 (2000) 1531-4.

Peters, J., Wroe, S.F., Wells, C.A., Miller, H.J., Bodle, D., Beechey, C.V., Williamson, C.M. und Kelsey, G., A cluster of oppositely imprinted transcripts at the Gnas locus in the distal imprinting region of mouse chromosome 2, Proc Natl Acad Sci U S A 96 (1999) 3830-5.

Plagge, A., Gordon, E., Dean, W., Boiani, R., Cinti, S., Peters, J. und Kelsey, G., The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding, Nat Genet 36 (2004) 818-26.

Plagge, A., Isles, A.R., Gordon, E., Humby, T., Dean, W., Gritsch, S., Fischer-Colbrie, R., Wilkinson, L.S. und Kelsey, G., Imprinted Nesp55 influences behavioral reactivity to novel environments, Mol Cell Biol 25 (2005) 3019-26.

Plagge, A. und Kelsey, G., Imprinting the Gnas locus, Cytogenet Genome Res 113 (2006) 178-87.

Polak, P. und Domany, E., Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes, BMC Genomics 7 (2006) 133.

Qiu, J., Epigenetics: unfinished symphony, Nature 441 (2006) 143-5.

Reik, W. und Walter, J., Genomic imprinting: parental influence on the genome, Nat. Rev. Genet. 2 (2001) 21-32.

Reik, W., Murrell, A., Lewis, A., Mitsuya, K., Umlauf, D., Dean, W., Higgins, M. und Feil, R., Chromosome loops, insulators, and histone methylation: new insights into regulation of imprinting in clusters, Cold Spring Harb Symp Quant Biol 69 (2004) 29-37.

Robertson, K.D., DNA methylation and human disease, Nat Rev Genet 6 (2005) 597-610.

Ronaghi, M., Pyrosequencing sheds light on DNA sequencing, Genome Res 11 (2001) 3-11.

Rutsch, F., Ruf, N., Vaingankar, S., Toliat, M.R., Suk, A., Hohne, W., Schauer, G., Lehmann, M., Roscioli, T., Schnabel, D., Epplen, J.T., Knisely, A., Superti-Furga, A., McGill, J., Filippone, M., Sinaiko, A.R., Vallance, H., Hinrichs, B., Smith, W., Ferre, M., Terkeltaub, R. und Nurnberg, P., Mutations in *ENPP1* are associated with 'idiopathic' infantile arterial calcification, Nat. Genet. 34 (2003) 379-381.

Salas, M., John, R., Saxena, A., Barton, S., Frank, D., Fitzpatrick, G., Higgins, M.J. und Tycko, B., Placental growth retardation due to loss of imprinting of Phlda2, Mech Dev 121 (2004) 1199-210.

Sambrook, J., Fritsch, E.F. und Maniatis, T. (1989). "Molecular cloning: a laboratory manual," Cold Spring Harbor Press, Cold Spring Harbor.

Sanger, F., Nicklen, S. und Coulson, A.R., DNA sequencing with chain-terminating inhibitors, Proc Natl Acad Sci U S A 74 (1977) 5463-7.

Schmidt, J.V., Matteson, P.G., Jones, B.K., Guan, X.J. und Tilghman, S.M., The Dlk1 and Gtl2 genes are linked and reciprocally imprinted, Genes Dev. 14 (2000) 1997-2002.

Schuster-Gossler, K., Simon-Chazottes, D., Guenet, J.L., Zachgo, J. und Gossler, A., Gtl2lacZ, an insertional mutation on mouse chromosome 12 with parental origin-dependent phenotype, Mamm Genome 7 (1996) 20-4.

Schuster-Gossler, K., Bilinski, P., Sado, T., Ferguson-Smith, A. und Gossler, A., The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA, Dev. Dyn. 212 (1998) 214-228.

Sleutels, F., Zwart, R. und Barlow, D.P., The non-coding Air RNA is required for silencing autosomal imprinted genes, Nature 415 (2002) 810-813.

Sleutels, F., Tjon, G., Ludwig, T. und Barlow, D.P., Imprinted silencing of Slc22a2 and Slc22a3 does not need transcriptional overlap between Igf2r and Air, Embo J. 22 (2003) 3696-3704.

Smith, R.J., Dean, W., Konfortova, G. und Kelsey, G., Identification of novel imprinted genes in a genome-wide screen for maternal methylation, Genome Res. 13 (2003) 558-69.

Smith, R.J., Arnaud, P. und Kelsey, G., Identification and properties of imprinted genes and their control elements, Cytogenet Genome Res 105 (2004) 335-45.

Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G.B., Silengo, M.C. und Riccio, A., Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome, Nat Genet 36 (2004) 958-60.

Strachan, T. und Read, A.P. (2005). "Molekulare Humangenetik," Spektrum Akademischer Verliag, Heidelberg.

Strahl, B.D. und Allis, C.D., The language of covalent histone modifications, Nature 403 (2000) 41-5.

Strausberg, R.L., et al., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences, Proc Natl Acad Sci U S A 99 (2002) 16899-903.

Strichman-Almashanu, L.Z., Lee, R.S., Onyango, P.O., Perlman, E., Flam, F., Frieman, M.B. und Feinberg, A.P., A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes, Genome Res 12 (2002) 543-54.

Sun, A., Ge, J., Siffert, W. und Frey, U.H., Quantification of allele-specific G-protein beta3 subunit mRNA transcripts in different human cells and tissues by Pyrosequencing, Eur J Hum Genet 13 (2005) 361-9.

Sunahara, S., Nakamura, K., Nakao, K., Gondo, Y., Nagata, Y. und Katsuki, M., The oocyte-specific methylated region of the U2afbp-rs/U2af1-rs1 gene is dispensable for its imprinted methylation, Biochem Biophys Res Commun 268 (2000) 590-5.

Surani, M.A., Barton, S.C. und Norris, M.L., Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis, Nature 308 (1984) 548-50.

Surani, M.A., Barton, S.C. und Norris, M.L., Nuclear transplantation in the mouse: heritable differences between parental genomes after activation of the embryonic genome, Cell 45 (1986) 127-36.

Szabo, P.E., Hubner, K., Scholer, H. und Mann, J.R., Allele-specific expression of imprinted genes in mouse migratory primordial germ cells, Mech Dev 115 (2002) 157-60.

Tada, M., Tada, T., Lefebvre, L., Barton, S.C. und Surani, M.A., Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells, Embo J 16 (1997) 6510-20.

Takada, S., Tevendale, M., Baker, J., Georgiades, P., Campbell, E., Freeman, T., Johnson, M.H., Paulsen, M. und Ferguson-Smith, A.C., Delta-like and gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12, Curr. Biol. 10 (2000) 1135-1138.

Takeda, H., Caiment, F., Smit, M., Hiard, S., Tordoir, X., Cockett, N., Georges, M. und Charlier, C., The callipyge mutation enhances bidirectional long-range DLK1-GTL2 intergenic transcription in cis, Proc Natl Acad Sci U S A 103 (2006) 8119-24.

Talley, E.M., Solorzano, G., Lei, Q., Kim, D. und Bayliss, D.A., Cns distribution of members of the two-pore-domain (KCNK) potassium channel family, J Neurosci 21 (2001) 7491-505.

Tatusova, T.A. und Madden, T.L., BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences, FEMS Microbiol Lett 174 (1999) 247-50.

Thompson, J.D., Higgins, D.G. und Gibson, T.J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res 22 (1994) 4673-80.

Tsai, T.F., Armstrong, D. und Beaudet, A.L., Necdin-deficient mice do not show lethality or the obesity and infertility of Prader-Willi syndrome, Nat Genet 22 (1999) 15-6.

Uhlmann, K., Brinckmann, A., Toliat, M.R., Ritter, H. und Nurnberg, P., Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis, Electrophoresis 23 (2002) 4072-9.

Van Limpt, V.A., Chan, A.J., Van Sluis, P.G., Caron, H.N., Van Noesel, C.J. und Versteeg, R., High delta-like 1 expression in a subset of neuroblastoma cell lines corresponds to a differentiated chromaffin cell type, Int J Cancer 105 (2003) 61-9.

Velculescu, V.E., Zhang, L., Vogelstein, B. und Kinzler, K.W., Serial analysis of gene expression, Science 270 (1995) 484-7.

Velculescu, V.E., Vogelstein, B. und Kinzler, K.W., Analysing uncharted transcriptomes with SAGE, Trends Genet 16 (2000) 423-5.

Verona, R.I., Mann, M.R. und Bartolomei, M.S., Genomic imprinting: intricacies of epigenetic regulation in clusters, Annu Rev Cell Dev Biol 19 (2003) 237-59.

Vogelstein, B. und Gillespie, D., Preparative and analytical purification of DNA from agarose, Proc Natl Acad Sci U S A 76 (1979) 615-9.

Waddington, C., Der Epigenotypus, Endeavour 1 (1942) 18-20.

Walter, J. und Paulsen, M., Genomic Imprinting - Evolution eines neuen Konzepts der Genregulation bei Säugetieren, Medizinische Genetik 17 (2005) 270-274.

Wang, Z., Fan, H., Yang, H.H., Hu, Y., Buetow, K.H. und Lee, M.P., Comparative sequence analysis of imprinted genes between human and mouse to reveal imprinting signatures, Genomics 83 (2004) 395-401.

Wang, Z.Q., Fung, M.R., Barlow, D.P. und Wagner, E.F., Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene, Nature 372 (1994) 464-7.

Weksberg, R., Shuman, C. und Smith, A.C., Beckwith-Wiedemann syndrome, Am J Med Genet C Semin Med Genet 137 (2005) 12-23.

Wilkins, J.F. und Haig, D., What good is genomic imprinting: the function of parent-specific gene expression, Nat Rev Genet 4 (2003) 359-68.

Wilkins, J.F., Genomic imprinting and methylation: epigenetic canalization and conflict, Trends Genet 21 (2005) 356-65.

Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., Maconochie, M. und Peters, J., Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster, Nat Genet 38 (2006) 350-5.

Yan, Y., Frisen, J., Lee, M.H., Massague, J. und Barbacid, M., Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development, Genes Dev 11 (1997) 973-83.

Yang, T., Adamson, T.E., Resnick, J.L., Leff, S., Wevrick, R., Francke, U., Jenkins, N.A., Copeland, N.G. und Brannan, C.I., A mouse model for Prader-Willi syndrome imprinting-centre mutations, Nat Genet 19 (1998) 25-31.

Yatsuki, H., Watanabe, H., Hattori, M., Joh, K., Soejima, H., Komoda, H., Xin, Z., Zhu, X., Higashimoto, K., Nishimura, M., Kuratomi, S., Sasaki, H., Sakaki, Y. und Mukai, T., Sequence-based structural features between Kvlqt1 and Tapa1 on mouse chromosome 7F4/F5 corresponding to the Beckwith-Wiedemann syndrome region on human 11p15.5: long-stretches of unusually well conserved intronic sequences of kvlqt1 between mouse and human, DNA Res. 7 (2000) 195-206.

Yu, S., Gavrilova, O., Chen, H., Lee, R., Liu, J., Pacak, K., Parlow, A.F., Quon, M.J., Reitman, M.L. und Weinstein, L.S., Paternal versus maternal transmission of a stimulatory G-protein alpha subunit knockout produces opposite effects on energy metabolism, J Clin Invest 105 (2000) 615-23.

Zwart, R., Verhaagh, S., Buitelaar, M., Popp-Snijders, C. und Barlow, D.P., Impaired activity of the extraneuronal monoamine transporter system known as uptake-2 in Orct3/Slc22a3-deficient mice, Mol Cell Biol 21 (2001) 4188-96.

7. Anhang

7.1. Danksagung

Die vorliegende Arbeit wurde erst durch die Ideen und die Unterstützung vieler Persönlichkeiten möglich, denen ich an dieser Stelle von ganzem Herzen meinen Dank ausspreche. Dazu zählen:

Prof. Dr. Peter Nürnberg, der als mein Doktorvater das Thema vorgeschlagen hat, mich dazu ermuntert hat, das Doktorandenstipendium des Max-Delbrück-Centrums zu beantragen, mir die Gelegenheit gegeben hat, in weiteren wissenschaftlichen Projekten zusätzliche Erfahrung zu sammeln, mich während der Doktorarbeit immer unterstützt hat und mir als Mentor durch die Qualität seiner wissenschaftlichen Arbeit und seine Menschlichkeit stets ein Vorbild bleiben wird.

