

Aus der Abteilung
für zelluläre Immunologie
am Heinrich-Pette-Institut

Kommissarischer Leiter: Prof. Dr. K. Harbers

Untersuchungen zum Einfluss von
Aktivitätsinduzierter Cytidin Deaminase auf
die Transkription in antikörperproduzierenden
B-Zellen mittels Mikroarrays

Dissertation zur Erlangung des Grades eines Doktors der Medizin

dem Fachbereich Medizin der Universität Hamburg vorgelegt
von Nina Gartmann aus Hamburg

Hamburg 2003

Inhaltsverzeichnis

| | | |
|----------|-----------------------------------------------------------------------------------------------|-----------|
| 1 | Arbeitshypothese und Fragestellung | 4 |
| 2 | Einleitung | 5 |
| 2.1 | Die Antikörpervielfalt beruht auf somatischen Rekombinationsmechanismen | 5 |
| 2.1.1 | V(D)J-Rekombination | 7 |
| 2.1.2 | Genkonversion | 8 |
| 2.1.3 | Somatische Hypermutation | 9 |
| 2.1.4 | „Class-switch-recombination“ | 10 |
| 2.1.5 | DNA-Doppelstrangbrüche und Reparatur während der somatischen Rekombination | 11 |
| 2.1.6 | Immundefekte und Tumorentwicklung als Folge von gestörten Rekombinationsmechanismen | 12 |
| 2.2 | Die DT40-Zelllinie als Modell für Rekombination im Immunsystem | 13 |
| 2.2.1 | Abkömmlinge von DT40 | 13 |
| 2.3 | Aktivitätsinduzierte Cytidin Deaminase | 14 |
| 2.3.1 | AID und die autosomal rezessive Form des Hyper-IGM-Syndroms | 16 |
| 2.3.2 | AID-Knockout in DT40 Cre1 | 16 |
| 2.4 | Mikroarrays zur Transkriptionsanalyse | 17 |
| 2.4.1 | Bedeutung von Transkriptionsprofilen für die medizinische Diagnostik und Therapie | 18 |
| 3 | Material und Methoden | 19 |
| 3.1 | Material | 19 |
| 3.1.1 | Filter | 19 |
| 3.1.2 | Software | 19 |
| 3.1.3 | Radioisotope | 19 |
| 3.1.4 | Agarosegelelektrophorese | 19 |
| 3.1.5 | Kits | 19 |
| 3.1.6 | „Stripping-Solution“ | 20 |
| 3.1.7 | „Amasino-Puffer“/ Hybridisierungspuffer für die Oligo-Hybridisierung | 20 |
| 3.1.8 | „Oligowash“ für die Oligo-Hybridisierung | 20 |
| 3.1.9 | 50x Denhardt's | 20 |
| 3.1.10 | Denhardt's Hybmix/ Hybridisierungslösung für die cDNA . | 20 |
| 3.1.11 | Waschpuffer I-III | 20 |
| 3.1.12 | Weitere Lösungen und Puffer | 21 |
| 3.1.13 | Weitere Reagenzien | 21 |
| 3.1.14 | Kulturmedium für DT40 Zellen | 21 |
| 3.1.15 | Gefriermedium für DT40 Zellen | 21 |

| | | |
|----------|---------------------------------------------------------------------------------------------------------------|-----------|
| 3.1.16 | Oligonukleotid-Primer für semiquantitative RT-PCR | 22 |
| 3.1.17 | Geräte | 23 |
| 3.2 | Methoden | 23 |
| 3.2.1 | cDNA-Bibliotheken und ESTs | 23 |
| 3.2.2 | Die cDNA-Bibliothek und die Datenbank bursaler B-Zellen dkfz 426 | 23 |
| 3.2.3 | Anfertigung von Mikroarrays | 24 |
| 3.2.4 | Herstellung der RZPD Filter | 24 |
| 3.2.5 | DT40 Zellen in Zellkultur | 26 |
| 3.2.6 | Zählung der Zellen | 26 |
| 3.2.7 | Aufbewahrung von DT 40 Zellen | 26 |
| 3.2.8 | Auftauen von Zellen | 26 |
| 3.2.9 | RNA-Isolation und mRNA Gewinnung | 27 |
| 3.2.10 | cDNA-Synthese | 30 |
| 3.2.11 | „Stripping“ der RZPD Filter | 32 |
| 3.2.12 | Hybridisierung der RZPD Filter mit Oligonukleotiden | 32 |
| 3.2.13 | Hybridisierung der RZPD Filter mit den cDNA Sonden | 33 |
| 3.2.14 | Auswertung der Filter mit dem „Advanced Image Data Analysingprogram (AIDA)“ | 34 |
| 3.2.15 | „Noise“ | 34 |
| 3.2.16 | „AIDA Array Metrix“ | 35 |
| 3.2.17 | „AIDA Array Compare“ | 36 |
| 3.2.18 | Statistische Auswertung wiederholter Hybridisierungen | 36 |
| 3.2.19 | t-Test-Statistik | 37 |
| 3.2.20 | SAM, „significance analysis of microarrays“ | 37 |
| 3.2.21 | Bestimmung der Sequenzen der differentiell exprimierten ESTs und Annotation durch Datenbanksuche | 38 |
| 3.2.22 | Semiquantitative RT-PCR zur Bestätigung der differenti- ellen Expression | 38 |
| 4 | Ergebnisse | 40 |
| 4.1 | Experimente | 40 |
| 4.1.1 | RNA-Isolation von AID-positiven und AID-negativen DT40- Zellen | 40 |
| 4.1.2 | Menge und Reinheit der Gesamt-RNA und mRNA | 41 |
| 4.1.3 | cDNA Synthese und Qualitätskontrolle | 41 |
| 4.1.4 | Die spezifische Aktivität der Hybridisierungssonden | 41 |
| 4.1.5 | Hybridisierung | 41 |
| 4.1.6 | Hybridisierung der Filter mit Oligonukleotiden | 42 |
| 4.1.7 | Hybridisierung der Filter mit den cDNA-Sonden | 43 |
| 4.2 | Auswertung mit dem AIDA Programm | 43 |
| 4.2.1 | Auswertung mit AIDA Array Metrix | 43 |

| | | |
|----------|------------------------------------------------------------------------------------------|-----------|
| 4.2.2 | Vergleich der AID-positiven und AID-negativen Filter mit AIDA Array Compare | 47 |
| 4.3 | Statistik und Datenbanksuche | 58 |
| 4.4 | Analyse der ESTs mit Hilfe der Datenbank „dkfz 426“ | 66 |
| 4.5 | Semiquantitative RT-PCR | 68 |
| 5 | Diskussion | 72 |
| 6 | Zusammenfassung | 75 |
| 7 | Abkürzungen | 77 |
| 8 | Danksagung | 88 |
| 9 | Lebenslauf | 89 |

1 Arbeitshypothese und Fragestellung

AID, Aktivitätsinduzierte Cytidin Deaminase, ist ein Schlüsselenzym für die Rekombination im Immunsystem, da es für Genkonversion, somatische Hypermutation und „Class-switch-recombination“ erforderlich ist. In welcher Weise AID wirkt und welche Faktoren noch eine Rolle spielen könnten, war zu Beginn dieser Arbeit nicht bekannt. Da Genkonversion kontinuierlich in der Zelllinie DT40 stattfindet, hat man mit ihr ein Modellsystem, an dem sich dieser Rekombinationsmechanismus studieren lässt. Das Gen für AID wurde durch homologe Rekombination im DT40-Abkömmling DT40 Cre1 ausgeschaltet. Im Rahmen dieser Arbeit soll mittels Mikroarrays die Transkription in AID-positiven Zellen (DT40 Cre1) und in AID-negativen Zellen (DT40 Cre1 AID^{-/-}) verglichen werden, um Gene zu finden, die möglicherweise differentiell exprimiert sind. Diese differentiell exprimierten Gene können wichtige Hinweise auf den zur Zeit noch ungeklärten Wirkmechanismus von AID liefern. Folgende Szenarien zur Wirkungsweise von AID sind denkbar:

1. AID könnte ähnlich wie sein sequenzhomologes APOBEC-1 die RNA editieren. Daraus könnte eine Bildung alternativer mRNAs resultieren, die wiederum unterschiedliche Genprodukte nach sich ziehen. Die Darstellung von Transkriptionsunterschieden kann hier schwierig sein, wenn es sich um eine oder wenige veränderte mRNAs handelt.
2. AID könnte als globaler Transkriptionsfaktor die Transkription von Genen induzieren. Dann würden Transkriptionsunterschiede zwischen Wildtyp- und Knockoutzelle auf den Arrays sichtbar werden. So sollten die von AID regulierten Gene detektiert werden können.
3. AID könnte die DNA editieren, wodurch Mutationen induziert werden. Dann wäre primär keine Veränderung des Transkriptionsprofils zu erwarten. Sekundär könnte aber die Expression einzelner Gene beeinflusst werden.

Falls es sich mit AID um einen Transkriptionsfaktor handelt, der die Transkription von Genen induziert, die an Genkonversion beteiligt sind, sollte sich dies in einem veränderten Transkriptionsprofil zeigen. Ein wichtiger Teil der Arbeit ist die Etablierung der Mikroarray-Technik mit den hier verwendeten Filtern, da es sich dabei um eine sehr störanfällige Technik handelt. Es gibt viele Variablen, die sich unspezifisch auf das Ergebnis auswirken (siehe „Noise“ auf S. 34). Um falsch signifikante Expressionsunterschiede herauszufiltern, werden zwei statistische Methoden angewandt. Die differentielle Expression der durch die Arrays gefundenen Kandidaten wird anschließend durch semiquantitative RT-PCR überprüft.

2 Einleitung

In diesem Abschnitt werden zunächst die somatischen Rekombinations- und Mutationsmechanismen beschrieben, mit denen die große Variabilität der Antikörpermoleküle erreicht wird (2.1). Im Weiteren wird die Modellzelle DT40 eingeführt, in der einer der B-zellspezifischen Rekombinationsmechanismen stattfindet. Durch gezielte Genzerstörung von AID, Aktivitätsinduzierte Cytidin Deaminase, kann dieser Mechanismus ausgeschaltet werden (2.2). Anschliessend wird die Funktion von AID diskutiert und das Ziel dieser Arbeit definiert (2.3 und Arbeitshypothese 1). Unter 2.4 wird die Mikroarray-Methode eingeführt, mit der die Transkription analysiert wurde. Diese Methode könnte in Zukunft in der klinischen Diagnostik und Therapie eine wichtige Rolle spielen, da sie zu einer molekularen Charakterisierung von Zellen und Geweben beiträgt. Es zeigt sich, dass die Methode empfindlich unspezifischen Einflüssen gegenüber ist. Mit Mikroarrays erzeugte Daten müssen durch andere Methoden bestätigt werden, da selbst durch sehr genaues experimentelles Arbeiten sich nicht alle Artefakte eliminieren lassen. Daher sind statistische Methoden wichtig, um falsch signifikante Daten auszuschließen.

2.1 Die Antikörpervielfalt beruht auf somatischen Rekombinationsmechanismen

Durch Antigenrezeptoren auf ihrer Oberfläche erkennen Lymphozyten Antigene in ihrer Umgebung. Jeder von einem B-Lymphozyten sezernierte Antikörper hat eine einzigartige Spezifität, die durch eine einzigartige Aminosäurezusammensetzung der Antigenbindungsstelle zustande kommt. Die Antigenbindungsstelle wird von den variablen Regionen der Leicht- und Schwereketten kodiert. Jeder variablen Region wird eine konstante Region angefügt, die für die Effektorfunktion des Antikörpermoleküls entscheidend ist, je nachdem an welchem Ort oder zu welchem Zeitpunkt die Antigen-Antikörperreaktion im Körper stattfinden soll [40].

Man schätzt die Zahl der möglichen Antikörpersequenzvarianten, die in Wirbeltieren erzeugt werden können, auf 10^{11} bis 10^{15} . Die hohe Variabilität wird durch Rekombinationsmechanismen erreicht, die die variable Region des Antikörpers verändern [80]. Zunächst werden die V-, (D-) und J-Gensegmente der variablen V-Region durch V(D)J-Rekombination rekombiniert (siehe 2.1.1). Die Immunglobulinleichtkettengene werden dabei durch ein V- („variable“) und ein J- („joining“)-Segment gebildet. Das Schwereketten gen enthält ein zusätzliches D- („diversity“)-Segment. Dabei hat dieser primäre Schritt je nach Spezies eine unterschiedliche Bedeutung für die Variabilitätsbildung, da er nicht in jedem Organismus entscheidend zur Repertoirebildung beiträgt (siehe Tab.1). Fast alle Wirbeltiere haben Mechanismen entwickelt, die Variabilität der V-Region im Anschluss an die V(D)J-Rekombination durch verschiedene Mechanismen zu er-

| Spezies | Primärer Mechanismus / Organ | Sekundärer Mechanismus in den Keimzentren der sekundären lymphatischen Organe |
|---------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Mensch | V(D)J-Rekombination / Knochenmark | Somatische Hypermutation |
| Maus | V(D)J-Rekombination / Knochenmark | Somatische Hypermutation |
| Kuh | V(D)J-Rekombination, somatische Hypermutation / Ileale Peyer Plaques | Somatische Hypermutation, Genkonversion? |
| Schwein | V(D)J-Rekombination, Genkonversion / ? | Somatische Hypermutation, Genkonversion? |
| Schaf | V(D)J-Rekombination, Somatische Hypermutation / Ileale Peyer Plaques | Somatische Hypermutation, Genkonversion? |
| Hase | V(D)J-Rekombination, Genkonversion, Somatische Hypermutation / Appendix | Somatische Hypermutation, Genkonversion |
| Huhn | V(D)J-Rekombination, Genkonversion, Somatische Hypermutation / Bursa Fabricii | Somatische Hypermutation, Genkonversion |

Tabelle 1: Diversifikationsmechanismen verschiedener Spezies während der primären Repertoirebildung und der sekundären antigenabhängigen Immunantwort. Während V(D)J-Rekombination bei Menschen und Mäusen deutlich zur primären Repertoirebildung beiträgt, ist dieser Mechanismus bei anderen Spezies, wie Kuh, Schwein, Schaf, Hase und Huhn von untergeordneter Bedeutung. Hier sind es Mechanismen wie Genkonversion und somatische Hypermutation, die entscheidend zur primären Repertoirebildung beitragen.

höhen, sowohl in den primären lymphatischen Organen antigenunabhängig, als auch später, antigenabhängig, bei der Affinitätsreifung in den Keimzentren der sekundären lymphatischen Organe [8]. Die beiden grundlegenden Mechanismen für die somatische Mutation der Immunglobulin-Gene sind Genkonversion und somatische Hypermutation. Bisher scheint es keine klare Verbindung zwischen der Position eines Organismus in der Phylogese und der Wahl des jeweiligen Mutationsmechanismus zu geben. Genkonversion ist ein Prozess, der meist im späten fetalen und frühen postnatalen Stadium antigenunabhängig auftritt (siehe 2.1.2). Somatische Hypermutation (2.1.3) kommt gelegentlich perinatal bei der Bildung des primären Immunglobulinrepertoires antigenunabhängig vor und ist später der Hauptmechanismus bei der antigenabhängigen Affinitätsreifung in den Keimzentren der sekundären lymphatischen Organe [74]. Die verschiedenen Antikörperklassen entstehen durch „Class-switch-recombination“, dem Mechanismus,

mit dem die für die konstanten Regionen kodierenden Gene rekombiniert werden. Durch die so entstandenen unterschiedlichen Effektormoleküle werden Art und Ort der Immunantwort bestimmt (siehe 2.1.4).

2.1.1 V(D)J-Rekombination

Während der Differenzierung in den primären lymphatischen Organen rekombinieren B-Zellvorläufer zunächst antigenunabhängig ihre V-, (D-) und J-Gene [80]. Je nach Vorhandensein der Gensegmente findet im Schwerkettenlokus eine VDJ- und im Leichtkettenlokus eine VJ-Rekombination statt. Sobald ein funktionelles Gen zusammengebaut ist, wird jede weitere Rekombination des gleichen Gens gestoppt. Nur B-Zellen, die funktionelle Leicht- und Schwerkettengene exprimieren, werden als naive IgM- oder IgD-positive B-Zellen in die Peripherie entlassen.

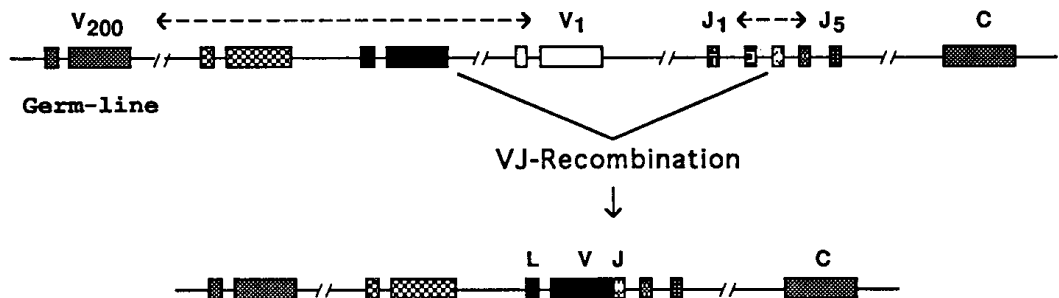


Abbildung 1: V(D)J-Rekombination in der Maus. Der Maus- κ -Leichtkettenlokus enthält ca. 200 V-Segmente, die sich über eine Region von 3000 kbp erstrecken, sowie 5 J-Segmente, die 20 kbp 3' des letzten V-Segmentes beginnen und ein C-Segment. VJ-Rekombination verbindet hier ein V-Gensegment mit einem der J-Segmente [80], [10].

Die Rekombination findet zwischen sogenannten „heptamer-spacer-nonamer“-Signalerkennungssequenzen, die die V-, D- und J-Gensegmente einrahmen, statt [80], [29]. Dies sind nicht kodierende, zum Teil konservierte Rahmensequenzen, die durch ihre Struktur dafür sorgen, dass die kodierenden Segmente in richtiger Weise rekombiniert werden können. Die V(D)J-Rekombinase, ein Proteinkomplex, besteht aus den lymphoidspezifischen Proteinen RAG1 und RAG2 (RAG=„recombination activating gene“) und weiteren ubiquitären DNA-Reparaturproteinen [69], [54], [65], [26]. Durch die RAG-Proteine werden zunächst die Erkennungssequenzen gebunden. Anschließend kommt es zu Doppelstrangbrüchen und schließlich zur Ligation der Gensegmente, nach dem durch die TdT (terminale Desoxytransferase) Nukleotide eingefügt wurden. Beteiligte Proteine des V(D)J-Rekombinase-Komplexes sind ferner Ku70, Ku80, XRCC4, DNA-Proteinkinase (DNA-PK) und DNA-Ligase IV. Eine fehlende Aktivität der katalytischen Untereinheit der Proteinkinase (PK_{cs}), kann zu einem schweren kombi-

nierten Immundefekt in Mäusen führen, dem SCID-Phänotyp (SCID= „severe combined immunodeficiency“) [12]. Die Mutation der DNA-Proteinkinase zieht ferner einen generellen Defekt in der Doppelstrangbruch-Reparatur nach sich (siehe dazu auch 2.1.4 und 2.1.5). Damit konnte eine Verbindung zwischen V(D)J-Rekombination und Doppelstrangbruch-Reparatur festgestellt werden [28], [36].

2.1.2 Genkonversion

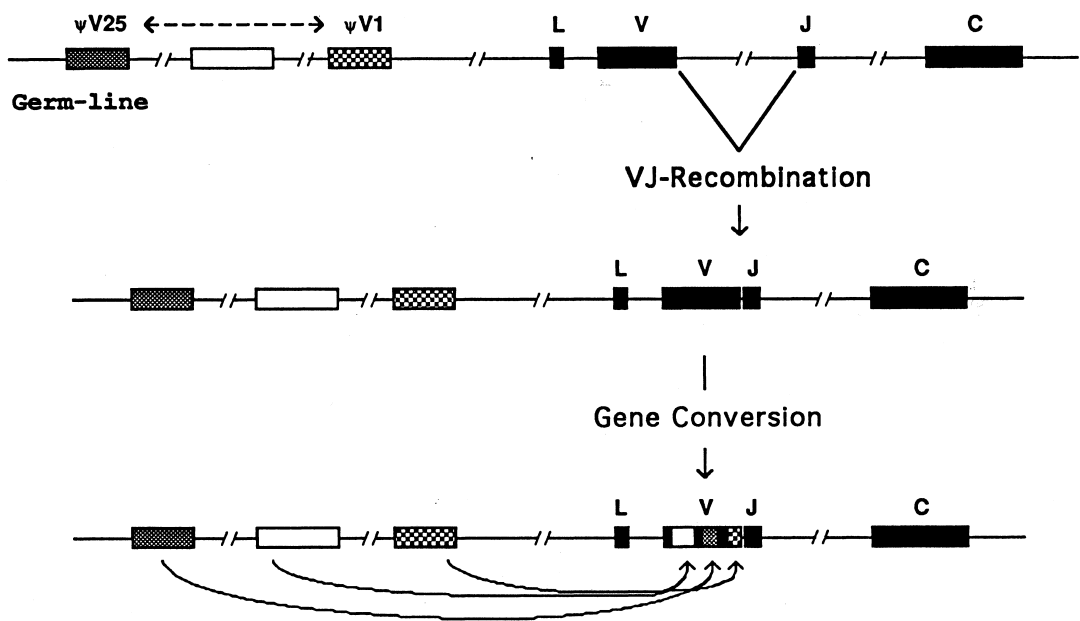


Abbildung 2: Genkonversion. Erzeugung des Leichtketten-Immunglobulinrepertoires im Huhn. Der Locus erstreckt sich über nur 30 kbp und enthält ein einziges V-, J- und C-Gen unterhalb eines Pseudogenpools mit 25 Pseudogenen ($\Psi 1 - \Psi 25$). Nach dem die V- und J-Segmente durch VJ-Rekombination verbunden sind, werden durch Genkonversion in der Bursa Fabricii V-Sequenzen durch Pseudogensequenzen substituiert [10].