Dr. Ulrich Zechner, der als fachlicher Betreuer die thematische Idee eingebracht hat, durch seine hohe wissenschaftliche und soziale Kompetenz meine Begeisterung für genomische Prägung und Epigenetik geweckt, mich bei den Besuchen in Mainz wie selbstverständlich in seinem Haus aufgenommen und mich zuverlässig während der gesamten Zeit unterstützt hat.

Prof. Dr. Ulf Stahl, für seine konstruktiven Ratschläge und die Übernahme der Fachvertretung vor der Fakultät III für Prozesswissenschaften der Technischen Universität Berlin.

Dr. Sylvia Bähring und Prof. Dr. Dr. Friedrich C. Luft für ihre Bereitschaft mich im April 2005, bedingt durch den Umzug meiner Arbeitsgruppe von Berlin nach Köln, in ihrer Gruppe am MDC aufgenommen und mir die nötigen Freiräume gelassen zu haben sowie die stete Unterstützung und Diskussionsbereitschaft.

Philippe Luedi und sein Doktorvater Prof. Dr. Randy Jirtle vom Department of Radiation Oncology des Duke University Medical Center in Durham (USA) für die unkomplizierte und erfolgreiche Kooperation.

Die Verantwortlichen vom Max-Delbrück-Centrum für die Bewilligung des dreijährigen Doktorandenstipendiums.

Alle ehemaligen Kollegen vom Gene Mapping Center des Max-Delbrück-Centrums, den Kollegen der Arbeitsgruppe Bähring/Luft sowie Prof. Dr. Thomas Haaf und seinen Mitarbeitern vom Institut für Humangenetik der Universität Mainz für jegliche Unterstützung und die freundliche Zusammenarbeit.

Anne Hempelmann, Dr. Birgit Uhlenberg und Liane Kulick für die konstruktiven Diskussionen und das Korrekturlesen der Dissertation.

Meinen Eltern Gudrun und Wolfgang Ruf sowie meiner Schwester Daniela Ruf für ihr Vertrauen und ihre Unterstützung.

7.2. Publikationen

1) Rutsch F*, **Ruf N***, Vaingankar S*, Toliat MR, Suk A, Höhne W, Schauer G, Lehmann M, Roscioli T, Schnabel D, Epplen JT, Knisely A, Superti-Furga A, McGill J, Filippone M, Sinaiko AR, Vallance H, Hinrichs B, Smith W, Ferre M, Terkeltaub R, Nürnberg P. Mutations in ENPP1 are associated with 'idiopathic' infantile arterial calcification. Nat Genet. 2003 Aug;34(4):379-81.

2) Uhlenberg B, Schülke M, Rüschendorf F, **Ruf N**, Kaindl AM, Henneke M, Thiele H, Stoltenburg-Didinger G, Aksu F, Topaloglu H, Nürnberg P, Hübner C, Weschke B, Gartner J. Mutations in the gene encoding gap junction protein alpha 12 (connexin 46.6) cause Pelizaeus-Merzbacher-like disease. Am J Hum Genet. 2004 Aug;75(2):251-60.

3) Tinschert S*, **Ruf N***, Bernascone I*, Sacherer K, Lamorte G, Neumayer HH, Nürnberg P, Luft FC, Rampoldi L. Functional consequences of a novel uromodulin mutation in a family with familial juvenile hyperuricaemic nephropathy. Nephrol Dial Transplant. 2004 Dec;19(12):3150-4.

4) **Ruf N**, Uhlenberg B, Terkeltaub R, Nürnberg P, Rutsch F. The mutational spectrum of ENPP1 as arising after the analysis of 23 unrelated patients with generalized arterial calcification of infancy (GACI). Hum Mutat. 2005 Jan;25(1):98.

5) Malkin I, Dahm S, Suk A, Kobyliansky E, Toliat M, **Ruf N**, Livshits G, Nürnberg P. Association of ANKH gene polymorphisms with radiographic hand bone size and geometry in a Chuvasha population. Bone. 2005 Feb;36(2):365-73.

6) Cheng KS, Chen MR, **Ruf N**, Lin SP, Rutsch F. Generalized arterial calcification of infancy: different clinical courses in two affected siblings. Am J Med Genet A. 2005 Jul 15;136(2):210-3.

7) Suk EK, Malkin I, Dahm S, Kalichman L, **Ruf N**, Kobyliansky E, Toliat M, Rutsch F, Nürnberg P, Livshits G. Association of ENPP1 gene polymorphisms with hand osteoarthritis in a Chuvasha population. Arthritis Res Ther. 2005;7(5):R1082-90.

8) Zheng W, Rosenstiel P, Huse K, Sina C, Valentonyte R, Mah N, Zeitlmann L, Grosse J, **Ruf N**, Nürnberg P, Costello CM, Onnie C, Mathew C, Platzer M, Schreiber S, Hampe J. Evaluation of AGR2 and AGR3 as candidate genes for inflammatory bowel disease. Genes Immun. 2006 Jan;7(1):11-8.

9) Ciana G, Trappan A, Bembi B, Benettoni A, Maso G, Zennaro F, **Ruf N**, Schnabel D, Rutsch F. Generalized arterial calcification of infancy: two siblings with prolonged survival. Eur J Pediatr. 2006 Apr;165(4):258-263.

10) **Ruf N**, Dünzinger U, Brinckmann A, Haaf T, Nürnberg P, Zechner U. Expression profiling of uniparental mouse embryos is inefficient in identifying novel imprinted genes. Genomics. 2006 Apr;87(4):509-19

11) **Ruf N**, Martelli M, Weschke B, Uhlenberg B. Oligodendroglial transcription factor (OLIG1 and OLIG2) mutations are not associated with Pelizaeus-Merzbacher-like leukodystrophy. Am J Med Genet B (*in press*)

* gleichberechtigte Erstautoren

7.3. Oligonukleotide

Tabelle 13. Sequenzierungsprimer zur SNP-Identifizierung.

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
36	C7-39658ex2-F	TTTCCATCTTCTACCTCCTCTACC	58	500
	C7-39658ex2-R	TCATGGAAGATAGCAGACATAAACTC		
37	C7-39658ex4-F	AAGTGCAGTGGTGCAAGATG	60	289
	C7-39658ex4-R	TCCACAGGAAGAAGGATTGC		
38	C7-39658ex5-F	GGGACATCCTTGCACATTTC	60	778
	C7-39658ex5-R	GGACTTTGCGCACTCCTTAC		
39	C7-38886-A-F	CATTTCCCAGAAAGGACTGC	60	328
	C7-38886-A-R	CACTGGATTGGGGTACCAAC		
40	C7-38886-B-F	CTGGCATTCAGGGAGCTTAG	60	669
	C7-38886-B-R	CTCAGGTCCTTCTCCAGCAC		
41	C7-028202-Ex1-F	GCCTTGAAGGGTGTCTGC	60	446
	C7-028202-Ex1-R	GCACCTCTGCCTGACTTCTC		
42	C7-028202-Ex5-F	GCCAGCATGTGTATCTGTGC	60	526
	C7-028202-Ex5-R	TGTGACTCACCCCACACTTC		
43	C7-028202-Ex7-F	TTATTCCGAGTGGCTCCTTC	60	746
	C7-028202-Ex7-R	CTGAAGGCTTTCCCACAGTC		
44	C7-046096ex4-F	GTTGGGCTTTGTTTCTGAGC	60	429
	C7-046096ex4-R	CAAATAAACAACCCCCTCCAC		
45	C7-046096ex5-F	GAATGCCAAGAAAGGGTGAG	58	740
	C7-046096ex5-R	TTTGCAAAATAATGGCAACG		
46	C7-046096ex2-F	TCTGGATTTTGGAGTAGTTTTGG	60	484
	C7-046096ex2-R	ATGAAGAAATGGCTCCGTTC		
47	C7-36368ex7/8-F	GGTGCTGTCTGTAGGGAAGG	60	505
	C7-36368ex7/8-R	GCCAGATGGAGGACTAGCAG		
48	C7-36368ex9-F	TGCTGGCTCTCTGGTTCC	60	689
	C7-36368ex9-R	GGGGTCCCACTCCCACTC		
49	C7-36368ex12-F	GGAACAGGTGGTGAGGTAGC	60	572
	C7-36368ex12-R	GCTTCTGCCTACCCCATAAAG		
50	C7-17603ex1-F	CCTGCACTCAACAGTGATGTC	60	477
	C7-17603ex1-R	GGGTGGAACACTTCATTTCG		
51	C7-17603ex3-F	AAAGGAGACTTGGCCCTCTG	60	586
	C7-17603ex3-R	GGCAGGTTGAGAGGAGACAG		
52	C7-17603ex6-F	CAACTTCAGGGTGGCTCTTC	60	586
	C7-17603ex6-R	TAGAGGTCCCCATTCTGTGG		
53	C7-17603ex12-F	AGGCTGCAAGGTTATGGATG	60	505
	C7-17603ex12-R	CCCAGCTCCACCTACTTCAC		
54	C7-36717ex1-F	AGATGGATGGCTGCTTATGG	60	426
	C7-36717ex1-R	AAGTTGTGGCCTCTGTGGTC		
55	C7-36717ex2-F	CCGCACTGAGAAGAGAGGTC	60	495
	C7-36717ex2-R	CCCTCTTCCACTCTGCAGTC		
56	C7-36717ex9-F	CCATTTCCCATTGAACTTGC	60	403
	C7-36717ex9-R	CAACAACACTGTGGGTGAGG		
57	C7-36717ex12-F	AGCAGGTGCCTACACCTCTG	60	599
	C7-36717ex12-R	TTCCAACATGATCGCACAAC		
58	C7-37730ex5-F	AGTGGTTCCATCTGGTTTCC	60	600
	C7-37730ex5-R	AGACCTGCGTAGGATGATGG		
59	C7-37730ex17-F	AGACAGCGGTACCTGACTCC	60	695
	C7-37730ex17-R	GGGGAAGAGCGTGAGGATAG		
60	C7-37730ex29-F	AGGGCCATTTATTGCTTTCC	60	533
	C7-37730ex29-R	AAGTGAATGGCACTGTTTGC		
61	C7-89294ex2-F	CATGTTGTGGCCATTTTGAG	60	378
	C7-89294ex2-R	CTCACCCCTCAAAGCAAGAC		
62	C7-89294ex5-F	GGAGCTGAGGCAGTATCGTC	60	594
	C7-89294ex5-R	CTGGCTCATGGCTTTCTCTC		
63	C7-89294ex6-F	TCCTTTGCTGACACTTCTGAAC	60	750
	C7-89294ex6-R	TGATCACACGGTACCCACAC		
64	C7-34864ex3-F	CTTCCAAGTCCACCCATGTC	60	501
	C7-34864ex3-R	CAGGCAAAACGCACATACAC		
65	C7-34864ex5-F	CTTGCTCAGGGCTCATTTG	60	759
	C7-34864ex5-R	GCTACCTCTCAGCCCACATC		
66	C7-34864ex2-F	TTGGTGAGTCATTCCTTTGC	60	403
	C7-34864ex2-R	GGCTGCATTACACTCTGTGG		
67	C7-17829ex1-F	TTGCTTTGTGGGTTGTGC	60	700
	C7-17829ex1-R	AAAACAATACGGACCCAACG		

Fortsetzung Tabelle 13

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
68	C7-17829ex3-F	CATGCACACAGCCTCTTGTC	60	425
	C7-17829ex3-R	AACAGGGTCCATGTGAGACTTAG		
69	C7-86590e5A-F	CCACACAAGCATCTGCATTC	59	466
	C7-86590e5A-R	TTTGGATTTGTGGAACTGCTC		
70	C7-86590e5B-F	CTGTGAGCATTTGTTGATCG	59	485
	C7-86590e5B-R	TGAGGCTAGTCTGGGCTACAC		
71	C7-86590e6-F	GCAGATGCTTAATGGGTCAC	59	496
	C7-86590e6-R	GGGCAGGAAGAATGTCACAC		
72	C7-038418ex1-F	GGGTGGGAGGTTATAATACAACTG	59	666
	C7-038418ex1-R	CTTCCTGCTTCCCATTGTTC		
80	C7-36368ex1-F	ACTACGCGTCCCACAATGC	59	627
	C7-36368ex1-R	CAACCTGACCCCAAAACTTG		
81	C7-36368ex2-F	TCCTGTGCTGGGTGACTAGG	61	226
	C7-36368ex2-R	CCAGAGCTCCACAGGACATC		
82	C7-36368ex3-F	CACTGAGGGGCTCTTTGTTC	59	237
	C7-36368ex3-R	GTACAGTCAGCAGCCACAGG		
83	C7-36368ex4-F	CAGAGCAGAGCCCCTGAC	59	241
	C7-36368ex4-F	CAGCCTGGAGGGTAAGACAG		
84	C7-36368ex5-F	CACCACACTAATGGCTCTGC	59	665
	C7-36368ex5-R	TGGCAAAAGGGGAACTAGAC		
87	C11-29538ex7-F	TTCAAGAACAGTTTGCCAAGTC	58	392
	C11-29538ex7-R	CCTTAACCCTTTCTCCCAGAAC		
88	C11-29538ex9-F	AATGGAAGAGAGGGAAGAAGG	58	493
	C11-29538ex9-R	GCAAACAAGACTTTGTCTCTATAAGC		
89	C11-29538ex14-F	GCCCCTCCTGTAAATTGTTAAG	58	398
	C11-29538ex14-R	CCATTTTGCCAGCCCAAG		
90	C11-29538ex16-F	TTCTTAGGAGTCTGTGTGTGACG	58	398
	C11-29538ex16-R	ATAAAGCCAATCCCCCATGT		
91	C12-14491ex1-F	GCTCAGTGGGACACCTAACC	58	279
	C12-14491ex1-R	AGAGAGTGCAGCTTGTTTTGC		
92	C12-14491ex2-F	CAGCTCCATCCCCACCTC	58	279
	C12-14491ex2-R	CCTAGGTTCCTAGTGTGCTGT		
93	C12-14491ex3-F	CTTGAAGGCCTCAGGATCAC	60	398
	C12-14491ex3-R	AGGCCAGGGATCAGCTCTAC		
94	C12-14491ex4-F	AGACTCCATTTGGTGTGTGC	58	300
	C12-14491ex4-R	AGTCTGTCACCCGTCTCTGC		
95	C12-14491ex5-F	GGAGGCAGCCAATAAGTGTC	58	744
	C12-14491ex5-R	CTAATGCCTTCCCCACTGAC		
96	C14-14850ex4-F	ACACCCTGTTGGTGAGGTTG	58	334
	C14-14850ex4-R	AAGGCTTGTTTAGATGCAGTTC		
97	C14-14850ex7-F	GTGCCAGGTTTCAAGGACAC	60	399
	C14-14850ex7-R	TGTGTGTGTGTTTGTGGTGTG		
98	C14-14850ex8-F	GGAAGCTCAATAAAAACCAAGC	60	710
	C14-14850ex8-R	AGGAGGGTGCACGAGCTAAC		
99	C17-38428ex1-F	AACAAGGCCCATTTCATCTG	58	283
	C17-38428ex1-R	TCTCCATGGACTCAGCTAATAAAG		
100	C17-38428ex2-F	TCCTTTTGTGGAGAAGATGG	58	300
	C17-38428ex2-R	CATGCATGACTCTGTGATGG		
101	C17-38428ex3-F	AGGGCACTGTCCATTGTACC	60	723
	C17-38428ex3-R	GGCCGCACAGTGTTTATTATG		
102	C17-45985ex1-F	TCACCAGCGTCCTTTATTCAC	60	497
	C17-45985ex1-R	CGTCCTACGTCCAGGGTTC		
103	C17-45985ex2-F	TGAGGTGAAGCACCGAAAAC	60	497
	C17-45985ex2-R	TCAAGTTTAACCAGAAACTGGAG		
104	C17-45985ex5-F	GATCTTTTCCAGACAGTCCTCAG	60	600
	C17-45985ex5-R	TCTTCATACTTATGTCTTTGGTAACTG		
105	C18-40367ex1-F	GGGTCCCTGAGAGTTTAGCC	60	397
	C18-40367ex1-R	GCTGCACAGAAGACCAACTTC		
106	C18-40367ex2-F	GGAAGCATGAAAGGTCTGTTG	60	498
	C18-40367ex2-R	GGGTACAGGCTGGAGTCTAGG		
113	C17-45985ex3-F	GCCCTATTCCATAGCAGCTC	59	581
	C17-45985ex3-R	ATCTTGAGGCAGGGTGAATG		
114	C17-45985ex4-F	TAGTTAGTGCCGCCTCACAG	59	577
	C17-45985ex4-R	ACCACCATCAAAGCTTCCTG		
116	C7-046096ex11-F	TGTGACATTGGCCTTAATGTTG	60	298
	C7-046096ex11-R	GCTCTGCTTGCGTGTATGG		
117	C7-046096ex12-F	GCATTTAACACTGAGCATATTTCC	58	281
	C7-046096ex12-R	CAGCCTTTTTGTACTGGGATAC		
118	C7-046096ex5B-F	TGAATGATGATGATGCCATTG	60	685
	C7-046096ex5B-R	TTGGTAAGAAGCAACAAAGGTG		

Fortsetzung Tabelle 13

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
119	C7-046096ex6-F	TTTCCATATTGTTTTGTTATTGCTG	59	372
	C7-046096ex6-R	GCTCCTGAAGCAGGGATATG		
120	C7-046096ex9-F	AAGGTTTTCTTGGGGTGGAG	60	543
	C7-046096ex9-R	CAGGACACAAAAGCACACATC		
148	C7-43056a-F	GCACCCAGCTAGCTATACCC	58	686
	C7-43056a-R	AATGGTGCTTGTTGCTCTTG		
149	C7-43056b-F	TTGATTTTGAATTGTTCAAGGTG	59	842
	C7-43056b-R	GGCAATCCATCAGACCCATA		
150	C7-43056c-F	ATTGTCCTTGTCTGGCTTGG	58	890
	C7-43056c-R	TGCTCCCTGGTGTCATTTAAC		
151	C7-79260a-F	TTTAAGGGTCCTGGGAGACC	60	840
	C7-79260a-R	TTCCCTTCTCTCCCACTTTG		
152	C7-79260b-F	GGTCACACCCACTCACACTG	58	795
	C7-79260b-R	ATGTGTCCCTCCTGCTTCAG		
153	C7-79260c-F	AGTTTATCGGAGCCCAGGAC	58	843
	C7-79260c-R	GCAGGTTTCCTCTGCTTCAC		
154	C7-34189a-F	AAATTCCAATGAAAGGGTGCT	58	815
	C7-34189a-R	GCTGTGTGCATCACTGGAAT		
155	C7-34189b-F	TTCTGCTCTTTCTGAAGCATTG	59	838
	C7-34189b-R	TTTCAAGACTATGAATACCACAACC		
156	C7-38900a-F	CCACACTCCCTTCATATTCCA	58	801
	C7-38900a-R	GAGGGTCAGTTTACTCTGATTGG		
157	C7-38900b-F	GCTAAGTTATTTCCTATTGCCAAG	58	892
	C7-38900b-R	GAGATGTGGAATATTTATGGCTTCA		
158	C7-34097a-F	GCTAAGTTATTTCCTATTGCCAAG	58	938
	C7-34097a-R	CAGTTGATATGTGTGTGGCTTT		
159	C7-34097b-F	TTTCAGAGAGTTGGTTCAGTGC	59	992
	C7-34097b-R	CGCTAGAAACAGGGCCTTAC		
160	C7-14392a-F	CTGCTCACTGCATCTTCTGC	58	891
	C7-14392a-R	GCTCTCTGGTCCCATCTCTG		
161	C7-14392b-F	TGGTCTCCTGTATGGCACAG	59	848
	C7-14392b-R	GGGAAAGATTTGGATGTGCT		
162	C7-39482a-F	TGCATCCAACCAATTCTTGA	58	975
	C7-39482a-R	AGAGCATCCAGTTTCCATGAG		
163	C7-39482b-F	TTGAGAATGGATCGCAGACA	58	644
	C7-39482b-R	ATGGATATCTTCCCGGCTCT		
164	9-34922a-F	TGTGGGTCTCTGTGTGTGTCT	60	980
	9-34922a-R	GCAAACCTGGTGTCTTAGGG		
165	9-34922b-F	GTCTGGCCTGGGTACAATCT	60	700
	9-34922b-R	GCCACAAACACCCATACACA		
166	10-37266a-F	AACCTTCAAAGATTCGTTCAT	55	944
	10-37266a-R	CCTTCTTATAGGGTTCAGGACA		
167	10-37266b-F	GGAGAAAGTGCGTGACATGG	58	976
	10-37266b-R	AAGACAGTTGGGAAACGACA		
168	10-38360a-F	TCAGATGGTGCTGGAAATGA	60	839
	10-38360a-R	CCCTGATTGGGTTTGGTTT		
169	10-38360b-F	TTGCCAGTTAGTGGAACCTTT	60	689
	10-38360b-R	GTGCACTGCCATCTAACACC		
170	12-39463-F	CCCTCCTGTTTGAGATAGGG	60	494
	12-39463-R	ATGCATACCTGGTGCCTTG		
171	12-18121a-F	AGTTTAGGGATGACGGCACA	58	839
	12-18121a-R	TCTGTCCAGACAGCCTCTCTTA		
172	12-18121b-F	GATGAGAGGCTGTGTCAGCA	60	834
	12-18121b-R	GTTGGCCACGGTGATTAACT		
173	12-18142a-F	GATAGGAGGGAAGGGCTCAG	60	688
	12-18142a-R	TGGTTTGGATTGGTTTGGTT		
174	12-18142b-F	GGTGACCAGACAGAGGGAGA	60	801
	12-18142b-R	TTCCCAGAGAGGAGACAAGG		
175	12-40459a-F	TGCTATCAGTGTACGGGTGA	58	843
	12-40459a-R	AGTAACAGCGAGTCGGAACC		
176	12-40459b-F	ACCAGTCTCCACATGCTGCT	58	844
	12-40459b-R	TTGGCGAATGTTTGTCTTCA		
177	14-38100a-F	CTGCCTCCTGAGTGCTGA	60	881
	14-38100a-R	GCACTCATTCCATTCCCAAG		
178	14-38100b-F	TTCCAGCTTTGTACCATGAGG	60	794
	14-38100b-R	GCCTGATGCCTGTTCCATAA		
179	14-37626a-F	TTAGCAACCCATTTCCCAGT	60	850
	14-37626a-R	GTGCTCAGTGGATAGGGACA		
180	14-37626b-F	TGGGAATGGCTTTGAACTAC	60	742
	14-37626b-R	TGGCACCATCAGTTCCTGTA		