Genkonversion ist ein weiterer Mechanismus der Rekombination von V-Genen. Bei einigen Spezies wie z.B. Hühnern, trägt die V(D)J-Rekombination nicht entscheidend zur Repertoireentwicklung bei. Durch Genkonversion wird jedoch trotzdem der Aufbau eines primären Repertoires von hoher Variabilität erreicht [84], [46], [76]. Oberhalb des V-Segmentes liegen über eine Region von 22 kbp verteilt 25 Pseudo-V-Gene, die homolog zu den V-Genen sind, aber keine eigenen Transkriptionsregulations- und Signalerkennungssequenzen tragen. Während der B-Zellproliferation im primären Immunorgan der Vögel, der Bursa Fabricii, kommt es durch homologe Rekombination zum Einbau von Pseudogensequenzen in die V-Segmente [62], [78]. Die Anzahl der Genkonversionsereignisse steigt mit der Zeit, in der B-Zellen sich in der Bursa aufhalten. Im Durchschnitt kommt es

alle 10 bis 15 Teilungen zu einem Konversionsereignis [84]. Dabei werden Pseudogene bevorzugt, die nah am V-Gen liegen, in Antisense-Orientierung vorliegen und homolog sind [63], [45]. Das initiiierende Ereignis für diesen Rekombinationsmechanismus ist noch nicht hinreichend verstanden. Es ist aber bekannt, dass das Enzym AID („Aktivitätsinduzierte Cytidineaminase“) für Genkonversion essentiell ist [5] (siehe auch Abschnitt 2.3).

2.1.3 Somatische Hypermutation

Während einer Antigenexposition wandern aktivierte B-Zellen in die B-Zell-Follikel der sekundären lymphatischen Organe. Dort expandieren sie und bilden oligoklonale Keimzentren, in denen sie ihre variablen V-Gene durch somatische Hypermutation verändern. Wie in Abschnitt 2.1 beschrieben, gibt es auch primäre antigenunabhängige somatische Hypermutationsereignisse, die in speziellen primären Immunorganen stattfinden, deren Mechanismus jedoch aller Wahrscheinlichkeit nach dem antigenabhängigen entspricht.

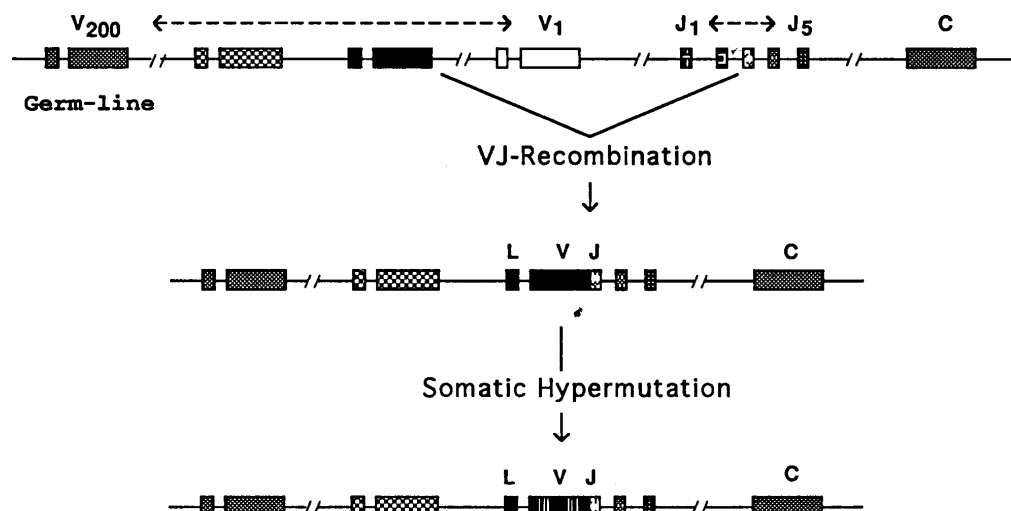


Abbildung 3: Somatische Hypermutation in der Maus. Nach V(D)J-Rekombination werden antigenabhängige Punktmutationen in das V-Segment eingefügt [83], [10].

Punktmutationen werden vorwiegend in den V(D)J-Regionen der IgH- und IgL-Gene induziert. Ein kleiner Anteil an B-Zellen, die Immunglobuline mit erhöhter Affinität zum Antigen produzieren, differenzieren zu antikörperproduzierenden Plasma- oder Gedächtniszellen aus. Diese Zellen bilden das sekundäre Immunglobulinrepertoire, das eine Affinitätsreifung der antigenspezifischen Immunglobuline durchgemacht hat [60], [39]. 50-60 % aller Mutationen befinden sich in RGYW-Motiven (R entspricht A oder G (Purin), Y entspricht C oder T (Pyrimidin) und W entspricht A oder T), so genannten „hotspots“. Die Mutationsrate wird auf $10^{-3} - 10^{-4}$ bp pro Generation geschätzt und ist somit 6

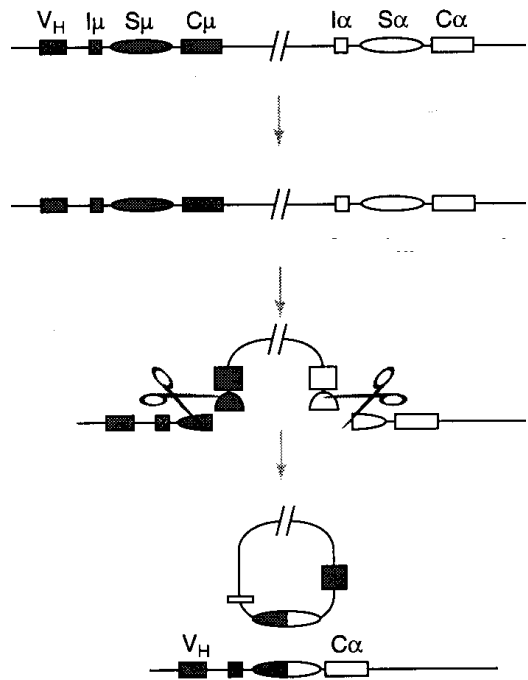
mal höher als die Rate spontaner Mutationen. Dabei werden Transitionen bevorzugt. Die Mutationen erstrecken sich über eine ungefähr 2 kb große Region, die HYM-Domäne („hypermutating region domain“), unterhalb des Ig-Promotors, wobei die Frequenz der Mutationen von 5' nach 3' um das erste Exon zunächst stark zunimmt und über das rearrangierte V(D)J-Exon in Richtung J-C-Intron wieder abnimmt [51]. Neben einem geringen Anteil an Deletionen, Insertionen und Duplikationen, handelt es sich meist um Punktmutationen (> 90%), wobei die Mutationsrate an die Transkriptionsaktivität geknüpft ist [30], [14], [27]. Als Reaktionszwischenprodukte werden Doppelstrangbrüche detektiert, die möglicherweise durch eine Nuklease induziert werden und im Anschluss durch eine fehleranfällige Polymerase repariert werden [67], [56], [14], [15]. Die Mutation wird durch die „intron enhancer/matrix attachment region“ kontrolliert [9]. „Cis-enhancer“-Sequenzen werden durch „trans-acting“-Faktoren aktiviert. Damit können in diesem Falle die Mutationen verstärkt werden. Es zeigte sich, dass das AID-Genprodukt in die Kontrolle von somatischer Hypermutation eingebunden ist [49]. Der genauen Funktion von AID konnte aber erst in jüngster Zeit auf die Spur gekommen werden (siehe Abschnitt 2.3 und 5). Unter bestimmten Bedingungen scheint somatische Hypermutation nicht auf die V(D)J-Region beschränkt zu sein, da dieser Mechanismus in Fibroblasten induziert werden kann und dort eine Mutationsfrequenz von 10^{-3} bp pro Generation hervorruft [55].

2.1.4 „Class-switch-recombination“

Während der antigenabhängigen Immunantwort in den Keimzentren der sekundären lymphatischen Organe werden die funktionell rearrangierten V(D)J-Regionen zunächst mit der C_μ -Region der Schwereketten gene verknüpft. Im weiteren Verlauf der Immunantwort kann durch externe Signale, wie von T-Zellen produzierte Cytokine oder durch Pathogene freigesetzte Mitogene, die Verknüpfung der selben V(D)J-Region mit anderen C-Genen (α , γ , ϵ) induziert werden. Die daraus entstehenden sekundären Isotypen (IgA, IgG, IgE) haben einen „Class-switch“ durchgemacht, der bei gleicher Antigenpezifität zu einer veränderten Form der Immunantwort führt [40].

Der Ig-Schwerkettenloкус umfasst 3' der V-Region ein Cluster aus C-Exons, die durch in den Introns gelegenen, aus repetitiven Tandemsequenzen bestehenden S-Regionen unterbrochen werden [50]. Diese repetitiven Sequenzen können ein Aneinanderlegen homologer Abschnitte erleichtern, ein initiales Ereignis homologer DNA-Rekombination. Durch die Rekombination von S-Regionen und das gleichzeitige Herausschneiden der dazwischen liegenden C-Regionen, kommt die V-Region in die Nachbarschaft einer neuen C-Region. Da die Rekombination zwischen Intron-Sequenzen stattfindet, ist gesichert, dass keine Leserahmenverschiebungen auftreten, die zu funktionslosen Proteinen führen würden [40]. Doppelstrangbrüche treten als Reaktionsintermediate auf, die wahrscheinlich durch NHEJ („non-homologous end-joining“, siehe 2.1.5.) repariert werden, wobei die

Abbildung 4: Mögliche „Class-switch-recombination“ von IgM zu IgA: Mit CD 40 Ligand oder Zytokinen stimulierte B-Zellen induzieren die Transkription ausgehend von einem Promotor 5' eines I α Exons. Zytokinstimulation induziert die Neusynthese von AID („Activation induced deaminase“) in B-Zellen. Ein unbekanntes DNA-schneidendes-Enzym schneidet durch Transkription zugängliche S-Regionen. Die zwischen den geschnittenen S-Regionen liegende DNA wird in Form eines zirkulären DNA-Stückes entfernt. Die abgeschnittenen Enden werden wahrscheinlich durch NHEJ („non-homologous end-joining“) repariert, wobei verschiedene Proteine und Polypeptide involviert sind: Die drei Untereinheiten der DNA-Proteinkinase „DNA-PK“ (DNA-PK_{cs}, Ku70 und Ku80), XRCC4 und DNA-Ligase IV [50].



DNA-Proteinkinase „DNA-PK“ mit ihren Polypeptiduntereinheiten Ku70 und Ku80 und der katalytischen Untereinheit „PK_{cs}“ wirksam ist [72] (siehe auch V(D)J-Rekombination S. 7).

2.1.5 DNA-Doppelstrangbrüche und Reparatur während der somatischen Rekombination

Bei Rekombinationsmechanismen kommt es oft zur Bildung von Doppelstrangbrüchen, die repariert werden müssen. Es liegt nahe, dass es gemeinsame Faktoren gibt, die sowohl essentiell für Reparatur-, als auch für Rekombinationsmechanismen sind. In Wirbeltieren gibt es zwei dominante Doppelstrangbruchreparaturmechanismen, die in unterschiedlichen Phasen des Zellzyklus wirken: Homologe Rekombination in der späten S- bis frühen G₂-Phase und NHEJ (non-homologous end-joining) in der G₁- bis frühen S-Phase [75]. Homologe Rekombination benötigt homologe DNA-Sequenzen, die auf dem homologen Chromosom, dem Schwesterchromatid oder, bei intrachromosomaler Genkonversion, auf demselben Chromosom liegen können. Während NHEJ werden die zwei Enden des unterbrochenen Doppelstranges repariert, ohne daß eine homologe Sequenz benötigt wird. NHEJ ist der häufigste Reparaturmechanismus in Wirbeltierzellen, mit dem oft vorkommende Doppelstrangbrüche, die z.B. durch Strahlung und reaktive Sauerstoffmetabolite erzeugt werden, repariert werden. Doppelstrangbrüche

bei der V(D)J-Rekombination werden durch NHEJ primär in Zellen, die sich in der G_1 -Phase befinden, repariert. Wahrscheinlich repariert NHEJ auch die Brüche bei „Class-switch-recombination“ [72]. Während somatischer Hypermutation induzierte Doppelstrangbrüche hingegen werden vor allem in Zellen in der G_2 -Phase gefunden und wahrscheinlich durch homologe Rekombination repariert [56]. Bei Genkonversion handelt es sich um eine spezielle Form der homologen Rekombination. Obwohl die Reparaturmechanismen von somatischer Hypermutation und „Class-switch-recombination“ unterschiedlich zu sein scheinen, weisen die Endprodukte in beiden Fällen vor allem Punktmutationen auf. Möglicherweise ist in beiden Fällen eine fehleranfällige Polymerase involviert, die die Punktmutationen erzeugt. [57].

2.1.6 Immundefekte und Tumorentwicklung als Folge von gestörten Rekombinationsmechanismen

Es gibt verschiedene Krankheitsbilder, die auf fehlgesteuerte Mutationsmechanismen zurückgeführt werden.

- Die hohe Mutationsrate während der einzelnen Schritte der Entwicklung und Reifung stellt gleichzeitig eine ständige Gefahr für die B-Zellen dar, maligne zu entarten [73]. Bei der somatischen Hypermutation kommt es mit einer Frequenz von 10^{-3} bis 10^{-4} pro bp pro Generation vor allem zu Punktmutationen, während Deletionen und Insertionen selten vorkommen [51]. In verschiedenen B-Zell-Lymphomen, zu denen Burkitt's Lymphom, DLBCL („diffuse large B-cell lymphoma“, siehe 2.7., S.18), monozytäres B-Zell-Lymphom und Hodgkin's Lymphom gehören, konnten in den V-Regionen Deletionen und Insertionen detektiert werden, die aber auch in B-Zellen ohne maligne Transformation vorhanden sind. Sie werden als „Nebenprodukt“ der somatischen Hypermutation diskutiert [30].
- „Common variable immunodeficiency“ CVID ist das am meisten verbreitete symptomatische primäre Antikörpermangelsyndrom. Der zu Grunde liegende Immundefekt ist noch nicht geklärt, doch konnte eine deutliche Reduktion an somatischer Hypermutation in zirkulierenden B-Gedächtniszellen festgestellt werden [41].
- Hyper-IgM-Syndrom (Siehe 2.3.1).
- „SCID“, schwerer kombinierter Immundefekt, der auf eine fehlende Aktivität der katalytischen Untereinheit der DNA-Proteinkinase zurückgeführt wird (siehe 2.1.1).

2.2 Die DT40-Zelllinie als Modell für Rekombination im Immunsystem

Die DT40-Zelllinie eignet sich besonders gut, den Mechanismus der Genkonversion zu studieren. DT40 ist eine durch das ALV (Avian leucosis Virus) transformierte Hühner-B-Zell-Linie aus der Bursa Fabricii, dem primären Immunorgan der Vögel. Ungefähr 85 % aller ALV-induzierten Tumoren haben das 3' Ende der viralen „Langen terminalen Wiederholung“ (LTR) oberhalb des c-myc Gens integriert, und erhöhen die Transkription dieses Gens damit ca. 100fach im Vergleich zu normalen bursalen Zellen [6], [35]. C-myc induziert die Synthese von Proteinen, die für die Tumorprogression bedeutend sind. Die DT40-Zelllinie kann man als eine im bursalen Stammzellstadium stehengebliebene Zelllinie beschreiben [6], [77], [46]. Damit existiert ein gutes Modellsystem, an dem sich ein B-zellspezifischer Prozess, wie Genkonversion, studieren lässt. Neben der hohen Rate an Genkonversionsaktivität kann man eine hohe Rate an gezielter Integration verzeichnen (mit 1:2 liegt sie bei weitem über jeder Säugetierzelllinie), so dass sich transfizierte Genkonstrukte mit einer hohen Trefferrate in endogene Loci integrieren lassen [18]. Aufgrund der raschen Proliferation können daraus resultierende veränderte Phänotypen schnell detektiert werden. Daher ist DT40 besonders gut geeignet für Knockout-Experimente. DT40 Zellen sind mit einem Durchmesser von ungefähr 10 μm klein und haben eine hohe Kern-Plasma-Relation. Ihre Verdopplungszeit beträgt ca. 12 Stunden. Sie tragen IgM auf ihrer Oberfläche, das mittels FACS-Analyse („Flourescence activated cell sorting“) sichtbar gemacht werden kann. Sie enthalten ein rearrangiertes und ein nicht rearrangiertes Leichtketten-Allel [18].

2.2.1 Abkömmlinge von DT40

Ausgehend von der oben beschriebenen Wildtypzelllinie DT40 wurden verschiedene Abkömmlinge isoliert, die sich im Hinblick auf Genkonversionsaktivität und Oberflächen IgM-Expression voneinander unterscheiden.

- Klon 18 (Cl 18) ist eine Variante von DT40, die aufgrund einer Leserahmenverschiebung im Leichtketten-V-Segment keine Oberflächen-IgM-Moleküle exprimiert. Daher ist sie primär Oberflächen-IgM-negativ. Diese V-Segment-Mutation kann durch Genkonversion repariert werden, so dass IgM-Moleküle wieder auf der Oberfläche erscheinen [17]. Die Genkonversionsrate ist in diesem Zellklon relativ niedrig und die IgM-Expression auf der Zelloberfläche liegt drei Wochen nach Subklonierung im Durchschnitt bei 1% (siehe Abb. 5 auf S. 14).
- DT40 Cre1 ist eine v-myb-Transformante von Cl 18. V-myb ist ein virales Onkogen und ein Transkriptionsfaktor. DT40 Cre1 zeigt eine fünf mal höhere IgM-Reversionsrate als Cl 18 und damit eine höhere Genkonversionsaktivität, vergleichbar mit der bursaler B-Zellen. DT40 Cre 1 enthält ein

durch Tamoxifen induzierbares Mer-Cre-Mer Rekombinasegen, durch das lox-P-flankierte Kassetten ausgeschnitten werden können, um Knockout-Mutanten zu erzeugen [5]. Daher eignet sich dieser Zellklon besonders für Knockout-Experimente (siehe Abb. 5 auf S. 14).

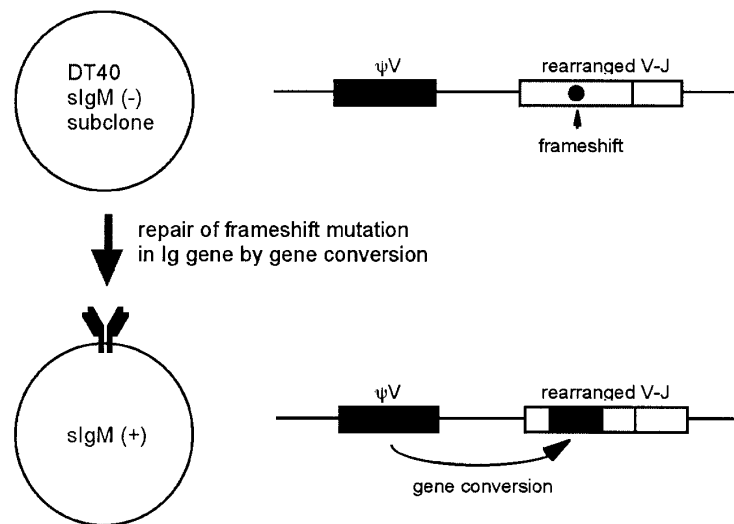


Abbildung 5: Die beiden Oberflächen-IgM-negativen Subklone Cl 18 und DT40 Cre1 können durch Genkonversion die Leserahmenverschiebung reparieren und tragen dann wieder IgM auf der Oberfläche. Dies kann durch FACS-Analyse sichtbar gemacht werden.

2.3 Aktivitätsinduzierte Cytidin Deaminase

Aktivitätsinduzierte Cytidin Deaminase („AID“) ist für die drei B-zellspezifischen Rekombinationsmechanismen somatische Hypermutation, Genkonversion und „Class-switch-recombination“ essentiell [48], [49], [5]. Die Expression von AID ist auf die Keimzentren der lymphatischen Organe beschränkt, also den Ort der Modulation der Immunglobulin-Gene [48]. Wie AID so unterschiedliche Mechanismen kontrolliert, deren Gemeinsamkeit Mutationen sind, ist Gegenstand vieler neuerer Untersuchungen. Die Frage ist, wie genau AID Mutationen induziert. Dabei könnte AID die Synthese eines DNA-schneidenden Enzyms hervorrufen, da DNA-Strangbrüche in allen Mutationsmechanismen eine Rolle spielen. AID könnte auch selbst Mutationen auf der DNA erzeugen, da es eine Deaminase-Aktivität hat, die an der DNA wirken könnte.

AID hat eine Domäne, die eine Sequenzhomologie zum RNA-editierenden APOBEC-1 (Apolipoprotein B editing catalytic subunit 1) aufweist [48]. APOBEC-1 ist die katalytische Untereinheit von APOBEC, einem RNA-editierenden Komplex, der zusammen mit einem akzessorischen Molekül Cytosin zu Uracil auf dem

apoB-Transkript deaminiert und dadurch neben dem Apolipoprotein 100, ein alternatives Apolipoprotein 48 erzeugt. Auch AID hat eine Deaminaseaktivität, für die allerdings das Substrat und mögliche akzessorische Moleküle bisher unbekannt sind. Die Sequenzhomologie, und die Tatsache, dass AID CTP zu UTP an freien Nukleotidtriphosphaten deaminiert unterstützen die Hypothese, dass auch AID Cytosin zu Uracil auf der mRNA editieren könnte, um damit die Synthese einer Endonuklease zu hervorzurufen, die DNA-Strangbrüche in den variablen V- bzw. konstanten S-Regionen katalysiert [48], [57]. Diese Strangläsionen könnten die Hypermutation induzieren (siehe auch Abb. 6 auf Seite 15). Auch wurde gezeigt, dass APOBEC-1 nur zusammen mit seinem Partnerprotein ACF (Apobec-1 Complementation Factor) die mRNA binden kann [47].

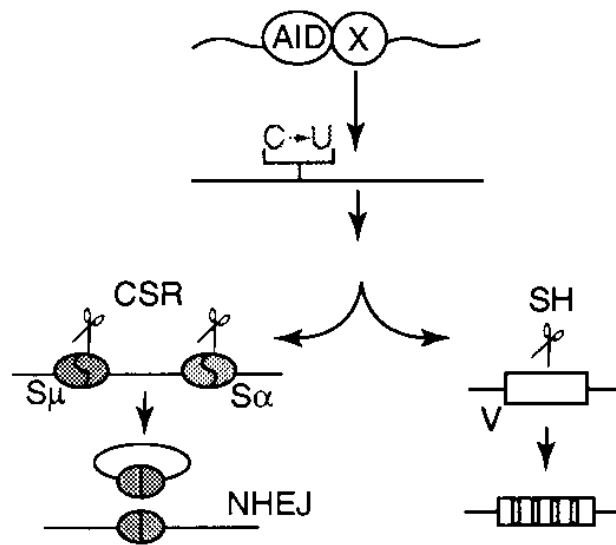


Abbildung 6: Durch die Sequenzhomologie von AID zu APOBEC-1 wird über die mögliche Aktivität von AID spekuliert: AID und ein akzessorisches Molekül (X) könnten eine mRNA binden und editieren, in diesem Falle ein C zu U deaminieren, und damit eine reife mRNA erzeugen, die ein DNA-schneidendes Enzym kodiert, das in „Class-switch-recombination“ und somatische Hypermutation involviert ist. CSR: Schneiden der S-Regionen und anschließende Reparatur durch NHEJ. SH: DNA Doppelstrangbrüche im V-Exon und anschließende Reparatur durch fehleranfällige Polymerasen, die Punktmutationen erzeugen [50].

Genkonversion, somatische Hypermutation und „Class-switch-recombination“ sind mehrstufige Prozesse. DNA-Strangbrüche und Reparatur konnten dabei nachgewiesen werden. Auch ist die Mutation von der Transkriptionsaktivität abhängig [37]. AID wirkt also während eines oder mehrerer dieser Zwischenschritte. Für die Reparatur durch NHEJ scheint AID nicht notwendig zu sein, da NHEJ auch bei

V(D)J-Rekombination der Reparaturmechanismus ist und die Abwesenheit von AID die V(D)J-Rekombination nicht beeinträchtigt [49]. Die in hypermutierenden Zellen detektierten Doppelstrangbrüche führten zu der Vermutung, dass diese die Mutationsmechanismen initiieren könnten [67], [56] [14]. Doppelstrangbrüche wurden aber auch in AID-defizienten B-Zellen gefunden, so dass die detektierten Doppelstrangbrüche nicht als alleinige Erklärung für die Mutation ausreichen [15],[19], [57].

AID könnte direkt an der DNA in der V- bzw. C-Region wirken, indem dort Cytosine zu Uracil deaminiert werden. Die so erzeugten Mutationen würden Reparaturprozesse induzieren, die zu Genkonversion, somatischer Hypermutation bzw. „Class switch recombination“ führen könnten [58].

In der Arbeitshypothese, Abschnitt 1, wird die mögliche Einwirkung von AID in drei Szenarien zusammengefasst. In den hier durchgeführten Experimenten wurde der Einfluss auf die Transkription als Möglichkeit vorausgesetzt und untersucht.

2.3.1 AID und die autosomal rezessive Form des Hyper-IGM- Syndroms

Das Hyper-IgM-Syndrom ist durch normale oder erhöhte Serum-IgM Werte und eine Abwesenheit von IgG, IgA und IgE charakterisiert, was mit einer stark erhöhten Anfälligkeit für bakterielle Infektionen einhergeht [53]. Die X-chromosomale Form des Syndroms (HIGM1) ist Folge einer Mutation im CD40-Ligand-kodierenden Gen [4]. Der CD40-Ligand ist ein auf Helfer-T-Zellen lokalisiertes Membranprotein, das durch Bindung von CD40 auf B-Zellen Proliferation, Adhäsion und Class-switch induziert. Bei der autosomal rezessiven Form (HIGM2) ist der CD40-Ligand intakt. Durch eine Mutation im AID-Gen kommt es zu dieser Form des Hyper-IgM-Syndroms mit folgenden Hauptmerkmalen:

- Abwesenheit von „Class-switch-recombinatin“ und somatischer Hypermutation
- Lymphknotenhyperplasie bedingt durch Riesenkeimzentren

[61]

2.3.2 AID-Knockout in DT40 Cre1

Die Funktion eines Gens kann durch gezielte Genausschaltung in DT40 Cre1 untersucht werden. DT40 Cre1 AID^{-/-} ist eine AID-Knockout-Mutante von DT40 Cre1, in der die Genkonversionsaktivität vollständig zum Erliegen kommt [5]. Mit DT40 Cre1 und DT40 Cre1 AID^{-/-} hat man zwei Zelllinien in der Hand, die sich hinsichtlich ihrer Genkonversionsaktivität und AID-Expression deutlich unterscheiden. Falls diese Unterschiede durch eine veränderte Transkription von

AID-Zielgenen bedingt sind, könnte man z.B. mittels „Northern blots“ gezielt nach differentiell exprimierten Genen suchen. Dafür müsste man Kandidatengene formulieren können, die dabei vermutlich eine Rolle spielen. Transkriptionsfaktoren, Nukleasen, Polymerasen wären als solche Kandidaten zu nennen. Man müsste dazu aber eine recht genaue Vorstellung davon haben, welche Gene durch AID beeinflusst werden. Besser wäre es in diesem Fall, sich einer ungezielten Methode, wie der im nächsten Abschnitt beschriebenen, zu bedienen bei der auch unbekannte Gene detektiert werden können, da man über die Weise, in der AID wirkt noch nicht viel weiß.

2.4 Mikroarrays zur Transkriptionsanalyse

Um das Genom eines Organismus zu analysieren, können verschiedene Methoden angewandt werden. Eine davon ist das Anlegen sogenannter EST-Datenbanken, die aus Sequenzen von ESTs, „expressed sequenced tags“ bestehen. Sie repräsentieren den Teil des Genoms, der tatsächlich transkribiert und gegebenenfalls auch translatiert wird, also hauptsächlich die Gene. Die relative Überschaubarkeit des menschlichen Genoms mit ca. 33 000 Genen [82] kann die Komplexität eines Organismus nicht erklären. Die Kenntnis der Basenabfolge auf der DNA reicht bisher nicht aus, um Aussagen über die Funktion einzelner Gene im Zusammenspiel mit anderen Genen, über deren Einbindung in zelluläre Prozesse, Wachstum, Differenzierung und Krankheit zu machen. Um diese Prozesse zu studieren, richtet man seinen Blick auf Transkription, Translation und deren Regulation. In den letzten Jahren wurde eine Methode entwickelt, mit der die mRNA-Population einer Zelle quantifiziert werden kann [16]. Dabei können gleichzeitig tausende transkribierter Gene einer Zelle unabhängig voneinander analysiert werden. Ein solches „Transkriptionsprofil“ zu einem bestimmten Zeitpunkt oder unter bestimmten Umständen zeigt wie eine Momentaufnahme die Transkriptionsaktivität der Zelle. Durch die Analyse und den Vergleich mehrerer solcher Profile zu verschiedenen Zeitpunkten oder unter verschiedenen Umständen kann durch die unterschiedliche Expression von Genen auf deren Einbindung in den beobachteten Prozess geschlossen werden. Durch Clusteralgorithmen lassen sich Gene zusammenfassen, die ein ähnliches Expressionsmuster zeigen. Da man zeigen konnte, dass Gene mit ähnlicher Funktion in ähnlicher Weise ihre Expression unter verschiedenen Bedingungen verändern, können Vermutungen über die Funktion unbekannter Gene angestellt werden, die mit schon bekannten Genen koexprimiert werden [23], [38]. Ein Transkriptionsprofil ist ein zelltypischer und zellzustandtypischer molekularer Phänotyp [87]. Im Gegensatz zum statischen Genom, das in jeder Chromatin enthaltenden Zelle eines Organismus vorhanden ist, gibt es so viele „Transkriptome“, wie es Zellarten multipliziert mit Umweltfaktoren gibt [13]. Die Beschreibung der Transkription von Genen ist also stets an eine detaillierte Beschreibung der Umstände geknüpft, unter denen sie gemessen wird.

2.4.1 Bedeutung von Transkriptionsprofilen für die medizinische Diagnostik und Therapie

Unter Krankheiten, die unter dem gleichen Namen laufen, gibt es oft große Unterschiede in Symptomatik, Verlauf und Ansprechen auf Therapie, die auf unterschiedliche molekulare Phänotypen zurückzuführen sind. Gerade unter den malignen Erkrankungen gibt es molekular sehr unterschiedliche Tumoren mit demselben Namen. Wenn es möglich wäre, diese molekularen Unterschiede in die Diagnosekriterien mit hinein zu nehmen, könnte man gezielter auf die jeweiligen Situationen reagieren und differenziertere Therapiepläne erstellen. DLBCL, „Diffuse large B-cell lymphoma“, ist der am weitesten verbreitete Subtyp des Non-Hodgkin-Lymphoms. Die Klassifikation von menschlichen Lymphomen hat sich von ihrer ursprünglichen Entdeckung durch Thomas Hodgkin 1832 bis heute basierend auf morphologischen und molekularen Eigenschaften laufend weiter entwickelt [64]. Das Klassifikationssystem der Wahl ist zur Zeit die REAL-Klassifikation, die „Revised European-American Lymphoma“-Klassifikation, nach der verschiedene morphologische Subtypen in Gruppen aufgeteilt werden [34], [52]. Klinisch ist DLBCL sehr heterogen, die Überlebensraten verschiedener Patientengruppen unterscheiden sich deutlich. Durch eine weitere Differenzierung könnte durch eine revidierte Diagnostik eine gezieltere Therapie erreicht werden. Mit der Anfertigung von Transkriptionsprofilen mittels DNA-Mikroarrays kann man innerhalb des DLBCL durch einen „Clusteralgorithmus“ zwei molekular unterschiedliche Subtypen mit differentiell exprimierten Genen beschreiben, die unterschiedlichen Stadien der Aktivität und B-Zellentwicklung zugeordnet werden können [23]. Die Überlebensraten und das Ansprechen auf die Therapie fallen in beiden Typen sehr unterschiedlich aus: Der eine Typ, mit einer deutlich höheren Überlebensrate, zeigt ein ähnliches Transkriptionsprofil wie im Keimzentrum befindliche B-Zellen. Der andere Typ hat ein ähnliches Muster wie *in vitro* aktivierte periphere B-Zellen und eine geringere Überlebensrate. DLBCL wird nach dieser Differenzierung zu zwei verschiedenen Krankheitsbildern, die unterschiedliche Therapien fordern [3]. Veränderte Diagnosekriterien durch Mikroarrays bei Brusttumoren beschreiben auch [71].

3 Material und Methoden

In diesem Teil werden die Materialien aufgeführt, mit denen die Experimente durchgeführt wurden. Der Methodenteil führt neben der Beschreibung der Experimente in die cDNA-Bibliothek dkfz426 als Quelle für die PCR-Produkte, die auf die Mikroarrayfilter aufgetragen wurden ein, beschreibt die Herstellung der Mikroarray-Filter, deren Hybridisierung und anschließende Auswertung mittels spezieller Software und Statistik.

3.1 Material

3.1.1 Filter

Ressourcenzentrum für Genomforschung, Berlin „rzpd“

3.1.2 Software

- AIDA Metrix Light
- AIDA Compare

von der Firma „raytest“

3.1.3 Radioisotope

α ^{33}P dCTP und γ ^{33}P dATP von Hartmann

3.1.4 Agarosegelelektrophorese

- Agarose (Sigma)
- EtBr (Boehringer Mannheim)
- Längenmarker λ HindIII Φ X174 HaeIII (New England Biolabs)

3.1.5 Kits

- Oligotex mRNA purification kit (Quiagen)
- Kinase-it-kit (Stratagene)
- Superscript Preamplification System (Gibco BRL)
- Expand long template PCR-System (Roche)
- Mobitec-Säulen (MoBiTec)

3.1.6 „Stripping-Solution“

- 5mM NaPi pH7,6 (Sigma)
- 0,1 % SDS (Sigma)

3.1.7 „Amasino-Puffer“/ Hybridisierungspuffer für die Oligo- Hybridisierung

- 130 mM NaPi pH7,2 (Sigma)
- 7% SDS (Sigma)
- 250 mM NaCl (Sigma)

3.1.8 „Oligowash“ für die Oligo-Hybridisierung

- 40 mM NaPi pH7,2 (Sigma)
- 1% SDS (Sigma)

3.1.9 50x Denhardt's

- 400 ml H₂O
- 5 g Ficoll (Type 400) (Serva)
- 5 g Polyvinylpyrrolodon (Merck)
- 5 g BSA (fraction V) (Boehringer)

3.1.10 Denhardt's Hybmix/ Hybridisierungslösung für die cDNA

- 20x SSC (Sigma)
- 50x Denhardt's
- 20 %SDS (Sigma)
- 100 µg/ml SsDNA (Sigma)

3.1.11 Waschpuffer I-III

- (I) 1 x SSC (Sigma) + 0,1 % SDS (Sigma)
- (II) 0,3 x SSC (Sigma) + 0,1 % SDS (Sigma)
- (III): 0,1 x SSC (Sigma) + 0,1 % SDS (Sigma)

3.1.12 Weitere Lösungen und Puffer

- Trizol Reagenz (Gibco BRL)
- DEPC (Serva)
- Chloroform (Sigma)
- Isopropylalkohol (Sigma)
- EDTA (Merck)
- Tris-Base + HCl (Serva)
- HCl (Serva)
- SDS (Serva)
- NaOH (Sigma)

3.1.13 Weitere Reagenzien

RNAsin Ribonukleaseinhibitor (Promega)

3.1.14 Kulturmedium für DT40 Zellen

- 87 % (500 ml) RPMI Medium 1640 mit 20 mM HEPES Puffer (Gibco BRL)
- 9 % (50 ml) FBS (Biochrome KG)
- 2 % (10 ml) Penicillin-Streptomycin (5 $\mu\text{g}/\mu\text{l}$ Penicillin G Sodium, 5 $\mu\text{g}/\mu\text{l}$ Streptomycinsulfat, 0,85 % Natriumchlorid) (Gibco BRL)
- 1 % (5 ml) 200 mM L-Glutamin/ 0,85 % Natriumchlorid (Gibco BRL)
- 1 % (5 ml) Hühnerserum (Sigma)
- 0,01 % (0,05 ml) 2M β -mercaptoethanol (Sigma)

3.1.15 Gefriermedium für DT40 Zellen

- 70 % RPMI Medium 1640 (Gibco BRL)
- 20 % FBS (Sigma)
- 10 % DMSO (Merck)

3.1.16 Oligonukleotid-Primer für semiquantitative RT-PCR

(bezogen von der Firma Invitrogene, in 5'-3'-Orientierung)

- AID/AI6 CTACCGCATCACATGGTTCACCTCCTGGAG
- AID/AI24 CCCAGATCTTGCTTGTGAAGTCTTCTTATTGCTG
- β -Aktin BA1 CCCCAAGCTTACTCCCACAGCCAGCCATGG
- β -Aktin BAI GGCTCTAGATAGTCCGTCAGGTCACGGCCA
- dkfz426 6h7r2 CTGTGTCACTTAACACTCCTTCTCTTGTCC
- dkfz426 6h7s2 ACAGAGAATGATGCAAAGATGGAAACAGCC
- dkfz426 14l11r2 GATGGTCTACTTCCTGGACCGGATCCTGCA
- dkfz426 14l11s2 AGGGGCTGTGGCTGAAAAGGTCTCCATGGC
- dkfz426 24f19r2 GTGGAGTGCATCCTGTTCAATGAGGAGTAC
- dkfz426 24f19s2 GTTGCCGCTCATGTTGTGGATGGTCAGGTT
- dkfz426 24l6r2 CTGAGGCTGCAGGCAATACCTTCAGACAGA
- dkfz426 24l6s2 CACCGTTTTAAGTGTGTTTGCTACCACCAT
- dkfz426 24m4r2 TGGTCGTGGAACATCTACCATAACCATGTC
- dkfz426 24m4s2 ACACAGCCATTGCTCAGCAGCTCCCTGCAG
- dkfz426 26d21r2 GATCTCCAAATAAACCCTTCCTGAACTAA
- dkfz426 26d21s2 TTCCTTGAAAAGGACTATCAGCTCCGTGTC
- dkfz426 12l20r2 GAGCTATGATACCCGATGGAGCCATGGGAA
- dkfz426 12l20s2 AGAGTGGCTTAACCAAAGCAGGCCCTAATC
- dkfz426 17o7r2 GGCAATGGTGTGTTATCAACAAATACTGAC
- dkfz426 17o7s2 TCAACTCATCATGTGTTAAAATTGCCACC
- dkfz426 32a11r2 CGAAATAAGTAGGTCCACACATTTTCAAGC
- dkfz426 32a11s2 CAAGCATTTACCCTTGA ACTCAGTTCTAG
- dkfz426 6k19r2 ACAAGAAGGTCATCGCAACGAAGGTTTTGG
- dkfz426 6k19s2 GGATATCGTCTGTAATGGTTACGGTCTGCT

3.1.17 Geräte

- Elektrophoresekammer (Bio Rad Laboratories)
- Photometer (Amersham Biosciences)
- Hybridisierungsöfen (Bachofer)
- Phosphorimager (raytest)
- Minitransilluminator (Bio Rad Laboratories)
- Video Printer P90 (Mitsubishi)

3.2 Methoden

3.2.1 cDNA-Bibliotheken und ESTs

Eine cDNA-Bibliothek ist eine Sammlung von in einen Vektor eingebaute cDNAs, die mittels reverser Transkription aus der mRNA-Population einer Zelle oder eines Gewebes gewonnen wurde. Teilsequenzen dieser cDNAs werden als ESTs („expressed sequence tags“) bezeichnet. Da in der Regel jede cDNA von einem bestimmten Gen abstammt, sind ESTs Teilsequenzen eines exprimierten Gens.

3.2.2 Die cDNA-Bibliothek und die Datenbank bursaler B-Zellen dkfz 426

Die cDNA-Bibliothek bursaler B-Zellen „dkfz426“ besteht aus Plasmiden mit durchschnittlich 1,3 kb langen cDNA-Insertionen. Teilsequenzen der cDNA-Insertionen von über 11 000 Klonen sind in der sogenannten bursalen EST-Datenbank gespeichert (<http://swallow.gsf.de/dt40Est.html>). Die Gesamtheit der EST-Information kann als Expressionsprofil bursaler B-Zellen bezeichnet werden und stellt damit ein wichtiges Werkzeug dar, Gene zu finden, die an B-zellspezifischen Rekombinationsvorgängen zur Repertoirebildung, insbesondere der Genkonversion beteiligt sind: Homologievergleiche der ESTs mit Genen von anderen Organismen liefern oft Hinweise auf die Funktion der durch die ESTs repräsentierten Gene. So werden potentielle Kandidatengene für einen bestimmten zellbiologischen Vorgang formuliert. Durch gezielte Ausschaltung dieser Kandidatengene in der Zelllinie DT40 und die anschließende Überprüfung des resultierenden Phänotypes, kann man die Rolle des Gens in dem zur Frage stehenden Vorgang überprüfen. Die Bibliothek wurde im DKFZ-Heidelberg (Deutsches Krebsforschungszentrum) hergestellt. Für die cDNA-Bibliothek wurde die RNA aus der Bursa fabricii zwei Wochen alter Hühner mit oligo dT geprimed, in cDNA umgeschrieben und gerichtet in einen Plasmidvektor bekannter Sequenz kloniert, um die Sequenzierung vom 5'-Ende des cDNA-Inserts zu erlauben. Diese Klone werden in Mikrotiterplatten aufbewahrt. Für die Sequenzierung wurde ein Primer verwendet, der in

der Polylinkerregion des Plasmides oberhalb des cDNA-Inserts anheftet. Die meisten der so erhaltenen Sequenzstücke haben eine Länge von 400-700 Basen. Ungefähr 15 % der sequenzierten Klone sind in voller Länge vorhanden und reichen somit an das 5'-Ende der durch das Gen kodierten Sequenz. Den anderen Klonen der cDNA-Bibliothek fehlt noch Information, die allerdings durch die RACE-Methode („rapid amplification of cDNA ends“) aufgeschlüsselt werden kann. Um die Sequenzinformation aus der EST-Datenbank nutzen zu können, kann man eine Homologien- oder Stichwortsuche gegen die öffentlichen Sequenzdatenbanken vornehmen (human EST database (BLASTN), Gen Bank/EMBL (BLASTN), PIR Protein database (BLASTX)) [1].