Fortsetzung Tabelle 13

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
181	17-39514a-F	CATTTGGGATGGTGCTTGTA	60	749
	17-39514a-R	CTTCATGGACAGAGGCCAAT		
182	17-39514b-F	AATCCCAGGTGTGGTAGGTA	60	700
	17-39514b-R	ACCTGGGAACAGGAGCTG		
183	17-44118a-F	GAGCCTGGTGTTGAACTCGT	60	827
	17-44118a-R	GCACAACACATGACTCAGAACA		
184	17-44118b-F	GTGAGCGTCAGTGCAAATGT	60	697
	17-44118b-R	CCCTTTAACTTCCAGAAGAGGA		
185	17-40032a-F	TGCTGGCAAGGTACAGTCTTT	60	1000
	17-40032a-R	CTGCCACCTTAACACAGAAGC		
186	17-40032b-F	TGTCATTTCCTGCTTCTCCA	60	825
	17-40032b-R	CCTGTGTCTGTTCCCATGAA		
187	C7-37653a-F	TGTCTATTTGTTGTGCTGTGAGG	59	822
	C7-37653a-R	CTCATGATCTCCATCCCTGA		
188	C7-37653b-F	TGCAGATACAATCCAGGCTTT	59	745
	C7-37653b-R	TCCTTTCTGGTTCCCTGTTT		
189	C7-38900c-F	TCTGAATAGAGATTGACAAATTCCTG	59	989
	C7-38900c-R	TTGGCAATAGGAAATAACTTAGCTT		
190	C7-38900d-F	GACATCATAGTGGATGCCAAA	59	982
	C7-38900d-R	GCACTGAACCAACTCTCTGAAA		
191	C7-34097c-F	AGGGCATTTGACTAGAGAACG	58	978
	C7-34097c-R	TCTGTCAGAAGTGAGAAATTGTG		
192	C7-43056d-F	GTGAACTGACTGGCGGTTT	59	850
	C7-43056d-R	GGAGGAGAATTTATAGCCTTGG		
226	sC7-38418b-F	TTCACAATCTCTGCCATTGTTT	59	833
	sC7-38418b-R	TTGCTTCAATTGTGCATAGGA		
227	sC7-38418c-F	GGCCAGCCTCTCTCTTTCTT	59	963
	sC7-38418c-R	TGCTCATCACTGTCCAGCTT		
228	sC7-38418d-F	AGAGGCTCACAAACACTGAAGA	59	972
	sC7-38418d-R	CAGCAGGAGGCTAGACAAGG		
242	sC6-34724/3a-F	AGTGGTCCTCACAGCGAGAT	60	933
	sC6-34724/3a-R	GGCTGGCTGCATCTGTAAAT		
243	sC6-34724/3b-F	GTCCCAAGCCTGTGTAAGGA	60	948
	sC6-34724/3b-R	TGACATCAGGACCTATCCCTTT		
246	sC2-38367/4-F	ACCCTCACAGACTCACATACG	57	699
	sC2-38367/4-R	ACTGGAAACTTCACAGGCTTT		
247	sC2-38367/5-F	ATGAGGCTGGAGGATCAATG	60	951
	sC2-38367/5-R	ACGCCCAGCGTGTCTTACT		
248	sC2-38367/2-F	TGCAGATGACTCAGCCTCTC	59	807
	sC2-38367/2-R	CAGGGTTCTCAGCGGTTAAG		
250	s6-34724/1-F	AACGCTGAGCTTCATGGTG	60	700
	s6-34724/1-R	GGCTCTGCAAAGGGAGAGT		
251	s6-34724/3c-F	CTGCAATGTCCTTCTGCTCA	60	395
	s6-34724/3c-R	GTGGGAAATGCAGAGGATGT		
259	HsUsp29A-F	CAGCTGCTTGCTTTGAAAGAT	60	958
	HsUsp29A-R	CAGTGAGTCAGCCACTGGAA		
260	HsUsp29B-F	GAGATAGGGCAATCGGTGAA	60	889
	HsUsp29B-R	GAAGTGAGGCCGAGTGAGTC		
261	hC19-43056-1F	CATGCAGCTGGGTACTACTGA	59	953
	hC19-43056-1R	CCTCCCAAATAAGCAGAAGG		
262	hC19-43056-2F	CCTTCTGCTTATTTGGGAGGT	58	929
	hC19-43056-2R	TGTTCATCCATTTGAGAAAGG		
263	Zim3ex5A-F	CGCTGACTACGCAGAAAGGT	59	478
	Zim3ex5A-R	GATTAATGCAGGATGATTTCCA		
264	Zim3ex5B-F	CACACGGGAGAGAAACCCTA	60	598
	Zim3ex5B-R	TGAGTGAGTATGCCGAATGG		
265	ZNF264ex2-F	TCTTGAGCCCAGCACAGAG	60	393
	ZNF264ex2-R	TGGGAAAGAGGCAAGGTAGA		
266	ZNF264ex4-F	GAGCCCTGAATGTGATGGTT	60	351
	ZNF264ex4-R	TTGAAGGCCTTTCCACACTC		
267	NB-43056-F	AGGCGTCCTTTCTCTTAGGG	60	319
	NB-43056-R	ATCTGCTCCCTGGTGTCATT		
268	hC19-43056-3F	TTGCGACATGAAGAAACGTC	60	436
	hC19-43056-3R	AGAGTCATGGCCTTGTTTGC		
269	hC19-43056-4F	CATACCCTGGTCTCCCAAAG	60	448
	hC19-43056-4R	CCCAGAAGTCCTAGCCAATG		
270	hC19-43056-5F	AAATTGAAGGCCAAAGCAGA	60	486
	hC19-43056-5R	GGGTCAGAGGAGATGAGCTG		
279	hPeg3A-F	AATACCCACACTGGCTGAGG	60	838
	hPeg3A-R	TGCATGAGAACCACTTCAACA		
280	hPeg3B-F	TCTGAATGCCTGCATCATGT	60	295
	hPeg3B-R	GGAGTTGAGAGCTACTGCCCTA		

Fortsetzung Tabelle 13

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
281	hZim2-F	CGGAAAGAAGCTATGGCAGT	59	831
	hZim2-R	AAGGCCTTCTCACACTCACAA		
282	mPeg3A-F	ACCCATTGATCAGGAGATGC	60	793
	mPeg3A-R	GCCCATACTCATAGGGCTCA		
283	mPeg3B-F	TGGTTGAGTAACCAGGCTGA	59	821
	mPeg3B-R	GATGGGTTGATTTGGGTCAC		
284	mUsp29A-F	GGCCTTACATACATCCTGGAA	59	849
	mUsp29A-R	TGAATCACAAACACACATGGA		
285	mUsp29B-F	AATCGCTCATGTCATCCAGA	59	846
	mUsp29B-R	CTCTCACCTGCAGGACAGTCT		
286	mUsp29C-F	CCACCATGCCTAACACAATG	59	840
	mUsp29C-R	CAGAAGAGGCCAGTTGATCC		
L1	Foxg1ex2A-F	GAGGTGCAATGTGGGAGAAT	59	838
	Foxg1ex2A-R	GCCTTTGAATTTCCTATAAGTTGA		
L2	Kcnk9ex2-F	CAGTCTGTTCGAAATGTCTGATG	59	694
	Kcnk9ex2-R	CACAGGGTTCTGCCGTATTT		
L3	Prdm16ex9A-F	CATCTGTGTTTGCCTCCTGT	59	841
	Prdm16ex9A-R	GTACCTCGGCCACACTGTTC		
L4	NppcEx3-F	ACAACTACCCTGTGGGATGG	59	649
	NppcEx3-R	TTGTCTTCTGCCTCTGCTCA		
L5	Wnt7aEx4A-F	GATGACACCCATTCTCTTCCA	60	843
	Wnt7aEx4A-R	GTGGCCTTTCTGTGTCCAGT		
L6	Dok1ex5A-F	GGTCAGGCCTGTGACTCATT	59	762
	Dok1ex5A-R	ATCTGAGCTGAAGCCTTTGC		
L7	Nkx6-2ex4-F	GAGTGCTGGTCGGGTGAG	60	969
	Nkx6-2ex4-R	TTGCAGAACTTTGGGTCTCC		
L8	Hes5ex3A-F	CACCTGACGCCTCTCCTC	59	823
	Hes5ex3A-R	CTTGCTGAAATGCAAACACAA		
L9	Cdk6ex7A-F	CCGTGCTATGTTTCACACCA	60	948
	Cdk6ex7A-R	TATAGGGCCTGGGACACATT		
L10	Rarres1ex6-F	TGCAAAGTTGTGTATCGAATTG	57	691
	Rarres1ex6-R	TCATGTGAGACAGTAGAGACAGTG		
L11	Camk2bEx24A-F	ATGTTGGGCTGAGAAAGGTG	60	842
	Camk2bEx24A-R	CTGCCCAGTCCTCTCTTCAT		
L12	Ntng2ex2-F	TTGCTGATGGCTAGCAAATG	60	822
	Ntng2ex2-R	CAGCAAGGTACCAGGTGACA		
L13	Disc1ex3A-F	GGAGTGTCCCACCATGAGTT	60	849
	Disc1ex3A-R	GGGCCTGTCTGAGCTAGATG		
L14	Irx4ex6A-F	ACTTGAGGCCCTCTGGAGAT	60	983
	Irx4ex6A-R	CACTGTGCAGGTGGGTGTTA		
L15	Stk32cEx12-F	GGGCTTCCAACAAGTGATGT	59	975
	Stk32cEx12-R	ATGGAAGACTGCAGACATGG		
L16	Ly6dEx3-F	TGGGAGCCAGGGTACACTAA	60	700
	Ly6dEx3-R	AGGTGCCTTTCAGGTCCAGT		
L17	Gad2-Ex2-F	TGGACTTAAAGACCTGGTTCA	59	250
	Gad2-Ex2-R	AGGCTGCAGGCTCAGAGG		
L18	GdnfEx3A-F	GCGAGTTCCCGTGACTTAGA	60	822
	GdnfEx3A-R	CCCTGAGCTTCTGGTGACAT		
L19	Hes5ex3B-F	TGTGCAGAGTTGTCATTTGG	597	58
	Hes5ex3B-R	ACTGTTCCACCTCCGGTCAC		
L20	Rarres1ex1-F	CTTCTGGTCCTTCTGCTTGG	486	60
	Rarres1ex1-R	ACAAGCTCAGGCTGGGAAC		
L22	GdnfEx3B-F	AGAACAGCCGTCTGAGCACT	698	60
	GdnfEx3B-R	GCATTGCATCGAGGAGCTT		
L23	GdnfEx3C-F	GGGCCTGTTGGTTTACAAAG	943	59
	GdnfEx3C-R	ACCATAGGGACGAAGCAGGT		
L24	Kcnk9ex1-F	GGTCCAACTTCCCTTTAGCA	391	59
	Kcnk9ex1-R	CGTGTGCGCTACATCTCCTA		
L25	NppcEx1-F	GCGGCAGGTTGGATTATAAAG	381	60
	NppcEx1-R	GCTTTGAGGGAGCAAGTCC		
L26	NppcEx2-F	GGATGCTGGACTCAGCTGTC	498	60
	NppcEx2-R	CTCTTGGGTGCAGAGCTAGG		
L28	Wnt7bEx4B-F	AACCACAAAGGGACCATGAG	995	60
	Wnt7bEx4B-R	TTACCCTGCCATGTGTTGTG		
L29	Wnt7bEx4C-F	TCCGTGGGAAGCTTAATGTC	843	59
	Wnt7bEx4C-R	CATGGAGACGTCACACACAC		
L30	Foxg1ex2B-F	CCCTCTACTGGCCCATGTC	821	60
	Foxg1ex2B-R	TGAAATGGCAAAGCAGTCAC		
L31	Foxg1ex2C-F	CCAGCATCACCATCATCATC	836	60
	Foxg1ex2C-R	ACGTGGTCCCGTTGTAACTC		
L32	Prdm16ex9B-F	CAGTGGAGAGCAAACCTGAG	700	59
	Prdm16ex9B-R	GAGGGAGGGCCATGTTAATAC		

Fortsetzung Tabelle 13

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
L33	Prdm16ex17-F	CCTCACCTGTCCTGTGTGTG	747	60
	Prdm16ex17-R	GGGTTTGCCAATAAGCTGAA		
L34	Nkx6-2ex1-F	CAGTGCCTTCAGCCAATCAG	825	60
	Nkx6-2ex1-R	CTGCGAAGAAGAGCCAGTG		
L35	Nkx6-2ex5-F	AGGCAGGAGAATAAGCAGCA	361	60
	Nkx6-2ex5-R	CCTGTTGCCTGGACTAGCAC		
L36	Camk2bEx24B-F	CTGCAGTGCCTACCCATGT	745	60
	Camk2bEx24B-R	CCATGAACCTCTGCAGTCAA		
L37	Camk2bEx24C-F	CCTTGGTGACTGAGGGAATC	676	59
	Camk2bEx24C-R	AGAGAGGAGGGCCACCAGTA		
L38	Disc1ex3B-F	TTCATCCAACTCTCCCTTGG	495	60
	Disc1ex3B-R	GGGACCAGCACTCAGTCG		
L39	Disc1ex12-F	GGGTGTGTGTGTGTGTGTGA	475	59
	Disc1ex12-R	CGTGAGGGTCTATGGAAACA		
L40	Gad2-Ex16-F	AACACCACTTGCTATGGGATG	983	59
	Gad2-Ex16-R	CATATCTGCCTCTGCAGTCTGT		
L46	Foxg1ex2C-F2	TGATTCCCAAGTCCTCGTTC	984	60
	Foxg1ex2C-R2	TTTGAGTCAACACGGAGCTG		
L47	Foxg1ex1-F	TTTCGCCCTACTGGTGTTTC	400	60
	Foxg1ex1-R	TGGGTGCGTTCTGTGGAG		
L48	Wntb7ex2-F	AGTGGAGGAGGTGTTTGGTG	379	60
	Wntb7ex2-R	CTTGGCCTAGACTTGGCTTG		
L49	Wntb7ex3-F	CCTGGCTCAGAACTTGAAGG	396	60
	Wntb7ex3-R	TCCCACACCCATCAGACC		
L50	Hes5ex1-F	GCCAATTCACAGGCAATTTA	546	59
	Hes5ex1-R	GGCTCACCTTTGCTGTGTTT		
L55	Kcnk9Ex1-F	CGCCGCTTACAACTTGGAG	526	60
	Kcnk9Ex1-R	GCTGCGTTTAACCCTCGAC		
L56	Kcnk9Ex2-F	CCCTCCTCACTTGCTCTCTC	959	60
	Kcnk9Ex2-R	CGAGTTGGACCAATGGAAAT		
L57	Kcnk9Ex3-F	GGGACCTCTGTGGTAACTGC	250	59
	Kcnk9Ex3-R	TGGGCTTCCTGTCTATGTGT		

Tabelle 14. Primer-Sets zur QUASEP-Analyse.

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
73	A28202Gex5-F	AGATGAGAAAAGGCCAGCAC	59	189
	A28202Gex5-R	Biotin-CTGTGTGTCGTGAACACTGC		
	A28202Gex5-S	GACTTAAGTGAGGAG		
74	A34864Gex3-F	GCCTTCCCACTGGAATACTG	59	160
	A34864Gex3-R	Biotin-TCCTCAATCACATGACCCATAG		
	A34864Gex3-S	TCCCCCTGACATCC		
76	A38886G_B-F	TGATCTGGAGCTATCCCGAAG	59	293
	A38886G_B-R	Biotin-GAAGGGACATGAAGGTCTGC		
	A38886G_B-S	GGTCCCCTGAGATCTG		
77	A89294Gex5-F	Biotin-TCTTTGGAGAAGCCACCAAG	60	225
	A89294Gex5-R	TCAGTGCCACCAAGCATTAG		
	A89294Gex5-S	CTTGTCCACCAGCTTC		
78	C17603Tex12-F	ATAAGACCAGACTGGACCCAAG	59	375
	C17603Tex12-R	Biotin-TTCAGGTGATGAAGGCAGTG		
	C17603Tex12-S	GAGCCCAGGGATGT		
79	C39658Tex2-F	TGCGTGTTCCAAGATCTCTC	59	172
	C39658Tex2-R	Biotin-TGTCATGTCTTATCCTCCTCCTG		
	C39658Tex2-S	GGAGCTTGTCTTTGTGT		
85	T36711Cex1-F	TACTAGGCCCAGAGGTTGGTTC	60	216
	T36711Cex1-R	Biotin-TCAGCTTTGATGTCCGTCAC		
	T36711Cex1-S	CCTGCTGTGACATGC		
109	T36368Cex5-F	ATGGCCGACAGTTTCTGAAG	60	150
	T36368Cex5-R	Biotin-AACAAAGCCATCCAGCTTGT		
	T36368Cex5-S	TGACAACAAGACTGACCC		
110	c11-29538e14-F	ATGGATGATCTGCACAAGCAC	60	167
	c11-29538e14-R	Biotin-CATTGAGCTCCCCAGTTTTG		
	c11-29538e14-S	TTTACTGGAACCAAATACAC		
111	c12-14491e2-F	Biotin-CTCTGCTGTGTGCTGTCACC	60	162
	c12-14491e2-R	TTCCCCGAAGGTGTAGAGAG		
	c12-14491e2-S	GCACAGAATCTCCTTCAC		
112	C14-14850e4-F	TGCCAAATATATGGAAGAAATGC	60	228
	C14-14850e4-R	Biotin-TGTCAATAGATGGATTTCAGAAGG		
	C14-14850e4-S	TCCTACAGGGATTCTTTC		

Fortsetzung Tabelle 14

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
115	C17-C38428Tex3-F	Biotin-CCAGTATCCCCAGGTCCAG	60	215
	C17-C38428Tex3-R	TACCAGAGCAGTCGGAGAGG		
	C17-C38428Tex3-S	GCATGCACACATGGT		
122	C7-G39658Tex5-F	ACCGGAACTTGGTCTCTGTG	60	294
	C7-G39658Tex5-R	Biotin-TGTGGGTTCCTTTGTGAGTG		
	C7-G39658Tex5-S	GAAATGCAAAGACCTTCT		
123	C18-G40367Aex2-F	Biotin-CCGAACTGAGAGCTGTAGCC	60	361
	C18-G40367Aex2-R	AAAGTCATAGCCGTCCTTTGG		
	C18-G40367Aex2-S	GGCCAGAATATTTACAGTT		
124	C2-19550ex5-F	ACGTGAGGAAGACCTGGATG	58	202
	C2-19550ex5-R	Biotin-GTTCTTATTCCGCCCCTCTG		
	C2-19550ex5-S	AACCTTTTGCATACACAC		
125	C2-20441ex2-F	TTGGTCATGAGCTTCGTCAAC	60	244
	C2-20441ex2-R	Biotin-CACACCATTAAGAGTCATCTGGAC		
	C2-20441ex2-S	CCCCTGTTAGGCCC		
126	C2-35073ex10-F	Biotin-CCCAACGGAAATACATTCAG	59	276
	C2-35073ex10-R	TTGTGGAGCTCTTCCCTCGTAG		
	C2-35073ex10-S	GCTCTTCCCTCGTAGC		
127	C2-45137ex4-F	Biotin-GAGAGGCTGCAGCATTGAG	60	136
	C2-45137ex4-R	ATCAGACACCATCAAGGAAAGG		
	C2-45137ex4-S	TCCCACCATGCAGA		
128	C2-79954ex4-F	Biotin-TCTCCAATACTCGAAGATAGACATTG	60	211
	C2-79954ex4-R	TTGCTAGACTTCAAGTGTGTTGCTC		
	C2-79954ex4-S	TTCAAGTGTGTTGCTCC		
129	C4-29920ex2-F	Biotin-CAGCTCTCGCGTCCTTTC	60	229
	C4-29920ex2-R	ATGCTGATGTCCAGGGAGAG		
	C4-29920ex2-R	CACACTCGGATGGCA		
130	C4-43873ex3-F	Biotin-GTCATGGCAGTGACCCAAC	59	174
	C4-43873ex3-R	TCGGTCCTTCTCCTTGGAAC		
	C4-43873ex3-S	ACGCTGCGCAGGC		
131	C12-A14491Ge5-F	TCAGCAACCACATCTTCATTG	59	182
	C12-A14491Ge5-R	Biotin-AAGGTGCTGCTCAGTCTTCG		
	C12-A14491Ge5-S	CCTTCAGGCCTCAGC		
132	C14-14850e4-F2	Biotin-AGGCTTTAAGGCTGTTGCAG	60	196
	C14-14850e4-R2	ACCAATTGAAAAGCCACAGG		
	C14-14850e4-S2	GGTGTTAAGTCTGGACAGTT		
133	G39658Tex3-F	Biotin-TCCAGGAGCTTGTCTTTGTG	60	210
	G39658Tex3-R	CGTTGAATGTCACTGCTTCC		
	G39658Tex3-S	TCTCTCTTGGAGAACAACT		
134	G86590Aex5-F	Biotin-GAGGCAGGAGTCAGCAGAAG	60	220
	G86590Aex5-R	TTGGCCTGTCATCACTTTTTG		
	G86590Aex5-S	TCAAACTCTGGGAACC		
136	T37730ex5-F	Biotin-AAGGCATGTCACCTTCATCC	60	188
	T37730ex5-R	CGCAGTTGGATGTTGAAGTC		
	T37730ex5-S	GCGGAACACGTCAC		
138	Zfp28ex9-cF	AAGCGAGAGCTGACAGGAAG	56	330
	Zfp28ex9-cR	Biotin-ATTGGCTCTTGCCTGAATTG		
	Zfp28ex9-cS	GGGATTCTGACTATGTGTT		
140	C2-79383ex3-F	Biotin-TTTGGGAGCACAACCACAGA	60	249
	C2-79383ex3-R	TTTGGAATGGTTTGCCCACT		
	C2-79383ex3-S	AGGAAGTTGCCTGGTA		
141	C2-79954ex4-F2	Biotin-CCTTCTCCAATACTCGAAGATA	58	327
	C2-79954ex4-R2	GGCAGGGTCATAAAGTGTTC		
142	C4-12218ex6-F	Biotin-TGAATGAAGCCAAGCCAAGTA	60	93
	C4-12218ex6-R	ACAGGGCAGGCTTTCCAGTA		
	C4-12218ex6-S	GCAATCAATTCAGCTCTA		
143	C4-14418ex2-F	ACGCCCAGGCAGAGTTCC	58	93
	C4-14418ex2-R	Biotin-CCTCCGGTCAGTGTTCAC		
	C4-14418ex2-S	TTTTGATGGATGCAAC		
144	C2-35073ex10-F3	Biotin-ACCCAACGGAAATACATTCAG	60	266
	C2-35073ex10-R3	TGGTTGTCTACACACTTAAACTA		
145	C2-37602-F	Biotin-AGCTCTTCTGGACTTAGAGTT	58	118
	C2-37602-R	GAAAACACGTCCATGTTCTG		
	C2-37602-S	GAAACAAACCCAGCC		
146	C3428Tex3-F3	Biotin-ACACCAGTACCCAGTATCC	58	108
	C3428Tex3-R3	AATGTTCCCCATAGGACCAC		
	C3428Tex3-S3	CACAGCTCAATCTTCCA		
147	C4-34678-F	GCATTCTCCGACCTCTGC	58	106
	C4-34678-R	Biotin-CACTGGGTGGGGAGTAGC		
	C4-34678-S	GGCCTGAGGACGCCC		

Fortsetzung Tabelle 14

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
193	pC10-37266-F	Biotin-CAACGGTATAATCAAGACTTTC	60	147
	pC10-37266-R	GTAATTCCGAGTCTTAGCTCT		
	pC10-37266-S	GACACTAAGTAGAGGTGGAA		
194	pC12-18142-F	AGCTGAGCCATTGTTGTCC	58	180
	pC12-18142-R	Biotin-TCCTTGCCTACCCCCTTC		
	pC12-18142-S	GGAAGAGAGGATGGC		
195	pC12-39463-F	Biotin-TCCAGAGGAACCCGGATAAT	58	127
	pC12-39463-R	CTTAACTACCGAGCCGTTGC		
	pC12-39463-S	TGCTGTGAGTTTCCC		
196	pC12-40459-F	ACGAGGAGATACATCCTTAAAC	58	235
	pC12-40459-R	Biotin-TATGCATCAAGACATGTAACAC		
	pC12-40459-S	AAAGCTGCTTGAAATG		
197	pC2-16545-F	CGGCCCCATTCACATGGT	58	118
	pC2-16545-R	Biotin-GGCTCACCAATCGATGACG		
	pC2-16545-S	GGCTATTACTGGACAATGTT		
198	pC2-18518-F	Biotin-GTCCAATTATCGCGCTGAAC	58	72
	pC2-18518-R	AGCAGGGTGGTATCGTTCTC		
	pC2-18518-S	CTCTGGGGGGACAGT		
199	pC4-34678-F	GCATTCTCCGACCTCTGC	58	113
	pC4-34678-R	Biotin-CACTGGGTGGGGAGTAGC		
	pC4-34678-S	GGCCTGAGGACGCCC		
200	pC4-35722-F	ATGTCAAGGGGTGGGAAGG	60	90
	pC4-35722-R	Biotin-CCACCTCCACCCCGAGAT		
	pC4-35722-S	GGTGGGAAGGAAGACT		
201	pC4-42496-F	Biotin-GTTGTGTGCCACTTTTTCTAC	58	174
	pC4-42496-R	GCTAGGCAAAGTACAACTCT		
	pC4-42496-S	GCAACCAACAGTGAAGAT		
202	pC6-39087-F	Biotin-GCCATGATGCTTCTGTCACTA	58	123
	pC6-39087-R	TACAAGCGGATGGCTTAGGAA		
	pC6-39087-S	ATGGCTTAGGAAGTACATAG		
203	pC7-079260-F	TTGGGAGATGGCCAGAAGG	58	98
	pC7-079260-R	Biotin-CTCAACTAAGTGGCATCAAC		
	pC7-079260-S	CTTCCTGGGCCATGT		
204	pC7-14392-F	Biotin-AGCCTAGACGTGTTATTAGTT	58	140
	pC7-14392-R	AGGATATGCAGGATGTCACA		
	pC7-14392-S	TGAAAAATTCCACATCAG		
205	pC7-34189-F	Biotin-AGGGCAACTTGAGACACAAA	58	172
	pC7-34189-R	TTTTACAGATCACCAGGAGAAT		
	pC7-34189-S	AAGCTAGCCGAATGTAA		
206	pC7-39482-F	TGCAAAAGAATATCAGTACACA	58	202
	pC7-39482-R	Biotin-TCAAGAATAATTACGAGGTAGC		
	pC7-39482-S	ATAATTTCATTTGAGGCA		
207	pC9-34922-F	TTGCTGTGACTGTACACTTC	62	97
	pC9-34922-R	Biotin-ATTCAAACACTCAGGGGCTAT		
	pC9-34922-S	TGTACACTTCACATGAGTTC		
208	pC10-38360-F	ATTTACTGGGGCAGAAGTTC	60	84
	pC10-38360-R	Biotin-CTGTCTTGTGAATGCCAATGA		
	pC10-38360-S	AGTGCCTGACAAGGAG		
209	pC14-37626-F	CCATTTTATTATGTGGGCTGTG	58	112
	pC14-37626-R	Biotin-GCTGGTTTATGGCTTTCTTTCT		
	pC14-37626-S	GTGGGCTGTGTAGTACAT		
210	pC14-38100-F	Biotin-TGGTTGTCTTTACTGGGTATAT	58	140
	pC14-38100-R	CAGATATTATTCCACTGTACATG		
	pC14-38100-S	ATCAGGGGTCACTGC		
211	pC17-39514-F	Biotin-CCCAATAAAGTAGCCCAAGA	60	123
	pC17-39514-R	GAGGTTCCTTCTTTCAGATAT		
	pC17-39514-S	TCAGAAAACTATCCAGTGTA		
212	pC17-40032-F	Biotin-TCCTGTCAGGTGCTCTGAA	60	152
	pC17-40032-R	AGGGGAGTGAGGCTGTGA		
	pC17-40032-S	AGTGCTGGCGAGGGT		
213	pC17-44118-F	Biotin-CCAGTTCCCAAAGTATAATGC	60	118
	pC17-44118-R	CCTTGGTTGAACACATAATAGT		
	pC17-44118-S	TTCAAGCCGTTATCAAT		
214	pC7-37653-F	Biotin-TGTTCAAAGCCACACCAGAAA	60	166
	pC7-37653-R	TTATGTTATCCCCCTCCAAAG		
	pC7-37653-S	CAAGAATAGAGCTGCTTAAC		
215	pC7-43056-F	Biotin-CCAGTTAGGTTTGGCTTTTTGT	60	129
	pC7-43056-R	GGGCAGATTGAAGTCTGTGA		
	pC7-43056-S	CAGCCAGGGCTACAC		
216	C10-37266-F2	Biotin-TTGTTTGGTTTTGCTCATTTAACT	60	168
	C10-37266-R2	CGTGAGAGCTTTTGAAATAATAT		
	C10-37266-S2	GCATTTAGCCAAAGAGATT		