3.2.3 Anfertigung von Mikroarrays

Einzelsträngige Nukleinsäuren werden in Form von cDNAs oder Oligonukleotiden durch ein Robotersystem auf unterschiedliche Untergründe, wie Glasplatten oder Nitrozellulosemembranen („Chips“) in hoher Dichte aufgetragen. Diese Chips werden anschliessend mit einer radioaktiv oder fluoreszenz-markierten einzelsträngigen Sonde hybridisiert. Es kommt zu einer Paarung komplementärer Basen. Mit einem Phosphoimager oder einem Laserscanner kann man dann die Signale detektieren. Eine digitale Bildsoftware verknüpft die gewonnenen Signale mit den auf dem Array positionierten Gensequenzen. Die Ergebnisse werden dann statistisch ausgewertet und mit anderen Expressionsmustern verglichen.

Für die hier durchgeführten Mikroarrayexperimente waren folgende sechs Schritte notwendig, wobei Schritt 1 und 2 vom Ressourcenzentrum für Genomforschung in Berlin ausgeführt wurden. Schritt 3-6 wurden im Rahmen dieser Arbeit nach den Protokollen 3.2.9 bis 3.2.14 durchgeführt.

1. ESTs aus der cDNA Bibliothek dkfz426 werden mittels PCR amplifiziert und gereinigt.
2. Die PCR Produkte werden mit Robotern als Duplikate auf eine Nylonfilter-Membran aufgetragen.
3. mRNA wird aus Zellen isoliert.
4. cDNA wird synthetisiert und radioaktiv mit ^{33}P markiert.
5. Hybridisierung der Membran mit der radioaktiv markierten Sonde.
6. Signaldetektion mittels Phosphoimager und anschliessende Analyse des „Transkriptionsprofils“ mit der Analysesoftware „AIDA“.

3.2.4 Herstellung der RZPD Filter

Auf jedem Nylon-Filter sind auf einer Fläche von 22 x 22 cm, in einem identischen Punktmuster 14 592 cDNA-Klone doppelt aufgetragen. Die Quelle der cDNAs

ist die bereits beschriebene Bibliothek bursaler B-Zellen dkfz426. Dabei wurden die cDNA-Insertionen aus 38 384er Klonplatten per PCR amplifiziert und als PCR-Produkte aufgetragen. Jedes PCR-Produkt liegt zur internen Kontrolle als Duplikat vor und ist in einer Menge von 10-20 ng aufgetragen.

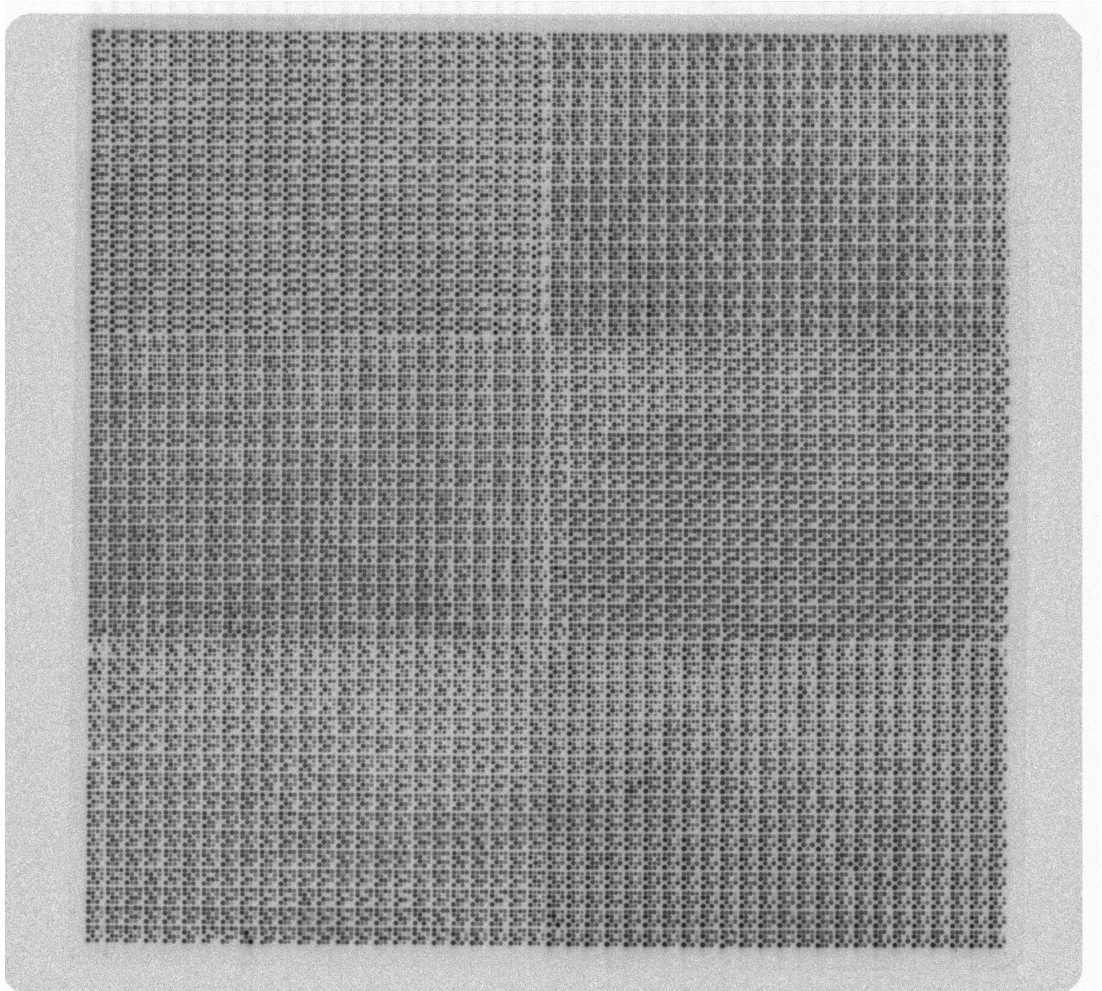


Abbildung 7: Die Punkte auf den RZPD Filtern liegen als Duplikate vor. Jeder Punkt enthält 10-20 ng DNA. Es gibt Positionen, auf denen kein PCR-Produkt aufgetragen wurde, Leerpunkte, die für die spätere Analyse wichtig sind. Die Filter sind im Ressourcenzentrum für Genomforschung (RZPD) angefertigt worden.

Von etwa 70 % der auf dem Filter aufgetragenen cDNA-Insertionen ist eine Teilsequenz in der bursalen EST-Datenbank vorhanden, von denen wiederum in etwa 60 % der Fälle signifikante Homologien mit schon bekannten Genen aus den öffentlichen Datenbanken aufweisen, und damit annotiert sind. Eine Vielzahl von ESTs überlappt mit anderen ESTs in der bursalen Datenbank. Die Information über diese zusammengehörenden ESTs kann als sogenannte „Clusterinformation“

auch der Datenbank entnommen werden (siehe 3.2.2). Man kann ein Filter bis zu fünfmal hybridisieren. Zwischen jeder Hybridisierung sollte das Filter von der Restaktivität befreit werden und für einen Zeitraum von ca. drei Wochen gelagert werden (siehe auch 4.1 im Ergebnisteil).

3.2.5 DT40 Zellen in Zellkultur

DT40 Zellen werden in einem angefeuchteten Inkubator bei 41 °C und 5 % CO_2 bis zu einer Menge von 300 ml und einer Zelldichte von 10^6 Zellen pro ml in Kultur gehalten. Zum Zeitpunkt der RNA-Isolation befinden sich die Zellen in der logarithmischen Phase ihres Wachstums und bilden unter dem Mikroskop charakteristische „Teilungs-Cluster“, wobei der Anteil von toten Zellen in der Kultur sehr gering ist.

3.2.6 Zählung der Zellen

20 μ l Zellsuspension werden mit 20 μ l Trypanblaufarbe gemischt. Ein Tropfen wird unter ein Deckgläschen auf einer Neubauer-Zählkammer gegeben. Intakte Zellen, die keine Farbe inkorporieren, werden unter dem Lichtmikroskop gezählt. Die durchschnittliche Zellzahl pro 0,1 mm^3 Feld wird mit 2×10^4 multipliziert, um die Dichte der Zellen pro ml Zellsuspension zu ermitteln.

3.2.7 Aufbewahrung von DT 40 Zellen

Für längere Lagerung werden DT40 Zellen in Gefriermedium suspendiert, das DMSO („Dimethylsulfoxid“) enthält, um eine Kristallbildung zu verhindern. 5-10 x 10^6 Zellen werden in 1 ml Medium aufgenommen und in einem Gefrierrohrchen zunächst bei -70 °C über Nacht eingefroren. Dabei umschließt man die Röhrrchen zunächst mit einem Styroporbehälter, um ein langsames Einfrieren zu gewährleisten. Nach zwei bis drei Tagen können die Zellen in einen Stickstofftank (-196 °C) transferiert, und bis zu 12 Monate gelagert werden.

3.2.8 Auftauen von Zellen

Beim Auftauen wird darauf geachtet, möglichst schnell eine Temperatur von 37 °C zu erreichen, um die Zellmortalität gering zu halten. Die Zellen werden bei 37 °C aufgetaut und direkt bei 1500 rpm zentrifugiert. Nach Absaugen des restlichen DMSO- haltigen Mediums werden sie in 1 ml Medium aufgenommen und in Kulturflaschen kultiviert.

3.2.9 RNA-Isolation und mRNA Gewinnung

- Allgemeines zum Umgang mit RNA:

Ribonukleasen sind sehr stabile Enzyme, die keine Cofaktoren benötigen, um zu funktionieren. RNAsen sind schwer zu inaktivieren. RNA kann ihnen innerhalb von Minuten zum Opfer fallen. Daher ist es beim Umgang mit RNA besonders wichtig, auf eine RNase freie Umgebung zu achten. Gefäße und Glaspipetten werden zunächst gebacken, anschließend in 0,1 % DEPC-Wasser über Nacht inkubiert und anschließend autoklaviert. Lösungen werden mit DEPC behandeltem Wasser angesetzt.

- RNA Isolierung mit TRIZOL-Reagenz:

300 ml Zellsuspension werden auf Falcontubes verteilt. Dann werden die Zellen bei 1500 Umdrehungen pro min. herunterzentrifugiert.

1. Homogenisierung

Für Zellen, die in Suspension gewachsen sind, nimmt man 1 ml Trizol für $5-10 \times 10^6$ Zellen. Man lysiert die Zellen durch repetitives Pipettieren und inkubiert sie anschliessend 5 min. bei Raumtemperatur. (Für 300 ml DT40 Zellen mit 10^6 Zellen pro ml werden 30 ml Trizol eingesetzt).

2. Phasenseparation

0,2 ml Chloroform werden pro 1 ml Trizol dazugegeben. Das Gemisch wird 15 s per Hand geschüttelt und 2-3 Minuten bei Raumtemperatur inkubiert. Dann wird für 30 min. bei 4 °C und 4000 Umdrehungen pro min. zentrifugiert. Nach diesem Schritt wird die RNA ausschliesslich in der oberen wässrigen Phase erwartet. (Für 30 ml Trizol werden 6 ml Chloroform eingesetzt).

3. RNA-Präzipitation (Ausfällung)

Die wässrige Phase wird vorsichtig in ein neues Falcontube gegeben. Dazu werden 0,5 ml Isopropylalcohol per 1 ml Trizol gegeben. Wieder wird für 30 min. bei 4 °C und 4000 Umdrehungen pro min. zentrifugiert. (Für 30 ml Trizol werden 15 ml Isopropylalkohol eingesetzt).

4. RNA-Waschen

Der Überstand wird abgegossen, das verbleibende RNA-Pellet mit 75 % Ethanol gewaschen und für 15 min. getrocknet. Die RNA wird zuletzt in RNase freiem Wasser gelöst und für 10 min. bei 55-60 °C inkubiert.

- RNA Qualitätskontrolle durch Gelelektrophorese:

Die Integrität und Größenverteilung der RNA wird im Agarosegel anhand der ribosomalen Banden überprüft. Die meisten RNAs einer Zelle, rRNAs,

tRNAs und leichte nukleäre RNAs, die in einer definierten Länge und Sequenz vorliegen, bilden auf einem Gelelektrophoresebild charakteristische Banden. Zwischen den ribosomalen Banden auf einem Elektrophoresebild liegt die mRNA, die einen Schmier zwischen einigen 100 Basenpaaren und mehreren Kilobasen bildet.

1. **Bereitung des 0,8 %igen Agarosegels**
In 200 ml TE werden 1,6 g Agarose gegeben und in der Mikrowelle erhitzt, nicht gekocht, bis sich die Agarose vollständig gelöst hat. Man läßt das Gemisch abkühlen. Dann setzt man 20 μl Ethidiumbromid (1 mg/ml) hinzu und gießt es anschließend in einen Elektrophoreseschlitten und wartet, bis das Gel erstarrt ist.
2. **Laden des Gels**
10 μg RNA und 2 μl 10 x Färbepuffer werden in einem Volumen von 20 μl pro Tasche geladen. Als Längenmarker wird 1 μg Lambda Hind III ΦX174 Hae III DNA verwendet.
3. **Elektrophorese**
Bei 120 V läuft die Elektrophorese 50 min. in 1 x TAE Puffer.
4. **Auswertung des Bildes mit 300 nm UV Transilluminator**

- **RNA-Quantifizierung und Reinheit:**

Um die Menge und Qualität einer RNA-Probe zu messen, bedient man sich am besten eines Spektrophotometers. Eine Absorption von einer Einheit bei 260 nm Wellenlänge entspricht einer Menge von 40 ng pro μl RNA in DEPC-behandeltem Wasser gelöst. Auch hier ist es wichtig darauf zu achten, dass man eine RNase-freie Cuvette benutzt. Man kann die Cuvette mit 0,1 M NaOH, 1 mM EDTA und RNase freiem Wasser (DEPC behandelt) auswaschen. Die Ausbeute an RNA, die aus 300 ml Zellsuspension (10^6 Zellen pro ml) gewonnen wurde, bewegte sich hier zwischen 500 μg und 1,5 mg RNA. Die Ausgangsmenge ist für die benötigte Menge an Oligotexpartikeln von Belang (s.u.). Die Reinheit der RNA kann durch die Messung des Verhältnisses zwischen der Absorption bei 260 nm und 280 nm abgeschätzt werden. Sie sollte zwischen 1,9 und 2,1 liegen.

- **Prinzip der mRNA Gewinnung mit dem Quiagen Kit:**

Eine typische Säugetierzelle enthält 10-30 pg RNA. Die meisten dieser RNAs sind tRNAs und rRNAs. Der Anteil der mRNA an der gesamten RNA einer Zelle beträgt 1-5 %, abhängig vom Zelltyp und physiologischen Zustand. Es gibt ca. 360 000 mRNA Moleküle in einer einzigen Säugerzelle. Seltene mRNAs, die insgesamt einen Anteil von 45 % der mRNA Population bilden, können mit 5-15 Kopien in einer sehr geringen Molekülzahl

vorliegen [2]. Um den relativen Anteil an mRNA in einer RNA-Population als Ausgangsmaterial für die cDNA Synthese zu erhöhen, bedient man sich patentierter Oligotexpartikel. Die meisten der eukaryotischen mRNAs werden posttranskriptionell polyadenyliert, rRNAs und tRNAs hingegen nicht. Da Poly(A)-Schwänze mit Oligo dT-Sequenzen hybridisieren können, hat man spezielle Polystyren-Latexpartikel entwickelt, an die kovalent dC₁₀T₃₀-Oligonukleotide gebunden sind.

- Protokoll der mRNA Gewinnung mit dem Quiagen Kit:

Man erhitzt die Oligotex-Suspension auf 37 °C in einem Heizblock. Nach der Erhitzung mixt man die Suspension mit einem Vortexer und lässt sie bei Raumtemperatur stehen. Den OEB-Puffer lässt man bei 70 °C in einem Heizblock stehen. Alle Schritte werden bei Raumtemperatur durchgeführt (20-30 °C). Alle Zentrifugationsschritte werden bei 14 000 x g durchgeführt. Die zuzusetzenden Mengen entnimmt man einer Tabelle. Sie beziehen sich auf die Menge der verwendeten RNA.

1. Hybridisierung

Nachdem die Menge der RNA bestimmt ist, pipettiert man sie in ein 1,5 ml Reaktionsröhrchen und füllt sie gegebenenfalls mit RNase freiem Wasser auf das im Kitprotokoll angegebene Volumen auf. Dazu gibt man die Oligotex-Suspension und OBB Puffer hinzu und mixt durch repetitives pipettieren. Dann inkubiert man das Gemisch für drei Minuten bei 70 °C im Heizblock. Die Sekundärstruktur der RNA wird so zerstört. Man entnimmt das Gemisch dem Heizblock und inkubiert bei Raumtemperatur (20-30 °C) für 10 Minuten. Jetzt können die dT₃₀-Oligotex-Partikel und der Poly(A)-Schwanz der mRNA hybridisieren.

2. Abtrennung des Oligotex-mRNA-Komplexes

Man zentrifugiert den Oligotex-mRNA-Komplex für zwei Minuten bei 14 000 x g) und pipettiert den Überstand ab.

3. Waschen des Oligotex-mRNA-Komplexes

Jetzt resuspendiert man das Oligotex-mRNA Pellet in der angegebenen Menge OW Puffer, pipettiert die Suspension auf eine Filtersäule, die man in ein 1,5 ml Reaktionsröhrchen plaziert und zentrifugiert für eine Minute bei maximaler Umdrehung. Man transferiert die Filtersäule auf ein neues 1,5 ml Reaktionsröhrchen und pipettiert nochmal die gleiche Menge OW-Puffer darauf. Dann zentrifugiert man wieder bei maximaler Umdrehung. Den Durchlauf gießt man weg.

4. Eluierung der mRNA

Die Filtersäule transferiert man nun in ein neues 1,5 ml Eppendorfröhrchen, pipettiert 70 °C heißen OEB-Puffer auf die Säule und mischt

durch repetitives pipettieren. Man zentrifugiert wieder bei höchster Umdrehungszahl. Erneut pipettiert man OEB-Puffer auf die gleiche Säule, mischt und zentrifugiert. Das Eluat enthält nun die aufgereinigte mRNA. RNA bewahrt man bei -20 bis -70 °C auf.

- Quantifizierung und Qualitätskontrolle der mRNA:

Die Menge der mRNA kann man, wie die der Gesamt-RNA, mittels Spektrophotometer ermitteln. Die Integrität und Grössenverteilung der aufgereinigten mRNA kann man durch denaturierende Agarosegelelektrophorese darstellen. Da die mRNA Menge hier gering war, wurde die Qualität nach der cDNA Synthese mit einem Gelbild getestet. Dabei braucht man nur eine geringe Menge an cDNA, da sehr wenig Radioaktivität ausreicht, einen Film zu schwärzen.

3.2.10 cDNA-Synthese

mRNA ist relativ instabil. Um eine stabile Sonde für das spätere Hybridisierungsexperiment zu erhalten, schreibt man die RNA mittels reverser Transkriptase in cDNA um. Dabei wird gleichzeitig ^{33}P markiertes dCTP eingebaut. ^{33}P hat eine kürzere Reichweite als ^{32}P und erzeugt schärfere Signale.

- Einzelstrang cDNA Synthese mit Superscript II RT:

0,5-1 μg RNA werden in einem Reaktionsröhrchen mit der gleichen Menge dTV Primer (dT₁₈) in einem Gesamtvolumen von 10,5 μl (evtl. aufgefüllt mit DEPC H_2O) auf 70 °C erhitzt, dann kurz auf Eis gekühlt. Bei Raumtemperatur werden in folgender Reihenfolge hinzugegeben:

1. 0,5 μl RNasin (RNase-Inhibitor)
2. 5,0 μl 5x RT-Puffer
3. 2,5 μl 0,1 M DTT
4. 0,5 μl 20 mM dGTP, dATP, dTTP
5. 5,0 μl α ^{33}P dCTP (10 $\mu\text{Ci}/\mu\text{l}$)

Das Gemisch wird für eine Minute bei 37 °C vorgewärmt. Jetzt wird 1 μl Superscript II RT hinzugegeben. In einem Heizblock wird für 2 Stunden bei 37 °C inkubiert.

- Hydrolyse der RNA, Gewinnung der Einzelstrang-cDNA:

Bei Raumtemperatur werden dem Reaktionsgemisch in folgender Reihenfolge hinzupipettiert:

1. 1 μl 0,5 M EDTA pH 8,0

2. 1 μl 10 % SDS
3. 3 μl 3 N NaOH

Bei 68 °C wird 30 Minuten inkubiert und hinterher auf Raumtemperatur abgekühlt. Bei Raumtemperatur werden hinzugefügt:

1. 1 μl 1 M Tris HCl pH 8,0
2. 3 μl 2 N HCl

Das Volumen beträgt jetzt 33,5 μl . Man gibt 16,5 μl H_2O hinzu, um auf ein Volumen von 50 μl zu kommen und so die Säulen-Reinigung der Sonde von nicht inkorporierten Nukleotiden zu erleichtern .