Fortsetzung Tabelle 14

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
217	pC17-38428-F	Biotin-GGCGTGTTTGTATGTTTTAGAC	58	71
	pC17-38428-R	AAGGCTCCAGCACATTCAAG		
	pC17-38428-S	TCAAGCCATCATTCG		
218	pC17-45985-F	TCGGGTTTGTGGTTGCTTGG	60	205
	pC17-45985-R	Biotin-CTGGAGTACCTCAGTTGACA		
	pC17-45985-S	AGGTAAGACAACTTAGCCT		
219	pC2-37521-F	Biotin-CAAAAGGCAAGGGTGGTTCT	60	94
	pC2-37521-R	ACAAGAGGGTCAACTGAGC		
	pC2-37521-S	CAGTGCAGCCTCCAA		
220	pC2-38677-F	TCCTCTTCGCCAGGCTCCA	62	132
	pC2-38677-R	Biotin-GCCGTGTAGCAGATGCAGT		
	pC2-38677-S	GCCCCCTGCTGGCGT		
221	pC2-45137-F	GCGCCTTCTCAGAGATCCAT	60	78
	pC2-45137-R	Biotin-AAGCACCACGTACCCACTC		
	pC2-45137-S	GAGATCCATAAGCTTCCT		
222	pC4-34678-F2	Biotin-GTAGTGAGAATTTGGGGACCT	60	97
	pC4-34678-R2	TGAAGCTCATTCCAGCCTAGA		
	pC4-34678-S2	AAGTTAGCTGGAGGATTAC		
223	pC4-45834-F	CGGTCTGACAGCCTACCTA	58	157
	pC4-45834-R	Biotin-AGCCCCTGCCTCTGTGAAA		
	pC4-45834-S	ACAGCCTACCTAGATTCTAC		
224	pC7-37730-F	GCACGTGGGGTGACTTCC	58	78
	pC7-37730-R	Biotin-GTGGCCCCGCAGTGTGAA		
	pC7-37730-S	CCATTACAAGACCTTTGA		
225	pC7-38886-F	Biotin-GCCTTGGAAGAGATGGATGT	60	84
	pC7-38886-R	GCCCCAATTAAGGGGTTGT		
	pC7-38886-S	ACAATTGGCGTAACCT		
229	pC2-79954-F3	Biotin-AGTTCAAATGCCAAGGATGC	60	163
	pC2-79954-R3	TTGCGTTAGGAAGCTAGACT		
	pC2-79954-S3	TCAAGTGTGTTGCTCC		
230	pC4-43873-F	Biotin-GGGCTACGCCAACAGCAG	58	90
	pC4-43873-R	CGACAGAGCTGGCGGAAG		
	pC4-43873-S	GGAAGGCGTAGAGAAC		
231	pC7-17829-F	Biotin-TGGCGAGAAAGTGGAACGG	60	139
	pC7-17829-R	AGGTACTGCAGGCGGGTAA		
	pC7-17829-S	CAGACTGGGATTGCTG		
232	pC7-38418-F	Biotin-ACTCTCTGCCTTTCTACAATA	60	222
	pC7-38418-R	GAAAGACTATCCCATCATGC		
	pC7-38418-S	CCTTGATGAGCAGAGAT		
233	pC2-80554-F	TTGATTGTGAACATGCTGTAAC	58	278
	pC2-80554-R	Biotin-AAACAGGACTTGCTGGTAAGTA		
	pC2-80554-S	TTGGTCTGTTATGTGAATG		
234	pC4-30742-F	TACTTTTGGGGCATAGTGGTA	58	116
	pC4-30742-R	Biotin-TGTGGACTTGCAGTTCACTCT		
	pC4-30742-S	CTGAGGATTGGTACCC		
235	sC6-38180-F	TGGAATATTGTTGGACCAAATAG	62	117
	sC6-38180-R	CCATGTGGTCGCTGAGAT		
	sC6-38180-R2	TTTAACTCTAGGCAAGGTTTCTTT		
236	C12-18142-F2	Biotin-CCATCTTGAGTGTGACTTAC	60	175
	C12-18142-R2	GCTGAAGGGTAGGCTTGAT		
	C12-18142-S2	GCCTCTAGGCTTCCA		
237	pC2-45137-F2	Biotin-TTGAGGCCACTCCAGACACT	60	109
	pC2-45137-R2	TTCCCTGGGACCAAGACATG		
	pC2-45137-S2	TTCCCACCATGCAGA		
238	pC7-38418-F2	ACCTTGGATTCCACTTGTCTA	60	72
	pC7-38418-R2	Biotin-CATGGGGGAATAAATGGAGTT		
	pC7-38418-S2	CAGAGCACCGGAAAT		
240	pC4-35229-F	ACGTGCTGGAATTGGTACCT	58	106
	pC4-35229-R	Biotin-TAGTCTGTGTCCGGGTCGA		
	pC4-35229-S	AACAGCTGCCCGGGA		
241	pC6-36818-F	Biotin-GCCCTCGCCAACCCTACG	62	184
	pC6-36818-R	CGAAGCGGGCTGCCTTCA		
	pC6-36818-S	ACCTAGGCTGGCGGTG		
244	pC6-38180-F	TTCAAATCTCAGCGACCACAT	62	148
	pC6-38180-R	Biotin-ACACTGTATCTGCAGTACAC		
	pC6-38180-S	CTCACAACCATCCGTA		
245	pC6-34724-F	Biotin-ACAATCTTCTCAGGGCCCAAT	60	87
	pC6-34724-R	AGTGGAGATGTTGGTGATATGT		
	pC6-34724-S	ATCATGGTTGTGGAACT		
252	pC10-38360-F2	TGTGGCTTTGTTAGAGTAGGTT	60	154
	pC10-38360-R2	Biotin-TGGCAGTTAGTTTCATACCTAAT		
	pC10-38360-S2	AGCCCAAGTCAGGCC		

Fortsetzung Tabelle 14

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
253	pC18-40367-F2	Biotin-CCCAGGCACATCTTATAGAAT	58	59
	pC18-40367-R2	TTTCAGCCCTCTAGTCTCAC		
	pC18-40367-S2	TCCAGCATCCTTGGTC		
254	pC7-28202-F2	Biotin-TGCTTCCAAGGCAGGACTC	60	84
	pC7-28202-R2	CCCTCTGAGAAGGCTCTGT		
	pC7-28202-S2	TCAAGTTTAGTGCTCCAA		
255	pC7-37730-F2	Biotin-GGGAGAAGCTGCCCTACA	60	205
	pC7-37730-R2	CACATAGCCACTGCTCTTC		
	pC7-37730-S2	GCTCTTCTCCATCAGGAGAC		
256	pC6-34724-F2	Biotin-CCTTCTTCCTTTAGAGTATTAG	58	105
	pC6-34724-R2	AAATCTAGAGACAGCCAAGAG		
	pC6-34724-S2	CCAAGAGAAGAGGGCA		
257	pC2-38367-F	GCCACAGCTTGGATTATTAATT	62	210
	pC2-38367-R	Biotin-CGTCTATAGGAACCATTCTTATT		
	pC2-38367-S	CCATGTTCTGCCAGG		
258	pC6-34724-F3	ATTCTGCGGATTCCTTCTTC	58	149
	pC6-34724-R3	Biotin-TGCATCTGTAAATCTAGAGAC		
	pC6-34724-S3	CCACAACCATGATCCC		
271	h43056-AC-F	Biotin-CATGGCCCATATGCTACTATTT	58	143
	h43056-AC-R	ACAGATAGATCAGGCAGACAAA		
	h43056-AC-S	TACTCATTCTTTTCTAATAT		
272	h43056-delA-F	GATTTCCTCTTTGCGACATG	58	100
	h43056-delA-R	Biotin-TATGGGCCATGCTATATGAG		
	h43056-delA-S	GACATGAAGAAACGTC		
273	USP-3764574-F	Biotin-TTTCTCTTCCTGTGATGTATGA	60	126
	USP-3764574-R	ATTGAACTCTACGAATGTATTG		
	USP-3764574-S	GAAGTTCTGGATGTTGAG		
274	ZIM-4801433-F	Biotin-ACACTGGGAAGAGAGCTTAT	58	100
	ZIM-4801433-R	GGCTTTTCCCCAGTATGGAT		
	ZIM-4801433-S	TATGGATTTTTTTATGTTGA		
275	ZNF-917340-F	Biotin-GAGCCCTGAATGTGATGGTTT	60	126
	ZNF-917340-R	ATCATGGGATCTTTACCTGAC		
	ZNF-917340-S	CTCACATGAGTTATGGC		
287	phuZim2-F	GTGGACTTCAGCCCAGAG	59	145
	phuZim2-R	Biotin-CCTCTTCCAGGCGTGAGAT		
	phuZim2-S	CAGCCCAGAGGAACT		
288	phuPeg3-F	CTTCTTTCTGCAGCCATTTC	59	102
	phuPeg3-R	Biotin-CCAATCAATCTGGGTCACAAA		
	phuPeg3-S	TGCAGCCATTTCATCT		
289	pmPeg3-F	CAGTTTTAGCTGTTCCCTATG	60	105
	pmPeg3-R	Biotin-TAGGTTTTCCCTCAATCTAGTT		
	pmPeg3-S	TCAGTGTATATTTGGGAAAG		
290	pmUsp29-F	Biotin-GCAGCCAAAATGGTAAGAGTA	59	142
	pmUsp29-R	AATTTGGGGAATAACACTTGTA		
	pmUsp29-S	AGAAGACACCCAGTGC		
L21	Dok-1209-F	Biotin-GGGGCCTTTATGATCTGCCT	60	103
	Dok-1209-R	TCATCGGTGGCAGGGTTGT		
	Dok-1209-S	AGCCCTCTTCCTTCA		
L27	Rarres150-F	Biotin-TTGGACAGGGAGGCGAGGA	60	180
	Rarres150-R	CTGGCATCGGGGGTTGAC		
	Rarres150-S	CGTGGATGGGGAAAG		
L41	Irx1318-F	Biotin-AGACCAAGCCCCGCATCT	58	127
	Irx1318-R	GTCGTCGCAGATACACCC		
	Irx1318-S	GAAACTCAGTCTGGCTC		
L42	Ntng540-F	Biotin-CCTGGCAGCCCTACCAGT	60	93
	Ntng540-R	GGGCGCTGGAGGGTGACA		
	Ntng540-S	CATGTCTCGGGCAC		
L43	Stk1506-F	TCCTCCAGGCCCAGAGAC	60	93
	Stk1506-R	Biotin-GGGCAGATGGAGCCACAC		
	Stk1506-S	GACGTCAGATATGACAGATT		
L44	Ly6d422-F	Biotin-CAGCCTGCTCACTGTTATG	58	101
	Ly6d422-R	TAAAGCATGGGCCGTGGG		
	Ly6d422-S	GGGGACACTGACGACT		
L45	Gdnf384-F	Biotin-ATGCCTGAAGATTATCCTGAC	58	165
	Gdnf384-R	ACCTTTCCCTCTGGAATTCT		
	Gdnf384-S	CTCTCTTCGAGGAAGC		
L51	Prdm2353-F	TGCGGAGTGCCCCTTCGA	60	75
	Prdm2353-R	Biotin-AGGGGGACCTTGGGTGCG		
	Prdm2353-S	CCACCAAACCAAAAGAG		
L52	Kcnk477-F	GCCTGTACCTTCACCTACC	58	78
	Kcnk477-R	Biotin-CTCCTCGCGCATCTCATG		
	Kcnk477-S	TGCCGCGGTGTTC		