- Entfernung der nicht inkorporierten Radionukleotide mittels „MoBiTec“ Zentrifugationssäulen und Messung der Inkorporation:

1-2 μl der vorbereiteten Sonde werden in ein Röhrchen mit Scintillationsflüssigkeit gefüllt. Mit einem Scintillationszähler werden die cpm gemessen. Jetzt bereitet man ein Röhrchen für die Abtrennung der nicht inkorporierten Nuklide vor:

1. Um das Harz zu resuspendieren, vortext man zunächst eine Zentrifugationssäule.
2. Man entfernt den Verschluss des unteren Endes und lockert den Deckel der Säule
3. Für eine Minute zentrifugiert man bei 2500 rpm über ein Reaktionsröhrchen und stellt die Säule hinterher in ein neues.
4. Dann lädt man die cDNA Sonde auf die Säule und zentrifugiert bei gleicher Umdrehungszahl für zwei Minuten. Die nicht inkorporierten Nukleotide sollen in der Säule hängenbleiben.

Man misst nun erneut die Radioaktivität in der Scintillationsflüssigkeit und berechnet die Inkorporation. Sie sollte ca. 50 % betragen. Zum Schluß hitzedenaturiert man die Sonde für 5 Minuten bei 100 °C, um Sekundärstrukturen zu zerstören und legt sie kurz auf Eis.

Für die Hybridisierung sollte man ca. 10^6 cpm/ml Hybridisierungspuffer benutzen.

- Qualitätskontrolle der cDNA mittels Agarosegelelektrophorese: Um die Größenverteilung der bereits radioaktiv markierten cDNA zu erhalten, kann man ein Agarosegel herstellen, das über Nacht auf einem Film exponiert wird. Das Gel wird nach dem oben beschriebenen Protokoll (siehe 3.2.9 auf S. 27) vorbereitet und mit cDNA beladen. Nach der Elektrophorese wird es auf einem Geltrockner getrocknet und über Nacht einem Film exponiert. Am nächsten Tag kann man die Größenverteilung als Schwärzung sehen.

3.2.11 „Stripping“ der RZPD Filter

Die Filter werden nach jeder Hybridisierung gestrippt, um verbliebene Radioaktivität zu entfernen damit die Filter wieder verwendet werden können, was bis zu fünf mal möglich ist. Dazu wird die „Stripping solution“ auf 100 °C in der Mikrowelle erhitzt und in einen Behälter gegossen. Der Filter wird langsam und vorsichtig beginnend von einer Ecke in die heisse Flüssigkeit getaucht und 45 Minuten bei leichtem Schütteln inkubiert.

3.2.12 Hybridisierung der RZPD Filter mit Oligonukleotiden

Die Oligohybridisierung dient als Kontrolle: Die Oligonukleotide hybridisieren dabei je nach Menge an aufgetragenen PCR-Produkten auf jedem Punkt gleich, da es sich bei ihnen um unspezifische 13 bp lange Oligonukleotide handelt. Die Oligohybridisierung kann im Vergleich mit der komplexen cDNA-Hybridisierung hintergrundproduzierende Variablen kennzeichnen, z.B. unsaubere Zonen auf dem Filter, oder Unregelmäßigkeiten. Eine Darstellung der Korrelation der Duplikate eines Filters, sowie die gleicher Werte auf zwei Filtern (siehe Abb. 24 und Abb. 31), dient im Vergleich mit den komplexen Hybridisierungen als Normierung. Auch werden die Oligo-Hybridisierungen in der statistischen Auswertung als weitere Sonden betrachtet, deren Wert in die Berechnung eingeht. Die beiden verwendeten Filter wurden vor der Hybridisierung mit den cDNAs mit Oligonukleotiden hybridisiert.

- Bereitung der Oligosonde mit dem „kinase-it-kit“ Mittels PNK (Polynukleotidkinase) und γ ^{33}P dATP werden die einzelnen Oligonukleotide mit ^{33}P markiert. Nacheinander pipettiert man in ein Reaktionsröhrchen:
 1. 0,4 μl M_{13} Oligonukleotide (25 pmol/ μl)
 2. 3,5 μl Kinase-Puffer
 3. 25,1 μl H_2O
 4. 5 μl γ ^{33}P dATP
 5. 1 μl T_4 Polynukleotidkinase (PNK)

Insgesamt 35 μl Reaktionsgemisch werden bei 37 °C 30 Minuten lang inkubiert. Anschliessend füllt man mit 65 μl gefiltertem H_2O auf 100 μl auf, um ein ausreichendes Volumen für die Säulen-Reinigung zu haben. Die überschüssigen Radionuklide entfernt man hinterher mit einer unter 3.7.3 beschriebenen „MoBiTec“ Zentrifugationssäule.

- Tag 1, Prähybridisierung; Hybridisierung:

Der Filter wird in einer Glasflasche plaziert und bei Raumtemperatur (25 °C) in „Amasino“ für 30 Minuten prähybridisiert. Im Anschluss daran wird

die radioaktiv markierte Oligosonde hinzugegeben und über Nacht bei Raumtemperatur hybridisiert.

- Tag 2, Waschen; Exponieren:

Der Filter wird zweimal für jeweils 30 Minuten bei Raumtemperatur mit „Oligowash“ in der Flasche gewaschen und anschliessend nach Antrocknen auf Zellstofftüchern in Frischhaltefolie (ohne Luftblasenbildung) eingepackt und über Nacht auf einem Phosphoimagefilm exponiert.

- Scanning:

Die Signale auf dem Phosphoimagefilm werden mit einem Phosphoimager detektiert und in Bilddaten umgesetzt.

3.2.13 Hybridisierung der RZPD Filter mit den cDNA Sonden

Für die Hybridisierung wird die cDNA aus unterschiedlichen RNA-Isolationsexperimenten gewonnen (siehe 4.1.1), wobei die Zellkonditionen gleich zu halten sind. Die RNA von DT40 Cre1 und DT40 Cre1 AID-/- wird für jede Hybridisierung einmal aus der jeweiligen Zellkultur isoliert und in cDNA umgeschrieben. Anschließend erfolgt die Hybridisierung nach dem folgenden Protokoll.

- Tag 1, Prähybridisierung und Hybridisierung:

1. 30 ml „Denhardt’s Hybmix“ werden bei 65 °C aufgetaut.
2. 7,5 μl (1 $\mu\text{g}/\mu\text{l}$) $(dA)_{40}$ wird 5 Minuten bei 100 °C hitzedenaturiert und auf Eis gelegt. Es soll durch Hybridisierung mit poly-dT-Residuen den Hintergrund reduzieren.
3. Der nasse Filter wird in einer Glasflasche plaziert, wobei wichtig ist, dass sich zwischen Glas und Filter keine Luftblasen befinden.
4. „Denhardt’s Hybmix“ und $(dA)_{40}$ werden in die Glasflasche gegeben.
5. Die Flasche mit Inhalt wird in einen Hybridisierungssofen gesetzt und bei 65 °C für zwei Stunden prähybridisiert.
6. Anschliessend wird die Sonde (10^6 cpm) hinzugegeben und über Nacht bei 65 °C hybridisiert.

- Tag 2, Waschen:

Beim Waschen wird sukzessiv die Stringenz erhöht (siehe Zusammensetzung der Waschlösungen I-III im Materialteil):

1. Die Waschpuffer werden bei 65 °C im Wasserbad vorgewärmt.

2. Der Filter wird zunächst für 20 Minuten in Lösung I bei 65 °C in der Flasche gewaschen.
3. Im zweiten Waschgang wäscht man für 10 Minuten in Lösung II bei 65 °C in der Flasche.
4. Der Filter wird nun in einen Behälter transferiert, in dem er erst für 10 Minuten in Lösung II, und dann für weitere 10 Minuten in Lösung III bei 65 °C gewaschen wird.
5. Der Filter wird mit der DNA Seite nach unten auf Tüchern getrocknet.
6. Dann packt man den Filter in Frischhaltefolie ein, wobei man darauf achten muss, dass zwischen der Folie und dem Filter keine Luftblasen sind.
7. Der Filter wird nun 18-24 Stunden auf einem Phosphoimagefilm exponiert.

- Tag 3, Scanning des Phosphoimagefilms:

Nach 18 Stunden wird der Filter mittels Phosphoimager gescannt.

3.2.14 Auswertung der Filter mit dem „Advanced Image Data Analysingprogram (AIDA)“

Die eingescannten Mikroarray-Filter werden mit dem AIDA-Programm ausgewertet, um für jedes Gen, bzw. EST einen vergleichbaren Expressionswert zu erhalten. Die Rohwerte sind das erste Ergebnis. Von ihnen wird der Hintergrund abgezogen, der durch eine eigene Signalintensität die Werte verzerrt, sowie eine Normalisierung vorgenommen, um die Expressionswerte vergleichbar zu machen. AIDA erzeugt Tabellen, in denen die Werte einzeln aufgestellt („AIDA Metrix“), oder einander gegenübergestellt („AIDA compare“) werden (siehe auch Abb. 23 auf S. 48 und Abb. 30 auf S. 53 im Ergebnisteil). Die differentielle Expression eines Gens wird durch das Verhältnis („ratio“) ausgedrückt, das durch die Division der zu vergleichenden hintergrundkorrigierten und normalisierten Rohwerte entsteht. Dieses Verhältnis wird in der Literatur als „Fold difference“ [31] bezeichnet, und beschreibt den x maligen Expressionsunterschied.

3.2.15 „Noise“

„Noise“ oder „Rauschen“ ist der Begriff für falsch positive, oder falsch negative Signale. Falsch negative Signale werden gar nicht detektiert. Falsch positive Signale können eine Signifikanz vortäuschen, die in Wirklichkeit in das Hintergrundrauschen fällt. Das Verhältnis zwischen Expressionswerten im niedrigen Bereich zur Noise ist niedriger als das von Werten im hohen Expressionsbereich. Daher kann man im niedrigen Bereich auch sehr schwer differentiell exprimierte Gene

detektieren. Die meisten Gene liegen in diesem Bereich niedriger Expressionsunterschiede, was man an der Form der Graphen auf Seite 55 sehen kann. Auch bei der Hybridisierung zweier gleicher cDNAs misst man unterschiedliche Signalwerte (siehe Abb. 32 auf S. 54). Solche falsch positiven Messwerte kommen in jedem Experiment vor und haben unterschiedliche Ursachen [33]:

- Biologische Varianz: Durch das Bemühen, die RNA der Zellen unter möglichst ähnlichen Bedingungen zu isolieren, wurde versucht, diese Varianz klein zu halten (siehe auch DT40-Zellen in Kultur).
- Varianz in der technischen Anfertigung des Mikroarrays, der Hybridisierungs- und Waschkonditionen, unterschiedliche Inkorporation der Radionuklide: Diese Varianz wurde durch möglichst sorgfältiges Arbeiten klein gehalten.
- Andere Varianzen, z.B. unpräzise Signaldetektion durch den Phosphoimager, konnten nicht gänzlich ausgeschlossen werden.

3.2.16 „AIDA Array Metrix“

Die eingescannten Grauwerte werden durch das Programm in Zahlenwerte umgesetzt. Ziel ist es, die Werte miteinander vergleichbar, und damit die differentielle Expression als Differenz zwischen zwei Zahlenwerten beschreibbar zu machen. Der Filter enthält die Information über einen Grauwert auf jedem Punkt, der als Koordinate festgelegt ist. Sichtbar ist ein Punktmuster mit unterschiedlich hellen Punkten, deren Grauwertspektrum von der Farbe des Hintergrundes bis schwarz reicht. Zunächst muss man über dieses Punktmuster ein dafür programmiertes Gitter mit den entsprechenden Koordinaten legen. Damit wird das *Alignment* vorbereitet. Beim Alignment legt das Programm hinter jedem Gitterfeld einen Grauwert fest, der mit einem Integral-Wert beschrieben wird, dem Rohwert. Verschiedene Parameter können dabei variiert werden, wie z.B. die Schwelle, die definiert, ab wann ein Punkt als Punkt erkannt werden soll, bzw. ab wann er in das Hintergrundrauschen fällt. Ferner kann man die erlaubte Abweichung eines Punktes von seinem Duplikat (hier 15 %) sowie eine bestimmte erlaubte innere Inhomogenität festlegen. Wenn der Alignmentvorgang abgeschlossen ist, erhält man eine Ergebnistabelle, in der jedem Punkt ein Integralwert zugeordnet wird, der sich aus der Gesamtheit aller integrierten Grauwerte ableiten lässt (siehe Abb. 23 auf S. 48 im Ergebnisteil). Desweiteren kann man die Korrelation der Duplikate graphisch darstellen (siehe Abb. 24 und 25-29 im Ergebnisteil, Seite 49-52).

In einem weiteren Schritt kann man von jedem Punkt, dem ein Integralwert zugeordnet wurde, einen Hintergrundwert abziehen. Dazu gibt es verschiedene Möglichkeiten. Hier wurde die gewichtete Methode angewendet. Dabei werden alle Punkte auf dem Filter, denen ein Leerwert zugeordnet wird, als Hintergrund

definiert. Von jedem integrierten Punkt wird nun ein Wert abgezogen, der sich aus dem Grauwert der Leerpunkte ergibt, wobei der Einfluß, den jeder Leerpunkt auf einen Punkt ausübt, davon abhängt, wie weit er von ihm entfernt ist. Nähere Punkte haben danach einen größeren Einfluß als weiter entfernt liegende. Diese Methode berücksichtigt lokale Unterschiede der Signalintensität auf dem Filter. Hinterher kann man den hintergrundkorrigierten Wert ablesen (Abb. 23 auf Seite 48).

Als nächstes normalisiert man jeden Wert eines Filters, um die Resultate zweier Filter vergleichbar zu machen. Dazu wird jeder Punktwert, Integral oder Integral minus Hintergrund auf den Mittelwert aller Punkte bezogen, wobei Leer- und Orientierungspunkte ausgeschlossen wurden. In der Tabelle erscheinen dann weitere normalisierte Werte, die das Verhältnis aus Integral zu Mittelwert aller Integrale darstellen (Abb. 23 auf Seite 48).

3.2.17 „AIDA Array Compare“

Mit diesem Programm lassen sich zwei vorher mit dem Metrix-Programm berechnete Filterdaten vergleichen. Dabei stehen einem mehrere Möglichkeiten der Darstellung zur Verfügung. Zunächst gibt es eine Tabelle, in der die Daten der einzelnen Filter aufgezeigt werden. Dabei werden die Expressionswerte eines „Master-Array“-Filters einem „Client-Array“-Filter gegenübergestellt. Das Verhältnis aus den Expressionswerten ist die differentielle Expression. Die Korrelation der Werte zweier Filter wird graphisch dargestellt (siehe Korrelationsgraphen auf Seite 49-57 im Ergebnisteil).

3.2.18 Statistische Auswertung wiederholter Hybridisierungen

Ob Änderungen im Expressionsprofil experimentell signifikant oder zufällig sind, wird durch statistische Methoden beschrieben. Entscheidend ist, dass man Mikroarray-Experimente wiederholt [79]. Die Analyse zweier gleicher Filter, die zu verschiedenen Zeiten hybridisiert wurden, zeigt eine deutliche Variabilität, die durch die schon beschriebenen „Noise“-erzeugenden Faktoren entsteht (siehe Ergebnisteil). Auch aus den Tabellen läßt sich feststellen, dass gleiche Werte aus verschiedenen Hybridisierungen unterschiedlich sind. Da es bei mehr als 14 000 Werten sehr aufwendig ist, jeden einzelnen Wert durchzugehen, wendet man daher statistische Mittel an, um tatsächliche Expressionsunterschiede von unspezifischen Unterschieden unterscheiden zu können. Die Basis für die Auswertung bilden die „Normalized-Bkg“-Werte (siehe Abb. 23 auf Seite 48), die mit dem Array-Matrix-Programm erzeugt wurden (siehe 3.2.16). 3x2 Wildtyp-, 2x2 Knockout-Werte und 2x2 Oligo-Werte werden in „Excel-Tabellen“ gegenübergestellt. Dabei fallen Dublikate, die über 15 % voneinander abweichen, sowie inhomogene Werte aus der Auswertung aus. (Siehe auch 4.2.1 auf Seite 43 im Ergebnisteil). Die t-Test-Analyse, sowie die SAM-Statistik wurden vom Insti-

tut für Bioinformatik an der Universität Turku/Finnland angefertigt. Das Protokoll für die SAM- Statistik erhält man ferner über die URL: <http://www-stat.stanford.edu/SAM/servlet/SAMservlet>.

3.2.19 t-Test-Statistik

Im Allgemeinen ermittelt man mit dem t-Test, ob zwei zu vergleichende Mittelwerte im Verhältnis zur Variabilität ihrer Werte statistisch unterschiedlich sind. In diesem Fall sind die zu vergleichenden Mittelwerte die zu vergleichenden Expressionswerte zweier Filter. Zwei wiederholte Hybridisierungen liefern durch den Duplikatauftrag insgesamt vier Werte für einen Expressionswert. Die Variabilität dieser Werte geht in die Berechnung mit ein. Man legt eine Wahrscheinlichkeit (P) fest, daß die Differenz in der Genexpression zufällig identifiziert wurde. Bei einer Wahrscheinlichkeit von $P = 0,01$ werden so schon 100 Gene aus einer Ausgangsmenge von 10 000 Genen zufällig identifiziert [31], so daß bei großen Datenmengen viele falsch positive Werte erzeugt werden.

3.2.20 SAM, „significance analysis of microarrays“

Um eine wirkliche differentielle Expression von der „Noise“ unterscheiden zu können und damit signifikante Expressionsunterschiede formulieren zu können, wurde eine Methode beschrieben, die jedem Gen einen „Score“ auf der Basis der Veränderung seiner Expression relativ zur Standardabweichung wiederholter Messungen zuordnet. Der Wert der differentiellen Expression jedes Gens aus einer Messung wird über alle Werte, die aus wiederholten Hybridisierungen gemessen werden gemittelt. In diesem Fall gibt es 3 Wildtyp-, 2 Knockout-, und zwei Oligo-Hybridisierungen mit jeweils zwei Meßwerten pro Gen, wobei auch hier viele Werte hinausfallen, die zu stark von ihrem Duplikat abweichen, oder inhomogen sind. Jeder einzelne Wert, der gemessen wird, wird mit dem erwarteten Wert aus der Mittelung verglichen. Daraus ergeben sich zwei Werte: Die erwartete relative Differenz und die gemessene, bzw. beobachtete relative Differenz. Die meisten Werte fallen in den Bereich hinein, wo sich die erwartete und die gemessene relative Differenz entsprechen, oder geringfügig voneinander abweichen. Die Abweichung der Werte voneinander wird als „Score“ jedem Gen zugeordnet. Gene, deren „Scores“ über einer festgelegten Schwelle liegen, werden als potentiell signifikant bezeichnet. Der prozentuale Anteil an Genen, die darunter liegen und somit nur scheinbar signifikant sind, wird als FDR, „false discovery rate“ beschrieben. Die Schwelle kann variiert werden. Hebt man sie, werden weniger Gene als signifikant identifiziert. Gleichzeitig steigt aber die Wahrscheinlichkeit, dass es sich tatsächlich um eine differentielle Expression handelt [31].

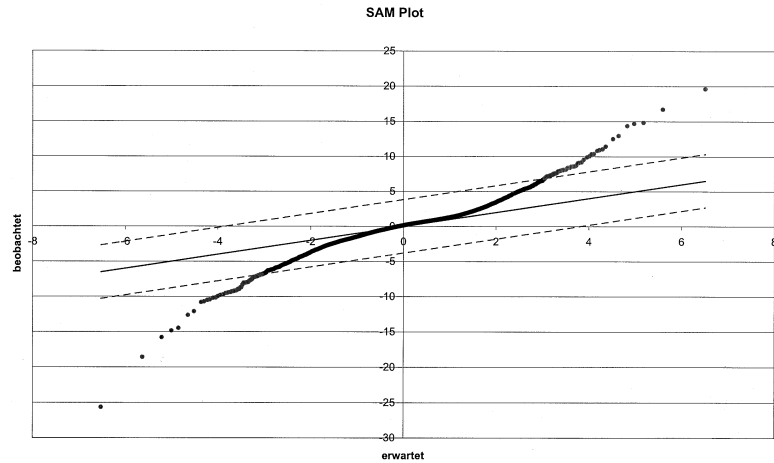


Abbildung 8: SAMplot von AID^{-/-} und DT40 Cre1. Auf der y-Achse liegen die gemessenen (=beobachteten) Werte, die gegen die erwarteten Werte aus der Mittelung aller Differenzwerte auf der x-Achse aufgetragen sind. Die Differenz zwischen erwartetem und beobachtetem Wert wird durch den „SAM-Score“ ausgedrückt. Der Schwellenwert, ab wann diese Differenz die differentielle Expression wahrscheinlich macht, wird hier auf $\pm 6,8$ festgelegt. Die meisten Werte fallen in den Bereich, in dem erwarteter und beobachteter Wert gleich, oder nicht sehr unterschiedlich sind.

3.2.21 Bestimmung der Sequenzen der differentiell exprimierten ESTs und Annotation durch Datenbanksuche

Nach Erstellung einer Tabelle mit den differentiell exprimierten ESTs, besteht die Möglichkeit über Homologiensuche in den öffentlichen Datenbanken („blsp“ (swiss protein) und „blge“ (gene bank)) verwandte Gene zu identifizieren, deren Funktion schon besser aufgeklärt wurde. Die Annotationen aus der bursalen EST-Datenbank beruhen auf diesem Prinzip. Vielleicht finden sich darunter Faktoren, die bekanntermaßen in Rekombinations- oder Reparaturprozessen eine Rolle spielen (siehe Tabellen unter 4.3).

3.2.22 Semiquantitative RT-PCR zur Bestätigung der differentiellen Expression

Um die durch AIDA und die beiden statistischen Methoden erhaltenen Werte durch ein weiteres Experiment zu bestätigen, wurde von 10 ESTs eine RT-PCR angefertigt. Dabei reicht es, zunächst nur 10 ESTs zu testen, um das Gesamtergebnis zu bestätigen (siehe auch Tab. 4 und 5). Mit dieser Methode kann man

die differentielle mRNA-Expression einzelner Gene in verschiedenen Zellen semiquantitativ darstellen und vergleichen. Dabei wird die eingesetzte cDNA aus initial 50 ng Gesamt-RNA auf 0,5 ng sukzessive erniedrigt. Man kann die Menge an amplifizierten PCR-Produkten in den verschiedenen Zellklonen vergleichen. Eine starke Bande bei schon geringen Mengen eingesetzter cDNA im Vergleich zu einer schwachen Bande bei der gleichen Menge an cDNA im anderen Zellklon, weist auf eine differentielle Expression hin.

- RNA-Isolation mit dem Trizol-Reagenz (siehe auch 3.2.9)
- cDNA-Synthese nach dem Protokoll 3.2.10 ohne radioaktive Markierung der cDNA
- Primer für die Kontrollen β -Aktin, AID und die ESTs (siehe 3.1.16)
- PCR-Protokoll: „Expand long template PCR-System“:
 1. 2 Minuten initiale Inkubation bei 93 °C
 2. 35 Zyklen, 10 Sekunden 93 °C, 30 Sekunden 65 °C, 5 Minuten 68 °C mit jeweils 20 Sekunden Verlängerung pro Zyklus
 3. Letzter Verlängerungsschritt für 7 Minuten bei 68 °C

4 Ergebnisse

Um eine mögliche differentielle Expression zwischen AID-positiver (Wildtyp-) und AID-negativer (Knockout-) Zelle darzustellen, wurden Transkriptionsprofile mittels der im Materialteil beschriebenen Filter hergestellt. Die zwei identischen Filter A und B (siehe auch 3.2.4) wurden zunächst mit den Oligonukleotidsonden zur Kontrolle und anschließend mit den komplexen cDNA-Sonden hybridisiert. Filter A wurde mit der cDNA der AID-positiven Wildtypzelle, Filter B mit der der AID-negativen Knockoutzelle hybridisiert. Zwischen den Hybridisierungen wurden die Filter nach dem Protokoll 3.2.11 gestrippt und für einen Monat zum Abklingen der Restaktivität gelagert, da die Halbwertszeit von ^{33}P ca. 25 Tage beträgt. Vor der nächsten Hybridisierung wurde mittels Phosphoimager kontrolliert, dass keine Signale mehr zu detektieren waren.

Nacheinander wurden die drei AID-positiven Wildtyp- und zwei AID-negativen Knockout-Hybridisierungen mit dem AIDA-Programm (siehe 3.2.14 im Materialteil) ausgewertet. Die gemessenen Expressionsunterschiede unterschieden sich von Hybridisierung zu Hybridisierung. Darum wurden die Ergebnisse der einzelnen Hybridisierungen mit den im Materialteil beschriebenen statistischen Methoden (siehe 3.2.18) ausgewertet. Damit wurde den Werten eine Signifikanz zugeordnet, die die Wahrscheinlichkeit anzeigt, mit der der gemessene differentielle Expressionswert im Vergleich zu denen wiederholter Hybridisierungen, auch wirklich differentiiell exprimiert war. Auch die Werte der Oligohybridisierungen gingen in die Berechnung ein, um die Filter A und B zu normalisieren.

Aus dieser statistischen Integration von Werten gleicher Positionen auf dem Filter aus unterschiedlichen Hybridisierungen, wurden einige Werte gefunden, die als differentiiell exprimiert bezeichnet wurden. Zunächst wurde in den Datenbanken nachgeschaut, ob die gefundenen Gene Homologien zu bekannten Genen zeigten und annotiert sind. Mittels RT-PCR sollten diese Werte anschließend bestätigt werden.

4.1 Experimente

4.1.1 RNA-Isolation von AID-positiven und AID-negativen DT40-Zellen

Die RNA-Quelle der drei Wildtyp-Hybridisierungen, DT40 Cre1 I, II und III ist die Zelllinie DT40 Cre1. Die Zellen stammen aus einem „batch“, d.h., sie entstammen einem Gefrierrohrchen. Zwei cDNA-Sonden, DT40 Cre1 I und II wurden aus einer RNA-Isolation genommen. Die dritte cDNA, DT40 Cre1 III, entstammt einer eigenen RNA-Isolation aus demselben „batch“. Die cDNA der Knockoutzelle, DT40 Cre1 AID^{-/-}, stammt aus einer einzigen RNA-Isolation aus der Zelllinie DT40 Cre1 AID^{-/-}. Die RNA wurde aus 300 ml Zellkultur, bzw 3 x 10⁸ Zellen gewonnen. Die Gesamt-RNA wurde nach dem beschriebenen Protokoll 3.2.9

extrahiert. Dabei wurden aus jeder Isolation zwischen 500 μg und 1,5 mg RNA gewonnen. Anschließend wurde ein Elektrophoresebild angefertigt, um die Integrität der RNA zu ermitteln. Aus der Gesamt-RNA von DT40 Cre1 I und DT40 Cre1 AID^{-/-} I wurde mit dem Oligotex-Kit (Protokoll 3.2.9) die mRNA isoliert. Die Ausbeute an mRNA betrug zwischen 5 und 10 μg . Die mRNA wurde sofort in radioaktiv markierte cDNA umgeschrieben, wobei jeweils 1,5 μg mRNA für eine Sonde genommen wurde. Die RNA von DT40 Cre1 II und DT40 Cre1 AID^{-/-} II und III wurde eingefroren und für die mRNA-Isolation aufgetaut. Dann wurde sie ebenso isoliert und in cDNA umgeschrieben. Die RNA für DT40 Cre1 III stammt aus einer weiteren RNA-Isolation. Aus ihr wurde sofort die mRNA isoliert, die anschließend in cDNA umgeschrieben wurde.

4.1.2 Menge und Reinheit der Gesamt-RNA und mRNA

Sowohl nach der RNA-Isolation als auch im Anschluß an die mRNA-Aufreinigung wurden Menge und Reinheit mit dem Photometer ermittelt. Das Verhältnis zwischen der Absorption bei 260 nm und 280 nm, mit der die Reinheit der RNA abgeschätzt wird, lag zwischen 1,9 und 2,1. Für die cDNA Synthese wurden jeweils 1,5 μg mRNA eingesetzt.

4.1.3 cDNA Synthese und Qualitätskontrolle

Die mRNA wurde, wie unter 3.2.10 beschrieben in radioaktiv markierte cDNA umgeschrieben. Eine Qualitätskontrolle der radioaktiven cDNA durch Gel-Elektrophorese zeigte eine Länge der cDNA-Fragmente zwischen 500 und 900 Basenpaaren, was für eine spezifische Hybridisierung ausreichend ist.

4.1.4 Die spezifische Aktivität der Hybridisierungs sonden

Die spezifische Aktivität betrug zwischen 0,5 und 1,4 $\times 10^6$ cpm/ μg cDNA. Es wurde darauf geachtet, dass die eingesetzte Menge an Radioaktivität bei allen Sonden ähnlich war. Es wurden ca. 1×10^6 cpm/ml Hybridisierungspuffer eingesetzt.

4.1.5 Hybridisierung

Die zwei identischen Filter A und B wurden in folgender Reihenfolge nacheinander hybridisiert: Zunächst wurden beide mit den Oligonukleotidsonden hybridisiert. Filter A wurde dann jeweils mit der cDNA aus der Wildtypzelle DT40 Cre1, und Filter B mit der cDNA aus der Knockoutzelle DT40 Cre1/AID^{-/-} hybridisiert. Dabei wurden DT40 Cre1 I zusammen mit DT40 Cre1 AID^{-/-} I, DT40 Cre1 II zusammen mit DT40 Cre1 AID^{-/-} II und DT40 Cre1 III zusammen mit DT40 Cre1 AID^{-/-} III hybridisiert, wobei DT40 Cre1 AID^{-/-} III wegen zu vieler

unspezifischer Signale nicht in die Auswertung genommen wurde. Zwischen den Hybridisierungen der Filter lag jeweils ein Monat.

4.1.6 Hybridisierung der Filter mit Oligonukleotiden

Zwei Filter wurden zunächst zur Kontrolle mit radioaktiv markierten Oligonukleotiden hybridisiert (siehe Abb. 9- 12):

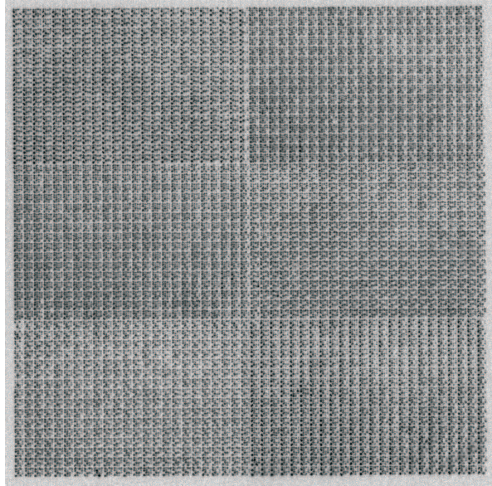


Abbildung 9: Oligohybridisierung A

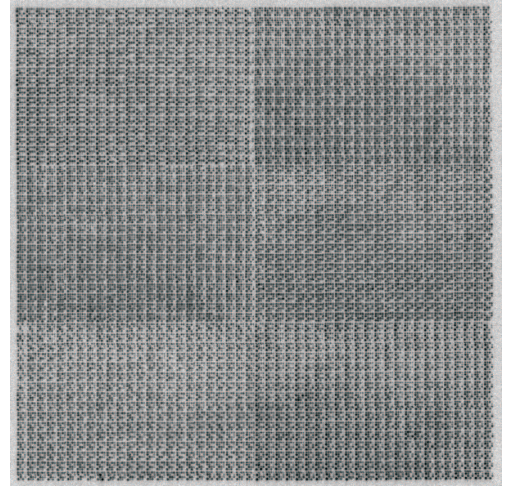


Abbildung 10: Oligohybridisierung B

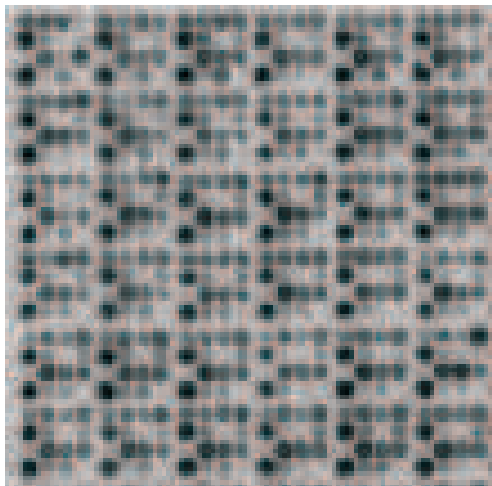


Abbildung 11: Ausschnitt aus Oligohybridisierung A

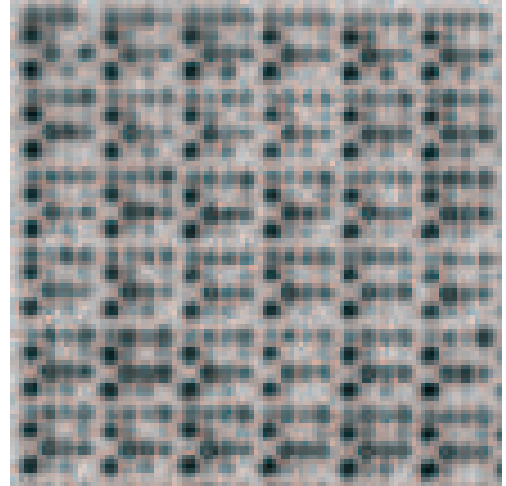


Abbildung 12: Ausschnitt aus Oligohybridisierung B

- Zunächst konnte man in der Übersicht schon sehen, dass sich ein regelmäßiges Muster zeigte, da die Oligonukleotide aufgrund der unterschiedlichen Mengen an aufgetragenem PCR-Produkt unterschiedlich binden.

- Die Duplikate auf einem Filter sollten sich idealerweise entsprechen. Ihre Abweichung voneinander konnte per Graphen dargestellt werden (siehe Korrelationsgraphen Abb. 24 auf S. 49). Man sieht an der Punktwolke, dass sich die Duplikate, also die zwei von Menge und Zusammensetzung gleichen Nukleotid-Punkte auf einem Filter, nicht genau entsprachen. Diese Abweichung der Duplikate voneinander diente als Norm für die Duplikatabweichung der komplexen Hybridisierung. Das AIDA-Programm ermöglicht es, einen Bereich zu definieren, in dem man Abweichungen zweier Duplikate voneinander toleriert. Hier wurden Punkte mit einer Abweichung von bis zu 15 % voneinander mit in die Berechnung hineingenommen.
- Werte im niedrigen Expressionsbereich streuten mehr, als im hohen Bereich. Nach Anfertigung beider Oligo-Hybridisierungen kam es zu einem solchen Ergebnis. Man kann davon ausgehen, dass aufgrund nicht auszuschließender Variablen, die sich unspezifisch auf das Ergebnis auswirken, gleiche Werte immer in einem Bereich streuen.
- Wichtig war die Normalisierung der Filter A und B, um die sich entsprechenden Werte auf zwei verschiedenen Filtern vergleichbar zu machen. Die Korrelation von Filter A und B konnte mittels Graphen in Beziehung gesetzt werden (siehe Korrelationsgraph Abb. 31). Durch die Standardabweichung des Graphen wurde die Korrelation der Oligo-Filter mit der der cDNA-Filter verglichen. Die Standardabweichung des Korrelationsgraphen der beiden Oligohybridisierungen war, wie zu erwarten, kleiner als die der komplexen Hybridisierung (siehe auch 4.2.1 und 4.2.2).
- Die Oligo-Hybridisierung wurde auch in die statistische Berechnung mit hineingenommen, wobei die Oligos als eine weitere zu vergleichende Sonde angesehen wurden (siehe 4.3 auf S. 58 und 3.2.18 im Materialteil).

4.1.7 Hybridisierung der Filter mit den cDNA-Sonden

Filter A und B wurden nach Oligohybridisierung mit den komplexen cDNA-Sonden nach dem Protokoll 3.2.13 hybridisiert. Das Transkriptionsprofil der Wildtypzelle DT40 Cre1 wurde danach dreimal auf Filter A, das der Knockoutzelle DT40 Cre1, AID^{-/-} zweimal auf Filter B dargestellt (siehe 4.1.1). Die Abb. 9-20 dienen nur der Übersicht, da sich die Expressionsunterschiede mit dem Auge natürlich nicht feststellen liessen (zur Auswertung siehe nächsten Abschnitt 4.2).

4.2 Auswertung mit dem AIDA Programm

4.2.1 Auswertung mit AIDA Array Metrix

Nach Hybridisierung und Einscannen der Bilddaten, wurden die Filter zunächst mit AIDA Metrix ausgewertet. Folgende Ergebnisse zeigten sich:

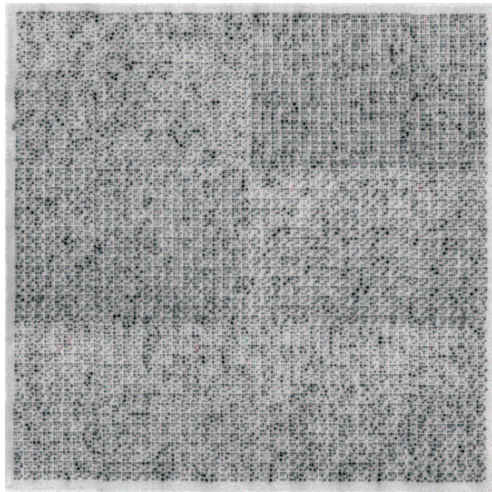


Abbildung 13: DT40 Cre1 I

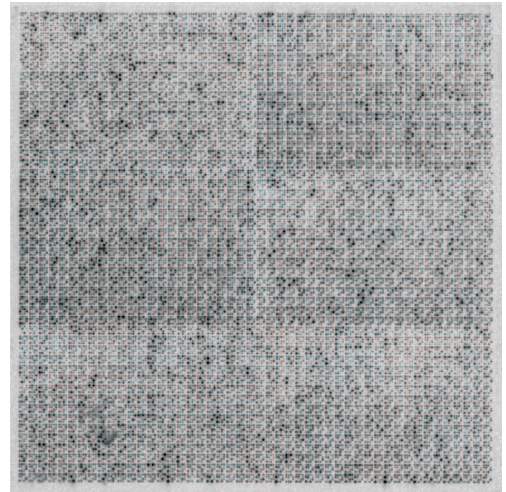


Abbildung 14: DT40 Cre1 AID^{-/-} I

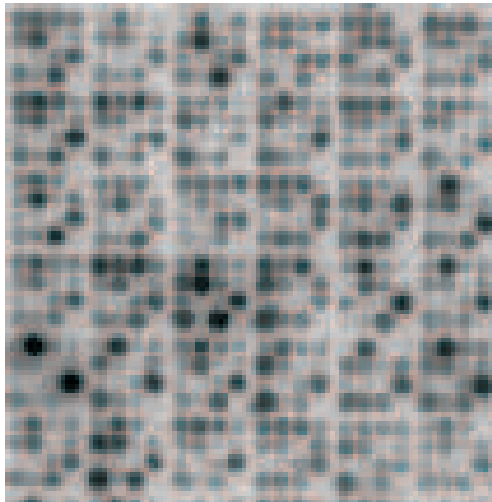


Abbildung 15: Ausschnitt aus DT40 Cre1 I

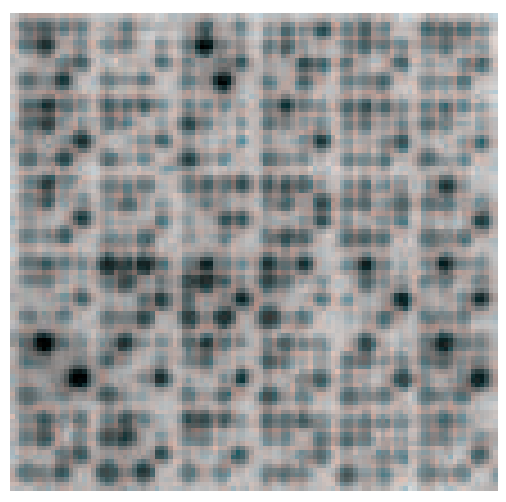


Abbildung 16: Ausschnitt aus DT40 Cre1 AID^{-/-} I

- Durch AIDA Metrix wurden in Tabellen sämtlichen Punkten Expressionswerte zugeschrieben (siehe dazu den Ausschnitt aus der Tabelle der DT40 Cre1 I-Hybridisierung, Abb. 23 auf S. 48 als Beispiel. Wegen der Grösse der Tabellen werden sie hier nicht vollständig aufgeführt).
- Nach Hintergrundkorrektur und Normalisierung (siehe 3.2.16 auf S. 35 im Materialteil) erhielt man mit „Integral normalized-background“ den hintergrundkorrigierten, normalisierten Integralwert. Dieser Wert wurde als Expressionswert verglichen und statistisch integriert.
- Es zeigte sich, dass die „Integral normalized-background“-Werte bei wie-

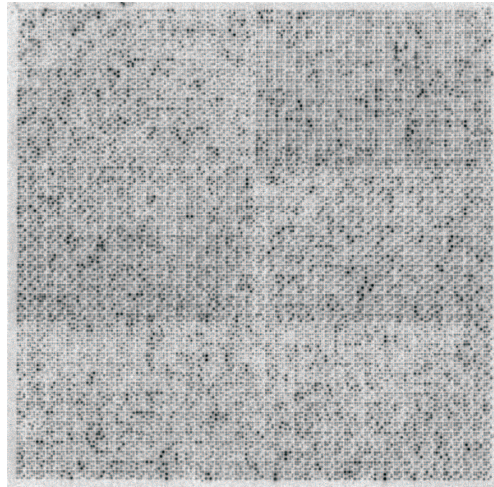


Abbildung 17: DT40 Cre1 II

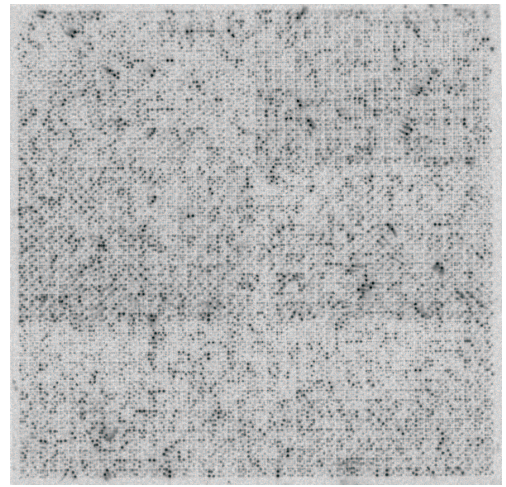


Abbildung 18: DT40 Cre1 AID^{-/-} II

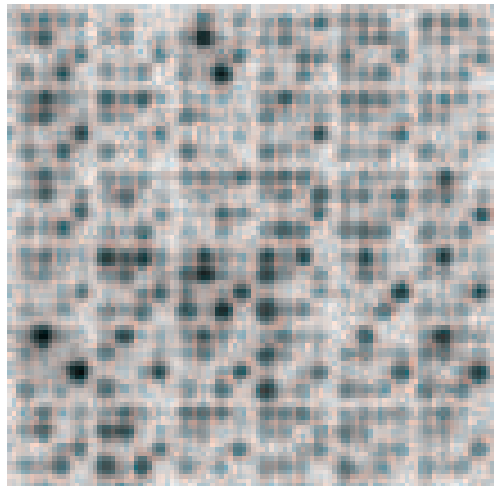


Abbildung 19: Ausschnitt aus DT40 Cre1 II

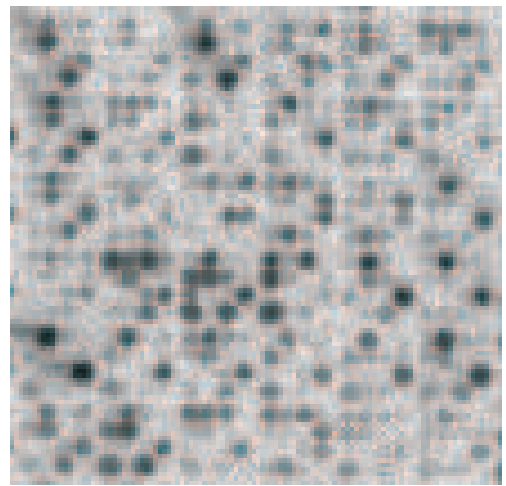


Abbildung 20: Ausschnitt aus DT40 Cre1 AID^{-/-} II

derholten Hybridisierungen in einem bestimmten Bereich variierten. Die Werte der Wildtyphybridisierung I und II, sowie der Knockouthybridisierung I variierten weniger stark, als die der Wildtyphybridisierung III und der Knockouthybridisierung II. Dies war ein erster Hinweis darauf, dass vielleicht keine differentielle Expression vorlag und die Varianz auf einem technischen Artefakt beruhte. Die Abweichung der Werte von Wildtyp I und II vs. Knockout I war geringer, als die der Werte von Wildtyp I und II vs. Wildtyp III. Die Statistik sollte die Werte herausfiltern, die in allen Fällen gleichermassen eine Expressionsdifferenz zwischen Wildtyp- und Knockoutzelle zeigte.