Fortsetzung Tabelle 14

Primernummer	Primername	Sequenz 5´- 3´	Tmelt [C°]	Länge [bp]
L53	Nppc200-F	Biotin-GGCACCATGCACCTCTCC	60	90
	Nppc200-R	TGGTGTCCCGGGCTTGGC		
	Nppc200-S	TGTCCCGGGCTTG		
L54	Cdk1056-F	Biotin-AGCTGAACACAGCCTGAGGT	62	72
	Cdk1056-R	ACTCGCAGCCGCCAATGTG		
	Cdk1056-S	AATGTGTTCAGATGACGA		
L58	Camk2b-2756-F	GGGAAGTTCCTCTCTGAAG	58	76
	Camk2b-2756-R	Biotin-TTCCCAGTATCCGTCCAGTA		
	Camk2b-2756-S	GTTGGCAGCAGACC		
L59	Gad2-2279-F	TGGTGTCACAGTAGAGTACAG	62	116
	Gad2-2279-R	Biotin-GGCCTTGAACTTAGAAATCTGT		
	Gad2-2279-S	GGGTGGTGGTGGC		
L60	Wnt7b-1661-F	Biotin-GGACACCCTTCCCTCACT	58	90
	Wnt7b-1661-R	AGCTTCCGTGCCAATTCTG		
	Wnt7b-1661-S	AGGGAGGGGCCGG		
L62	Foxg1-1165-F	AGTCCCGAGAAGCGCCTGA	62	92
	Foxg1-1165-R	Biotin-TTCTGCCAGCCCTGCTTGTT		
	Foxg1-1165-S	CGCTCAATGGCATCT		
L63	Nkx6-1113-F	TTAGCCCGTGTGGTGGCAG	62	159
	Nkx6-1113-R	Biotin-CGCGCCAACGTCCGCTTTA		
	Nkx6-1113-S	GGCCGGAGAACCC		
L64	hKCNK9ex2-F	TACCATTGGGTTCGGGGACTA	64	79
	hKCNK9ex2-R	Biotin-GCTAAAGGCCACGTAGAGCG		
	hKCNK9ex2-S	CCTGCAGACCAAGGG		

7.4. Identifizierte tSNPs der Kandidatengenen

Tabelle 15. Transkribierte SNPs der Nikaido-Kandidatengene.

Genname	EICO-ID	GenBank-Eintrag	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
unclassifiable	B130039H18	AK045137	prox2 (4.24)	2	147	A	G
					510	C	T
HPRP18 (PRE-MRNA SPLICING FACTOR SIMILAR TO S. CEREVISIAE PRP18) homolog [Homo sapiens]	9430083N14	AK035073	prox2 (4.5)	1	1242	A	G
f-box only protein 3	4921535A12	AK019550	cen2 (103.9)	3	610	C	T
					1104	T	C
					1574	G	C
integrin beta 4 binding protein	2700008E01	AK012218	dist2 (155.8)	4	95	G	T
					194	T	C
					203	C	G
					1107	C	T
unknown EST	A730085E23	AK080554	dist2 (156.2)	2	364	A	C
					384	A	T
unclassifiable	A130099D17	AK038367	dist2 (158.7)	12	2241	G	A
					2248	C	T
					2307	A	G
					2346	G	A
					2354	A	G
					2448	T	G
					2600	G	A
					2882	T	A
					2893	T	C
					2894	G	A
					2928	A	G
					2936	T	C
MELANOCORTIN-3 RECEPTOR (MC3-R)	A230054D12	AK038677	dist2 (172.3)	2	753	T	C
					795	T	C
					1089	G	A
bone morphogenetic protein 7	9430012M17	AK020411	dist2 (172.9)	4	750	C	T
					802	T	G
					857	C	T
					859	A	G
RAE1 RNA export 1 homolog (S. pombe)	9630054N05	AK079383	dist2 (173.0)	1	399	T	C
unknown EST	A530024C12	AK079954	dist2 (173.1)	1	570	C	T
hypothetical Cysteine-rich region containing protein	4932434E15	AK016545	dist2 (173.3)	3	207	A	G
					259	A	G
					260	C	G
unclassifiable	A130023E15	AK037521	dist2 (173.7)	1	467	G	A
unknown EST	9030418K01	AK018518	dist2 (178.1)	1	404	C	T
unknown EST	A130029E23	AK037602	dist2 (179.2)	1	491	G	A
nuclear receptor binding factor 1	B230313J17	AK045834	dist4 (130.0)	2	161	A	T
					393	C	G
unknown EST	A630096I18	AK042496	dist4 (130.1)	1	280	A	G
unknown EST	9430023J17	AK034678	dist4 (130.1)	4	929	A	G
					1057	T	C
					1124	C	T
					1182	C	T

Fortsetzung Tabelle 15

Genname	EICO-ID	GenBank-Eintrag	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
DELTA-TYPE OPIOID RECEPTOR	A830044L10	AK043873	dist4 (130.3)	2	813	A	G
					1047	G	T
unknown EST	9530096D07	AK035722	dist4 (130.5)	2	454	G	C
					456	A	G
hypothetical Type-1 copper (blue) domain/Leucine-rich region containing protein	4931440N07	AK029920	dist4 (132.5)	1	421	A	G
similar to CDNA FLJ13976 FIS, CLONE Y79AA1001594, WEAKLY SIMILAR TO HYALURONAN-MEDIATED MOTILITY RECEPTOR [Homo sapiens]	5730495A12	AK030742	dist4 (132.5)	1	1135	G	A
similar to CDNA FLJ10747 FIS, CLONE NT2RP3001799 [Homo sapiens]	A230080D12	AK038972	dist4 (133.2)	0			
unclassifiable	9530059D22	AK035517	dist4 (133.9)	0			
inferred: Mus musculus von Willebrand factor A-related protein (Warp) mRNA, complete cds	9530003H10	AK035229	dist4 (153.6)		86	A	G
stromal cell derived factor 4	3830404L10	AK014418	dist4 (153.8)	5	25	A	T
					49	C	T
					58	T	C
					107	G	A
					158	A	G
unknown EST	A230094E16	AK039087	prox6 (4.6)	1	355	A	G
pyruvat dehydrogenase 4	9930014N18	AK036818	prox6 (5.2)	3	70	C	G
					122	C	T
					559	C	G
unknown EST	A130084J17	AK038180	prox6 (29.6)	2	565	G	A
					573	T	A
hypothetical protein	9430029A11	AK034724	prox6 (32.3)	4	28	C	T
					813	A	G
					814	A	delA
					837	A	G
unclassifiable	A230072E17	AK038886	prox7 (no more data)	5	20	A	G
					170	A	G
					173	T	C
					271	A	G
					549	G	C
ZFP77P (FRAGMENT) homolog [Mus musculus]	A330079D15	AK039658	prox7 (no more data)	8	212	C	T
					229	C	T
					236	C	A
					269	T	G
					292	G	A
					607	G	T
					806	T	G
					915	A	G
zinc finger protein 28	2700045L05	AK028202	prox7 (no more data)	7	107	G	A
					223	G	A
					629	A	G
					636	T	C
					900	A	G
					923	A	G
					1274	C	G

Fortsetzung Tabelle 15

Genname	EICO-ID	GenBank-Eintrag	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
unclassifiable	A730051H19	AK043056	prox7 (no more data)	4	367	C	T
					1555	G	A
					1626	T	A
					2371	C	G
unclassifiable	9530064F02	AK079260	cent7 (46.3)	14	729	-	insT
					790	A	delA
					1037	G	A
					1054	T	C
					1068	A	G
					1078	G	A
					1089	C	T
					1112	G	C
					1167	G	A
					1292	A	C
					1352	T	C
					1488	A	G
					1515	T	C
					1569	G	A
unclassifiable	A130033P09	AK037653	cent7 (46.8)	2	701	G	T
					1042	A	G
unknown EST	A230073K19	AK038900	cen7 (46.8)	0			
unknown EST	9330162G02	AK034189	cent7 (46.8)	3	1947	A	C
					2145	A	G
					2671	C	A
unclassifiable	A230006K03	AK038418	cent7 (48.0)	12	1289	A	G
					1378	C	T
					1380	A	T
					1541	G	A
					1542	A	C
					1725	A	G
					1780	G	T
					1782	C	T
					2335	T	C
					2562	A	G
					2642	A	T
					2643	A	C
phosphatidylserine synthase 2	9630061G10	AK036368	dist7 (128.8)	1	606	T	C
hypothetical Leucine-rich repeat containing protein	5730427C23	AK017603	dist7 (128.8)	1	1968	C	T
GL008 homolog [Homo sapiens]	9830166G06	AK036717	dist7 (129.1)	2	70	T	C
					1636	G	A
similar to MUCIN (FRAGMENT) [Homo sapiens]	A130042M24	AK037730	dist7 (129.5)	2	325	T	C
					544	A	T
cathepsin D	F730002E02	AK089294	dist7 (130.0)	2	678	A	G
					936	A	G
unclassifiable	9430050L03	AK034864	dist7 (130.7)	6	845	C	T
					875	A	C
					1189	A	G
					1337	G	T
					1424	T	C
					1472	A	G

Fortsetzung Tabelle 15

Genname	EICO-ID	GenBank-Eintrag	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
weakly similar to ORF 2 PROTEIN (FRAGMENT) [Rattus norvegicus]	A330049H05	AK039482	dist7 (130.9)	6	1151	G	C
					1341	A	G
					1761	A	G
					1859	A	G
					2082	A	G
					2171	A	C
fibroblast growth factor 15	5730551C10	AK017829	dist7 (132.6)	1	236	T	C
unclassifiable	9430062P05	AK034922	cen9 (89.9)	7	2021	G	A
					2042	A	C
					2046	T	G
					2926	T	C
					2940	T	C
					2953	C	T
					3100	C	T
unknown EST	A130002D10	AK037266	prox10 (10.8)	6	857	A	G
					867	G	C
					1110	T	C
					1294	G	A
					1443	T	C
					2010	C	G
hypothetical protein	A130098G01	AK038360	prox10 (12.0)	2	1147	T	C
					1396	G	C
hypothetical protein	4921519G13	AK029538	prox11 (23.4)	1	1780	C	T
unclassifiable	A330045P15	AK039463	dist12 (103.7)	2	336	A	G
					338	T	C
unknown EST	6330404N04	AK018121	dist12 (104.2)	0			
unknown EST	6330408G06	AK018142	dist12 (104.2)	4	362	C	T
					363	G	A
					682	C	T
					836	C	A
unknown EST	A430099D21	AK040459	dist12 (105.3)	4	403	A	G
					426	G	A
					892	C	T
					1101	G	A
similar to X-RAY REPAIR CROSS-COMPLEMENTING PROTEIN 3 (X-RAY REPAIR COMPLEMENTING DEFECTIVE REPAIR IN CHINESE HAMSTER CELLS 3) [Homo sapiens]	4432412E01	AK014491	dist12 (106.5)	6	190	A	G
					597	G	C
					667	A	G
					809	A	G
					949	A	G
					952	A	G
unknown EST	A130078K24	AK038100	cent14 (64.4)	3	311	G	A
					427	A	G
					881	C	T
unclassifiable	A130030G06	AK037626	cent14 (65.1)	5	1954	T	G
					2068	G	A
					2256	A	G
					2410	T	C
					2587	A	G

Fortsetzung Tabelle 15

Genname	EICO-ID	GenBank-Eintrag	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
hypothetical Leucine-rich repeat/bZIP (Basic-leucine zipper) transcription factor family/Leucine-rich repeat, typical subtype containing protein	4921509B22	AK014850	cent14 (66.8)	1	1031	A	G
unknown EST	A230009B12	AK038428	prox17 (9.8)	5	306	C	T
					326	C	T
					390	C	T
					751	T	C
					786	T	C
unclassifiable	A330053J22	AK039514	prox17 (12.1)	9	178	A	G
					239	A	G
					271	A	G
					284	T	C
					419	G	A
					512	A	C
					557	C	G
					567	A	G
					584	G	A
unknown EST	A830092C01	AK044118	prox17 (12.4)	4	525	T	A
					1261	C	A
					1467	T	C
					1639	C	G
unknown EST	A430047I21	AK040032	prox17 (13.7)	5	1159	C	T
					1429	A	G
					2048	T	G
					2135	T	A
					2295	C	T
programmed cell death 2	B230330P13	AK045985	prox17 (14.3)	1	552+18 (IVS2)	T	C
Niemann Pick type C1	A430089E03	AK040367	cent18 (12.5)	3	239	G	A
					373	T	C
					423	C	T

Tabelle 16. Identifizierte tSNPs der Luedi-Kandidatengene.