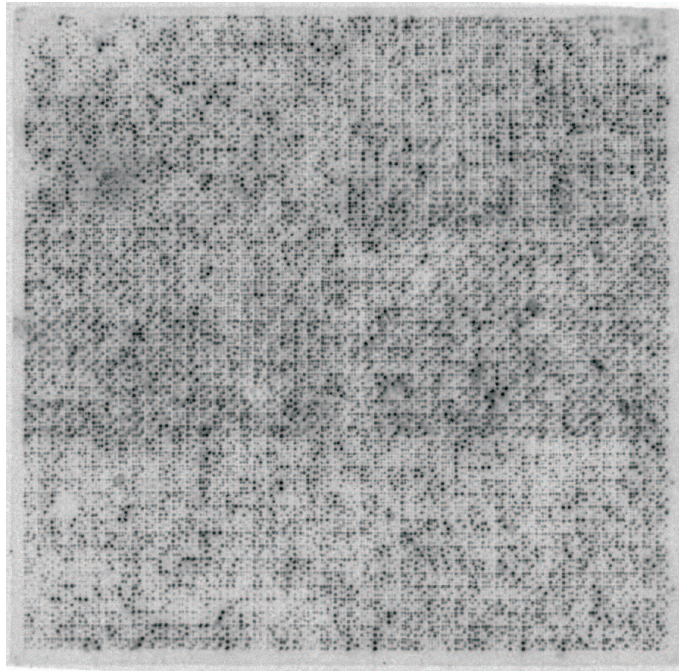


Abbildung 21: DT40 Cre1 III

- In den Abbildungen 24 und 25-29 auf den Seiten 49-52 ist die Korrelation der Duplikate eines Filters, die mit a und b bezeichnet werden, logarithmisch aufgezeigt. Diese Korrelation bot eine interne Qualitätskontrolle: Duplikate sollten von ihren Werten her möglichst gleich sein, oder sehr dicht beieinander liegen. An den Graphen kann man erkennen, dass die Duplikate von Wildtyp III und Knockout II stärker voneinander abwichen (vergleiche Abb.24-29).
- Das Programm ermöglicht es, dass Werte, die über ein bestimmtes Maß von ihrem Duplikat abwichen, hier 15 %, aus der weiteren Analyse ausgeschlossen wurden. Auch Punkte, die eine bestimmte innere Inhomogenität überschritten, wurden durch das Programm gekennzeichnet und aus der Auswertung herausgenommen.

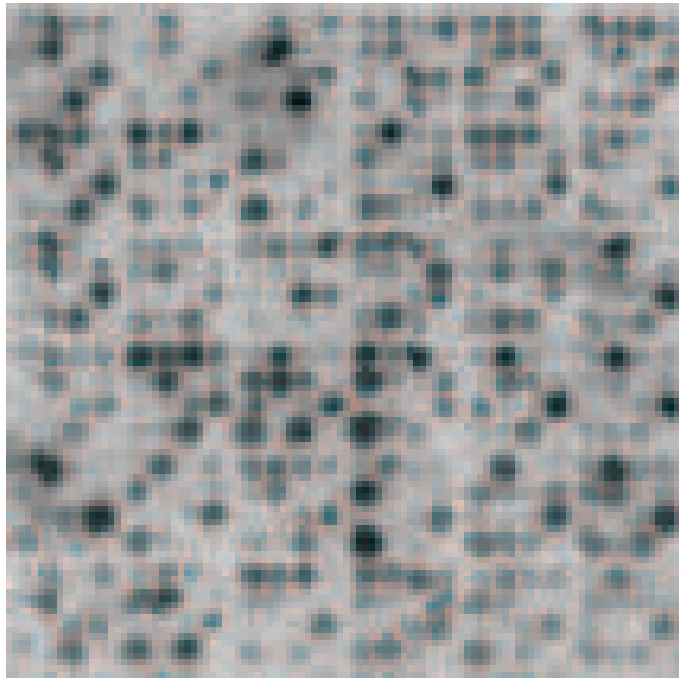


Abbildung 22: Ausschnitt aus DT40 Cre1 III

4.2.2 Vergleich der AID-positiven und AID-negativen Filter mit AIDA Array Compare

Die hintergrundkorrigierten und normalisierten Werte zweier Filter wurden mit AIDA Compare einander gegenübergestellt (siehe Wertetabelle, Abb. 30 auf S. 53). Zunächst wurden die Werte, die nach den Tabellen am höchsten differentiell exprimiert zu sein schienen verglichen. Dabei ergaben sich unterschiedliche Werte von Hybridisierung zu Hybridisierung (siehe 4.2.1). Mit den von AIDA produzierten Korrelationsgraphen konnte der Vergleich zweier Filter graphisch aufgezeigt werden. Ein („Master“)-Filter wurde einem („Client“)- Filter gegenübergestellt. Die „Master“-Daten wurden durch die x-Achse und die „Client“- Daten durch die y-Achse beschrieben. Dabei entstand eine Punktwolke, die die Korrelation der gegenübergestellten Werte aus den zwei Hybridisierungen repräsentierte. Je größer die Streuung der Werte war, desto mehr unterschieden sich die Profile voneinander. Das Maß der Streuung wurde durch die Standardabweichungen beschrieben. Die Form der Korrelationsgraphen lieferte eine Übersicht über das Ergebnis und lies folgende Aussagen zu:

- Abb. 31 zeigt die Korrelation der zwei Oligohybridisierungen A und B. Abb. 32 darunter, zeigt die Korrelation zweier Wildtyphybridisierungen. Die Oligohybridisierungen streuten erwartungsgemäß weniger als die komplexen Hybridisierungen, was an einer geringeren Standardabweichung in Abb. 31

| Index | Label | Type | Integral [PSL] | Bkg [PSL] | Integral-Bkg [PSL] | Normalized [%] | Normalized-Bkg [%] | Average-Bkg [%] | Deviation-Bkg [%] | Name |
|-------|-------|------|----------------|-----------|--------------------|----------------|--------------------|-----------------|-------------------|---------------|
| 1 | 1-1 | 10a | 273.0 | 14.1 | 259.0 | 132.9 | 132.8 | 122.8 | 8.1 | dkfz426_10p24 |
| 2 | 1-3 | 10b | 230.9 | 10.9 | 220.0 | 112.4 | 112.8 | 122.8 | 8.1 | dkfz426_10p24 |
| 3 | 1-5 | 10a | 227.9 | 13.6 | 214.3 | 110.9 | 109.9 | 102.6 | 7.1 | dkfz426_10p23 |
| 4 | 1-7 | 10b | 198.2 | 12.3 | 185.9 | 96.5 | 95.3 | 102.6 | 7.1 | dkfz426_10p23 |
| 5 | 1-8 | 16a | 82.3 | 13.8 | 68.6 | 40.1 | 35.2 | 36.1 | 2.5 | dkfz426_16p23 |
| 6 | 2-5 | 4a | 687.0 | 15.1 | 672.0 | 431.9 | 447.0 | 422.9 | 5.7 | dkfz426_4p23 |
| 7 | 2-6 | 22a | 75.8 | 13.0 | 62.8 | 36.9 | 32.2 | 32.5 | 0.9 | dkfz426_22p23 |
| 8 | 3-7 | 16b | 64.7 | 12.5 | 52.2 | 41.2 | 37.0 | 36.1 | 2.5 | dkfz426_16p23 |
| 9 | 4-5 | 4b | 792.8 | 14.9 | 777.9 | 386.0 | 398.8 | 422.9 | 5.7 | dkfz426_4p23 |
| 10 | 4-7 | 22b | 76.5 | 12.6 | 63.9 | 37.3 | 32.8 | 32.5 | 0.9 | dkfz426_22p23 |
| 11 | 1-9 | 10a | 294.7 | 15.8 | 278.9 | 143.5 | 143.0 | 137.8 | 3.8 | dkfz426_10p22 |
| 12 | 1-10 | 34a | 155.9 | 15.2 | 140.7 | 75.9 | 72.1 | 78.3 | 7.9 | dkfz426_34p22 |
| 13 | 1-11 | 10b | 273.1 | 14.5 | 258.7 | 133.0 | 132.6 | 137.8 | 3.8 | dkfz426_10p22 |
| 14 | 1-12 | 16a | 179.9 | 14.4 | 165.5 | 87.6 | 84.9 | 93.3 | 9.1 | dkfz426_16p22 |
| 15 | 2-10 | 22a | 196.8 | 16.7 | 180.1 | 95.8 | 92.4 | 102.5 | 9.9 | dkfz426_22p22 |
| 16 | 3-11 | 16b | 212.9 | 14.3 | 198.5 | 103.6 | 101.8 | 93.3 | 9.1 | dkfz426_16p22 |
| 17 | 3-12 | 34b | 178.3 | 13.4 | 164.9 | 86.8 | 84.6 | 78.3 | 7.9 | dkfz426_34p22 |
| 18 | 4-11 | 22b | 232.6 | 12.9 | 219.8 | 113.3 | 112.7 | 102.5 | 9.9 | dkfz426_22p22 |
| 19 | 1-16 | 16a | 133.7 | 13.6 | 120.1 | 65.1 | 61.6 | 60.5 | 1.7 | dkfz426_16p21 |
| 20 | 3-15 | 16b | 128.7 | 12.7 | 116.0 | 62.7 | 59.5 | 60.5 | 1.7 | dkfz426_16p21 |
| 21 | 1-17 | 10a | 103.7 | 14.6 | 89.1 | 50.5 | 45.7 | 47.6 | 3.9 | dkfz426_10p20 |
| 22 | 1-19 | 10b | 109.6 | 13.2 | 96.4 | 53.4 | 49.4 | 47.6 | 3.9 | dkfz426_10p20 |
| 23 | 2-18 | 22a | 169.9 | 14.7 | 155.2 | 82.7 | 79.6 | 82.5 | 3.5 | dkfz426_22p20 |
| 24 | 4-19 | 22b | 179.5 | 12.9 | 166.6 | 87.4 | 85.4 | 82.5 | 3.5 | dkfz426_22p20 |
| 25 | 1-21 | 10a | 143.1 | 14.3 | 128.7 | 69.6 | 66.0 | 65.8 | 0.3 | dkfz426_10p19 |
| 26 | 1-23 | 10b | 140.2 | 12.4 | 127.8 | 68.3 | 65.5 | 65.8 | 0.3 | dkfz426_10p19 |
| 27 | 2-21 | 4a | 970.0 | 16.5 | 953.6 | 472.3 | 488.9 | 511.2 | 4.4 | dkfz426_4p19 |
| 28 | 1-25 | 10a | 156.8 | 13.7 | 143.1 | 76.3 | 73.3 | 72.7 | 0.9 | dkfz426_10p18 |
| 29 | 1-26 | 34a | 117.0 | 13.2 | 103.8 | 57.0 | 53.2 | 49.6 | 7.4 | dkfz426_34p18 |
| 30 | 1-27 | 10b | 152.9 | 12.5 | 140.4 | 74.5 | 72.0 | 72.7 | 0.9 | dkfz426_10p18 |
| 31 | 1-28 | 16a | 195.4 | 13.5 | 182.0 | 95.1 | 93.3 | 101.4 | 8.0 | dkfz426_16p18 |
| 32 | 2-26 | 22a | 60.9 | 13.4 | 47.5 | 29.7 | 24.4 | 25.0 | 2.6 | dkfz426_22p18 |
| 33 | 3-27 | 16b | 226.4 | 12.7 | 213.7 | 110.2 | 109.6 | 101.4 | 8.0 | dkfz426_16p18 |
| 34 | 3-28 | 34b | 102.9 | 13.4 | 89.5 | 50.1 | 45.9 | 49.6 | 7.4 | dkfz426_34p18 |
| 35 | 4-27 | 22b | 62.7 | 12.7 | 50.1 | 30.5 | 25.7 | 25.0 | 2.6 | dkfz426_22p18 |

Abbildung 23: AIDA Metrix. Ausschnitt aus der Wertetabelle von DT40 Cre1. „Index“ = Laufende Numerierung. „Label“ = Filterkoordinate. „Type“ = Klonplattenummer und Bezeichnung der Duplikate durch a und b. „Integral“ = Rohwert. „Integral-Bkg“ = hintergrundkorrigierter Rohwert. „Normalized“ = normalisierter Rohwert. „Normalized-Bkg“ = hintergrundkorrigierter normalisierter Rohwert. „Average“ = Mittelwert zweier Duplikate. „Deviation“ = Abweichung vom Mittelwert. „Name“ = EST Name aus dem „bursal EST library“.

abzulesen ist.

- In Abb. 33 und 34 auf Seite 55 (DT40 Cre1 vs. DT40 Cre1 AID^{-/-}) entsprachen die Standardabweichungen ungefähr der in Abb. 32 auf Seite 54 (DT40 Cre1 vs. DT40 Cre1). Dies zeigt, dass die Transkriptionsprofile von Wildtyp- und Knockoutzelle nicht eindeutig voneinander abwichen, woraus man schliessen kann, dass AID kein globaler Transkriptionsfaktor sein kann, der eine deutliche Änderung des Profils hervorgerufen hätte.

Duplicate Correlation

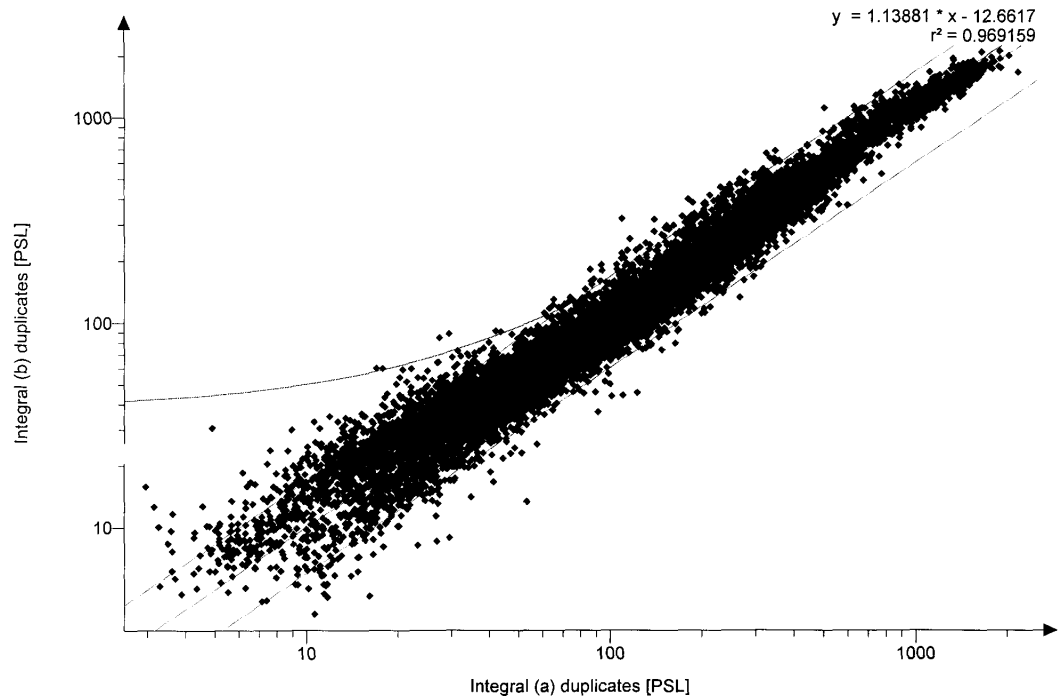


Abbildung 24: Korrelation der Duplikate der Oligohybridisierung A. Logarithmischer Auftrag. Die beiden äußeren Linien repräsentieren die festgelegte Schwelle von 15 %, bis zu der Duplikate voneinander abweichen können, ohne aus der Auswertung herausgenommen zu werden. Durch die angegebene Gleichung wird der Korrelationsgraph beschrieben. „ r^2 “ ist der Korrelationskoeffizient

- Die Abb. 35- 38 zeigten stärkere Abweichungen. DT40 Cre1 III trug zu einer größeren Abweichung bei. Die RNA von DT40 Cre1 III stammte aus einer anderen Isolation, als die von DT40 Cre1 I und II. Ein geringes Abweichen der Zellkonditionen schien sich deutlich auf das Ergebnis auszuwirken (siehe 4.1.1). Auch DT40 Cre1 AID^{-/-} II trug zu einer grösseren Abweichung bei. Die stärkeren Duplikatabweichungen von DT40 Cre1 III und DT40 Cre1 AID^{-/-} II (siehe Abb. 25- 29) zeigten, dass hier während der Experimentierschritte (RNA-Isolation, Auftauvorgang, und Hybridisierung) mehr unspezifische Werte erzeugt wurden.
- Man kann auch hier feststellen, dass die Streuung im Bereich niedriger Expressionswerte grösser war, so dass in diesem Bereich des Graphen die Werte sehr viel stärker divergierten (siehe dazu auch Abschnitt 3.2.15 auf Seite 34 und vergleiche Korrelationsgraphen auf S. 54- S. 57) Diese Tatsache erschwert es, differentiell exprimierte Gene im niedrigen Expressionsbereich zu detektieren.