Genname	GenBank-Eintrag (mRNA)	Chromosom (Mb)	Anzahl der SNPs	SNP-Position	C57BL/6-Allel	Cast/Ei-Allel
Foxg1	NM_008241.1	cent12 (50.2)	2	1165	T	C
				2411	T	C
Kcnk9	NM_053405.1	cent15 (110.2)	1	477	C	T
Prdm16	NM_027504.2	dist4 (153.3)	1	2353	G	T
Nppc	NM_010933.2	cent1 (88.5)	2	200	T	C
				362	G	C
Wnt7b	NM_009528.2	dist15 ((85.4)	2	1207	C	G
				1661	C	T
Dok1	NM_010070.3	cent6 (82.9)	3	770	T	G
				848	T	C
				1209	C	A
Nkx6-2	NM_183248.2	dist7 (139.4)	2	1092	C	T
				1113	G	A
Cdk6	NM_009873.1	prox5 (3.3)	1	1056	T	G
Rarres1	XM_130987.7	cent3 (67.6)	7	135	C	T
				150	C	A
				188	C	T
				300	A	G
				306	C	T
				397	G	T
				424	C	T
Camk2b	NM_007595.2	prox11 (5.9)	3	2756	T	C
				3088	G	T
				3142	C	T
Ntng2	NM_133500.1	prox2 (29.0)	2	525	C	T
				540	A	C
Irx4	NM_018885.1	cent13 (73.7)	2	915	A	C
				1318	A	G
Stk32c	NM_021302.1	dist7 (139.0)	3	1506	C	T
				1778	A	G
				1862	C	G
Ly6d	NM_010742.1	cent15 (74.5)	2	408	T	C
				422	C	A
Gad2	NM_008078.1	prox2 (22.4)	4	2279	T	C
				2293	G	A
				2369	A	C
				2370	G	A
Gdnf	NM_010275.1	prox15 (7.7)	5	384	C	T
				490	A	G
				1982	C	T
				1992	T	C
				2102	A	G

7.5. BLAST-Ergebnis AK043056 gegen das Humangenom

```
Query location     : MMU7prox                883 to     1362 (+)
Database location  : AC004792.1.1.36200     1929 to     2420 (+)
Genomic location   : 19                  62049222 to 62049713 (+)

Alignment score    : 142
E-value            : 2.5e-38
Alignment length   : 513
Percentage identity: 66.47

Query:  883 AACTGTCTTTTGAATGTAGATTTTGAAAATCTACTGCATTGCAGTCAAATACAACTAGTC 942
            |||| ||||||||||  ||||||  ||| |  | ||||||  | ||| |||  | | | |
Sbjct: 1929 AACTATCTTTTGAATACAGATTTCCAAATTA-ATTGCATTATACTCATATAGCA-TGGCC 1986

Query:  943 --TATGCGA---TTTGCCATTTCCGT-AAGGCGTCCTT--TC-TCTTAGGGTTTGATTGA 993
              ||||| |   |||  |||||  || || || |||||  |  ||||| | | || || |
Sbjct: 1987 CATATGCTACTATTTCACATTT--GTCAAAGCATCCTTGGTGATCTTAAGATATGGTTAA 2044

Query:  994 TTTTGAATTGTTCAAGGTGTTAGATAAGAAACAGTATTCTCTAATTGCTAGAAATTAATT 1053
            |||| ||||||||||  | ||||| |||||  |||||| |||||||| |||  || ||||
Sbjct: 2045 TTTT-AATTGTTCAATATATTAGAAAAGAATGAGTATTGTCTAATTGTTAGGTATCAATT 2103

Query: 1054 TGCATTTATATTCACAAGACCAAACTTGCTAATTATGTTTATATGTTCC--ACTGTTG-T 1110
            |  |||||| |||| | || ||| |||| ||| | |||| | || ||    | | ||| |
Sbjct: 2104 TCTATTTATGTTCATAGGATCAA-CTTGTTAACTGTGTTCAGATTTTTAGTATTCTTGGT 2162

Query: 1111 AATTTTTTGCCTGCTTCATCCATTCACTCTAAATAGTTCCAATTAGTACAAACT---AAT 1167
            ||||||||| |||| | ||| ||    |||| |||        || || || ||   |||
Sbjct: 2163 AATTTTTTGTCTGCCTGATCTATCTGTTCTAGATA--------TAATAAAATCTCCCAAT 2214

Query: 1168 ATGCTTCTGCTTTTTATCCATTGCTTCCTGTAATC-------TTTGCTTTAAGTATTTCA 1220
            || | | || |||| |||||   || | | ||||        |||||||||  ||| |
Sbjct: 2215 ATTCCTATGGTTTT-ATCCACCACTCCTTATAATTAATAGTTTTTGCTTTATATATATAT 2273

Query: 1221 AAGTCA---C-AT-TGTTAGGTGTGAACAAGTCCAGAGCC----CT-TTTGTTAAATGAC 1270
            |  |||   | || | ||  | ||| | ||  | || |||    || |||||| |   ||
Sbjct: 2274 ATTTCAAGGCTATGTTTTTAG-GTGCA-AA--CAAG-GCCATGACTCTTTGTTGAGAAAC 2328

Query: 1271 ACCAGGGA-GCAGATCCTTTTACACTAGTAATTATTGACATATTTATCTAATATGCCTCT 1329
            || ||||| |   |||   ||||| |  |||||||||  ||||| | |||||||| | |
Sbjct: 2329 AC-AGGGAAGTTAATCTGCTTACATTTTTAATTATTGATGTATTTGTTTAATATGCTTTT 2387

Query: 1330 TCTTTTTCATTTTCTCCTTCCCAGCTTTTTGCT 1362
              |  || |||||||||| || || || | ||
Sbjct: 2388 GGTCATTTATTTTCTCCTTTCCTGCATTCTACT 2420
```

7.6. Abkürzungsverzeichnis

AF	Allelfrequenz
AG	Androgenetisch
AS	Angelman-Syndrom
BWS	Beckwith-Wiedemann-Syndrom
cen	Central
C_T-Wert	Threshold cycle-Wert
dist	Distal
DMR	Differentiell metylierte Region
Dnmt(1)	DNA Methyltransferase(1)
E(11.5)	Embryonic day (11.5)
EICO	Expression-based Imprint Candidate Organizer
EST	Expressed sequence tag
H(3)	Histon(3)
HSA(19q)	Humanes Chromosom (19q)
IC(1)	Imprinting center(1)
IGF2	Insulin-like Growth Factor Type 2
IGF2R	Insulin-like Growth Factor Type 2-Rezeptor
K2P	Two pore potassium channel
LINE	Long interspersed nuclear elements
MER-Element	Medium reiteration frequency-Element
MIR-Element	Mammalian interspersed repetitive-Element
MMU(7prox)	proximales Maus-Chromosom (7)
ORF	Offener Leserahmen
PG	Parthenogenetisch
PP_i	Anorganisches Pyrophosphat
prox	Proximal
PWS	Prader-Willi-Syndrom
QUASEP	Quantification of allele-specific expression by Pyrosequencing
RACE	Rapid amplification of cDNA ends
RFLP	Restriktionsfragmentlängenpolymorphismus
SAGE	Serial analysis of gene expression
SD	Standardabweichung
SINE	Short interspersed nuclear elements
tSNP	Transkribierter Einzelnukleotidpolymorphismus
UPD	Uniparentale Disomie
ZNS	Zentrales Nervensystem

7.7. Abbildungsverzeichnis

Abbildung 1. Modifikationen an den Seitenketten der vier Kernhistone. 8
Abbildung 2. Die genomische Prägung von *Igf2* und *H19* auf dem distalen Maus-Chromosom 7 (MMU7dist) 10
Abbildung 3. Der Lebenszyklus der Imprints 12
Abbildung 4. Schematische Darstellung des *Gnas*-Locus auf dem distalen Maus-Chromosom 2 (MMU2dist) 13
Abbildung 5. Die Entstehung der genomischen Prägung während der Evolution der Arten 19
Abbildung 6. Vergleich der Expressionsniveaus von androgenetischen und parthenogenetischen Mausembryonen dargestellt als Histogramm 24
Abbildung 7. Schematische Darstellung des Pyrosequencing™-Prinzips 32
Abbildung 8. Pyrosequencing™-Pyrogramme 32
Abbildung 9. Eichgerade zur Bestimmung der tatsächlichen Allelfrequenz (AF) 33
Abbildung 10. Beispiel einer verwendeten Eichgerade zur Bestimmung der tatsächlichen Allelfrequenz (AF) 34
Abbildung 11. QUASEP-Ergebnisse, die die Eignung für die Imprinting-Analyse verdeutlichen 40
Abbildung 12. Gewebespezifische Prägung von *Slc38a4* 40
Abbildung 13. Identifizierung von vier geprägten Transkripten. 41
Abbildung 14. Identifizierung von vier stammspezifisch exprimierten Transkripten 42
Abbildung 15. Genomische Karte vom proximalen Maus-Chromosom 7 mit den bekannten geprägten Genen 46
Abbildung 16. Genomische Karte vom distalen Maus-Chromosom 7 mit AK039482 und vier benachbarten geprägten Genen 47
Abbildung 17. Genomische Karte vom distalen Maus-Chromosom 12 mit AK018142 sowie den bekannten geprägten Genen *Dlk1*, *Gtl2* und *Rtl1* 48
Abbildung 18. Genomische Karte vom proximalen Maus-Chromosom 17 mit AK039514 als Teil von *Air* und dem bekannten geprägten Gen *Igf2r* 48
Abbildung 19. Expressionsvergleich von AK043056 (links) und *Usp29* (rechts) 49
Abbildung 20. Allelische Expressionsanalyse für das humane *PEG14* in Gehirn und Testis 51
Abbildung 21. Genomische Karten von MMU7prox (oben) und HSA19q13.4 (unten) mit *Peg14* und den benachbarten Genen 52
Abbildung 22. Biallelische Expression von *USP29*, *ZIM3* und *ZNF264* sowie präferentiell allelische Expression von *PEG3* auf dem humanen Chromosom 19q13.4 54
Abbildung 23. Allelische Expressionsanalyse für das murine *Kcnk9*-Gen 56
Abbildung 24. Allelische Expressionsanalyse für das humane *KCNK9*-Gen im fetalen Gehirn 57

7.8. Tabellenverzeichnis

Tabelle 1. Geprägte Gene und die Phänotypen von Knockout-Mäusen. 15
Tabelle 2. Humane Imprinting-Erkrankungen (Morison *et al.*, 2005]. 17
Tabelle 3. Primer zur Amplifikation der Sonden für die Northern-Blot-Hybridisierung. 29
Tabelle 4. Primer für die RACE-PCR. 29
Tabelle 5. Ergebnisse der allelischen Expressionsanalyse der untersuchten Imprinting-Kandidatengene. 43
Tabelle 6. Geprägte Expression in verschiedenen adulten Geweben von F1-Maushybriden. 45
Tabelle 7. Position und Größe der Exons und Intron von *Usp29* 47
Tabelle 8. C_T-Werte der Untersuchung mit TaqMan® Gene Expression 50
Tabelle 9. Art und Zusammensetzung repetitiver Elemente in *Peg14* sowie dem humanen Ortholog. 53
Tabelle 10. Allelische Expressionsanalyse der untersuchten Imprinting-Kandidatengene (Luedi-Studie) 55
Tabelle 11. CpG-Inseln in der flankierenden Region des murinen und humanen *Kcnk9*-Gens. 58
Tabelle 12. Allelische Expression und Funktion der geprägten Gene auf MMU7prox / HSA19q13.4. 66
Tabelle 13. Sequenzierungsprimer zur SNP-Identifizierung. 89
Tabelle 14. Primer-Sets zur QUASEP-Analyse 94
Tabelle 15. Transkribierte SNPs der Nikaido-Kandidatengene 100
Tabelle 16. Identifizierte tSNPs der Luedi-Kandidatengene. 105