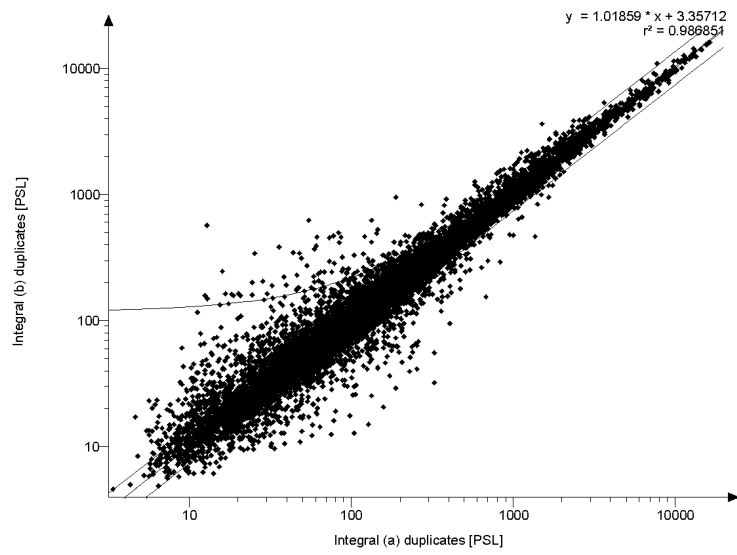


Abbildung 25: Duplikatkorrelation DT40 Cre1 I

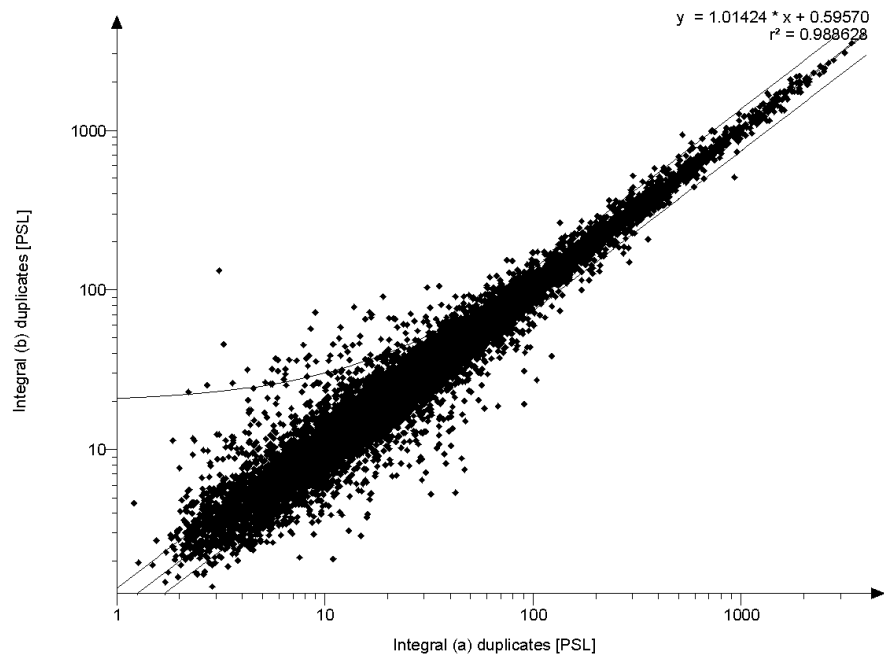


Abbildung 26: Duplikatkorrelation DT40 Cre1 II

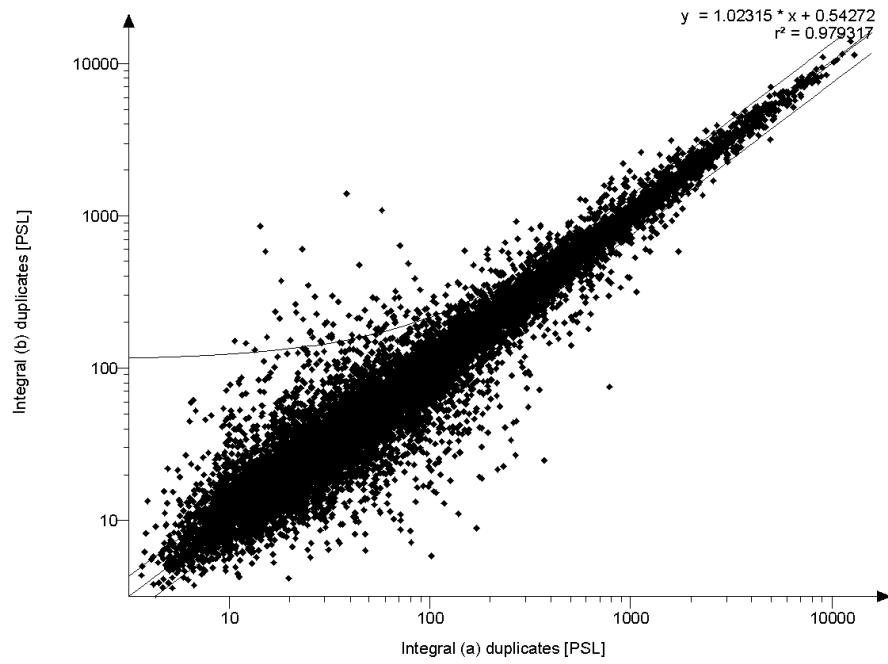


Abbildung 27: Duplikatkorrelation DT40 Cre1 III

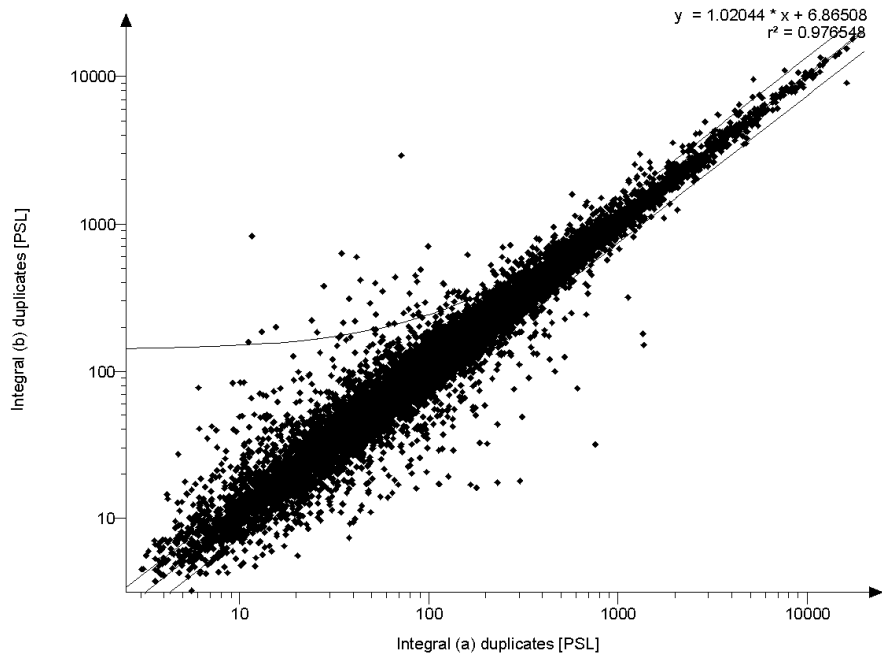


Abbildung 28: Duplikatkorrelation DT40 Cre1 AID^{-/-}I

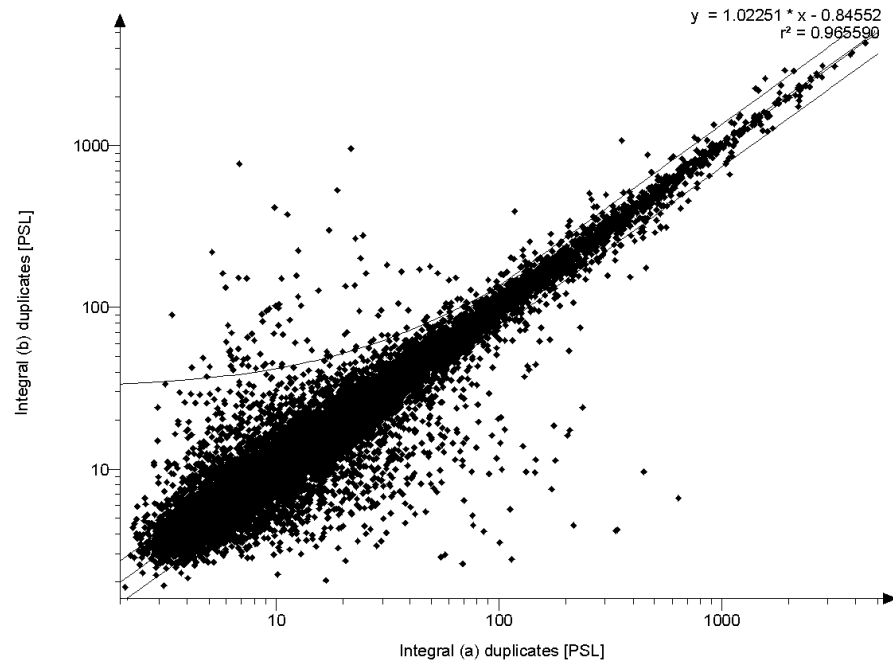


Abbildung 29: Duplikatkorrelation DT40 Cre1 AID^{-/-} II

Nach den Abb. 33 und 34 auf Seite 55 (DT40 Cre1 vs. DT40 Cre1 AID^{-/-}) und 32 auf Seite 54 (DT40 Cre1 vs. DT40 Cre1) konnte keine Veränderung des Transkriptionsprofils festgestellt werden, so wie man es von einem globalen Transkriptionsfaktor erwarten würde. Auch wies die stärkere Abweichung der zwei Wildtphybridisierungen I vs. III oder II vs. III, gegenüber der geringeren Abweichung von Wildtyp- vs. Knockoutzelle darauf hin, dass keine differentielle Expression vorlag.

Mit der Darstellung der Werte durch die Korrelationsgraphen lassen sich keine genauen Aussagen über eine differentielle Expression treffen. Durch die statistische Integration aller Hybridisierungen sollten die Gene, bzw. ESTs gefunden werden, die in allen durchgeführten Hybridisierungen eine hohe Expressionsdifferenz zeigten. Dazu wurde die im Materialteil beschriebene Statistik angewandt (siehe auch 4.3 und Tabellen 2 und 3 auf S. 60-68).

| Index | Label | Type | Master Normalized-Bkg [%] | Client Normalized-Bkg [%] | Ratio Normalized-Bkg | Name |
|-------|---------|------|---------------------------|---------------------------|----------------------|---------------|
| 1 | 170-8 | 29a | -0.4 | 52.5 | 100.00 | dkfz426_29f23 |
| 2 | 169-7 | 11b | 0.0 | 15.4 | 100.00 | dkfz426_11f23 |
| 3 | 169-5 | 11a | 0.2 | 12.3 | 56.03 | dkfz426_11f23 |
| 4 | 169-6 | 35a | 1.9 | 48.0 | 25.95 | dkfz426_35f23 |
| 5 | 183-92 | 35b | 0.6 | 13.9 | 23.11 | dkfz426_35c2 |
| 6 | 172-6 | 29b | 3.4 | 55.8 | 16.33 | dkfz426_29f23 |
| 7 | 169-4 | 17a | 0.7 | 10.5 | 16.05 | dkfz426_17f24 |
| 8 | 173-6 | 35a | 1.9 | 29.2 | 15.13 | dkfz426_35e23 |
| 9 | 187-56 | 35b | 0.9 | 13.3 | 14.19 | dkfz426_35b11 |
| 10 | 176-38 | 29b | 17.3 | 196.3 | 11.32 | dkfz426_29e15 |
| 11 | 169-8 | 17a | 55.5 | 599.9 | 10.81 | dkfz426_17f23 |
| 12 | 173-7 | 11b | 4.4 | 44.9 | 10.18 | dkfz426_11e23 |
| 13 | 172-7 | 23b | 80.5 | 739.7 | 9.19 | dkfz426_23f23 |
| 14 | 170-10 | 23a | 1.9 | 15.7 | 8.08 | dkfz426_23f22 |
| 15 | 172-5 | 5b | 9.9 | 75.4 | 7.64 | dkfz426_5f23 |
| 16 | 143-96 | 35b | 1.6 | 11.8 | 7.26 | dkfz426_35m1 |
| 17 | 173-8 | 17a | 5.9 | 43.0 | 7.26 | dkfz426_17e23 |
| 18 | 170-4 | 29a | 5.4 | 36.1 | 6.62 | dkfz426_29f24 |
| 19 | 144-95 | 23b | 2.3 | 15.0 | 6.48 | dkfz426_23m1 |
| 20 | 170-5 | 5a | 9.4 | 59.4 | 6.34 | dkfz426_5f23 |
| 21 | 148-131 | 27b | 12.6 | 78.4 | 6.21 | dkfz426_27l16 |
| 22 | 168-7 | 23b | 11.8 | 71.3 | 6.17 | dkfz426_23g23 |
| 23 | 173-10 | 35a | 4.5 | 26.7 | 5.99 | dkfz426_35e22 |
| 24 | 175-8 | 35b | 5.2 | 29.9 | 5.78 | dkfz426_35e23 |
| 25 | 181-70 | 35a | 4.0 | 22.6 | 5.64 | dkfz426_35c7 |
| 26 | 174-4 | 29a | 3.6 | 20.4 | 5.62 | dkfz426_29e24 |
| 27 | 169-9 | 11a | 6.3 | 34.9 | 5.59 | dkfz426_11f22 |
| 28 | 173-2 | 35a | 2.7 | 14.9 | 5.49 | dkfz426_35e24 |
| 29 | 175-4 | 35b | 3.0 | 16.3 | 5.48 | dkfz426_35e24 |
| 30 | 191-64 | 35b | 4.8 | 25.8 | 5.41 | dkfz426_35a9 |
| 31 | 179-72 | 35b | 2.0 | 10.3 | 5.17 | dkfz426_35d7 |
| 32 | 179-68 | 35b | 3.0 | 15.6 | 5.15 | dkfz426_35d8 |
| 33 | 192-75 | 23b | 7.3 | 37.1 | 5.08 | dkfz426_23a6 |
| 34 | 173-5 | 11a | 9.2 | 46.6 | 5.07 | dkfz426_11e23 |
| 35 | 171-4 | 35b | 56.5 | 284.3 | 5.03 | dkfz426_35f24 |

Abbildung 30: AIDA Compare. Ausschnitt aus der Wertetabelle von DT40 Cre1 vs. AID^{-/-}. „Index“ = Laufende Numerierung. „Label“ = Filterkoordinate. „Type“ = Klonplattennummer und Bezeichnung der Duplikate durch a und b. „Master Normalized-Bkg“ = DT40 Cre1. „Client Normalized-Bkg“ = AID^{-/-}. „Ratio Normalized-Bkg“ = Expressionsverhältnis. „Name“ = EST Name aus dem „bursal EST library“.

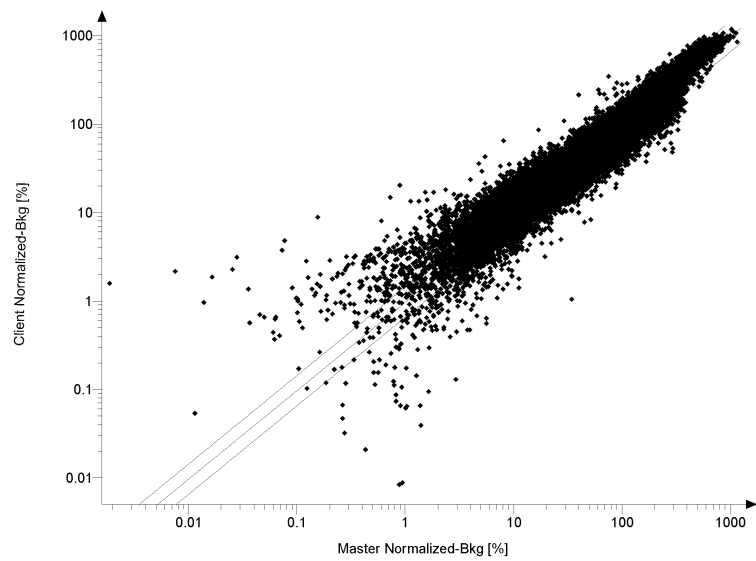


Abbildung 31: Oligohybridisierungen A („Master“) und B („Client“) Die Standardabweichungen ober- und unterhalb der Korrelationsgeraden sind angegeben ($SD=1,4/-1,5$)

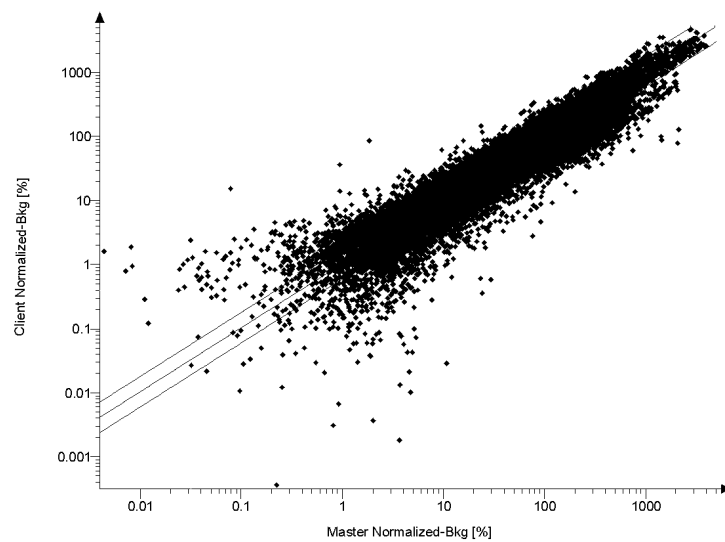


Abbildung 32: Korrelation: DT40 Cre1 I („Master“) und DT40 Cre1 II („Client“)($SD=1,8/-1,7$)

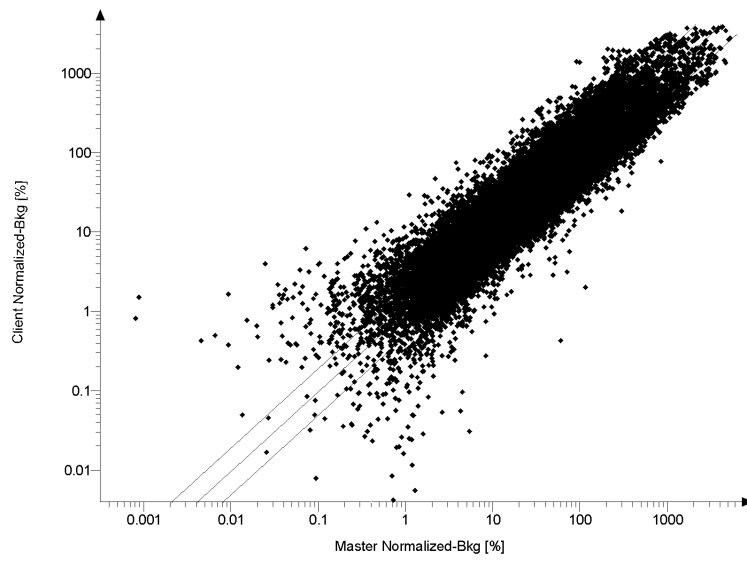


Abbildung 33: Korrelation: AID-/- I („Master“) und DT40 Cre1 I („Client“)(SD=1,9/-2,0)

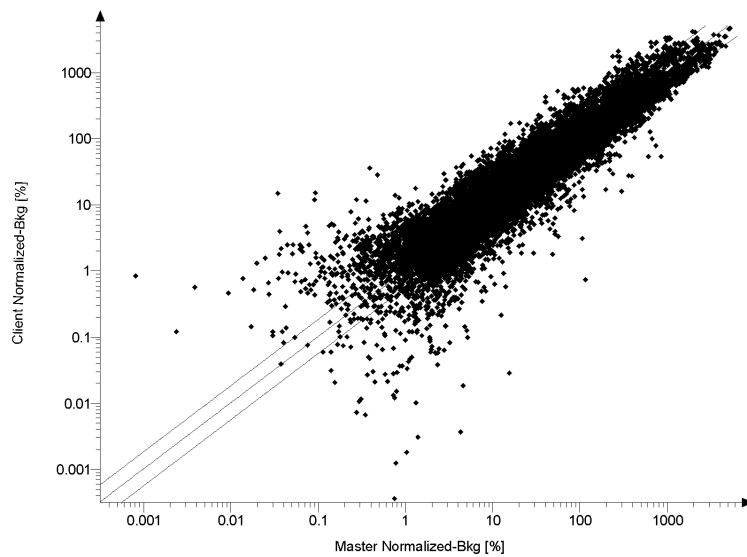


Abbildung 34: Korrelation: AID-/- I („Master“) und DT40 Cre1 II („Client“)(SD=1,9/-1,7)

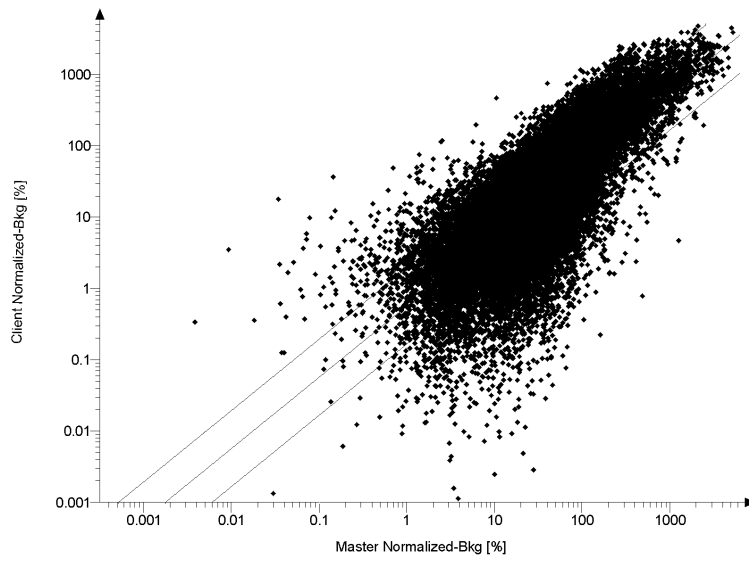


Abbildung 35: Korrelation: AID-/- I („Master“) und DT40 Cre1 III („Client“)(SD=1,9/-6,0)

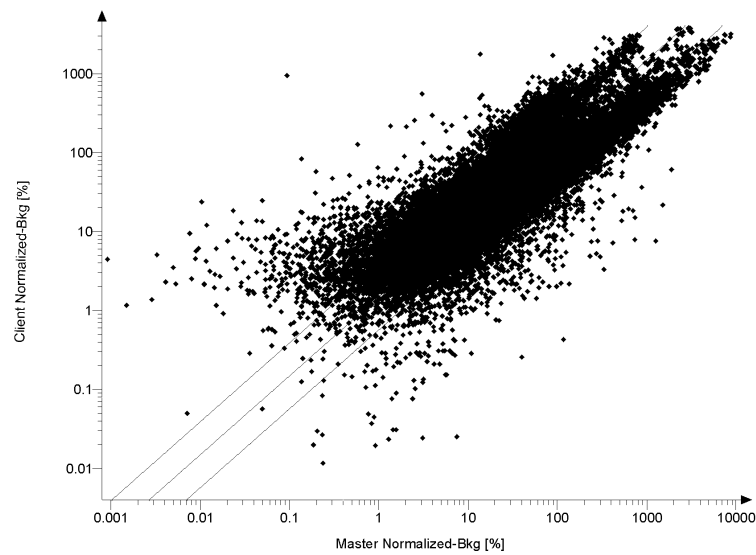


Abbildung 36: Korrelation: AID-/- II („Master“) und DT40 Cre1 I („Client“)(SD=3,8/-1,8)

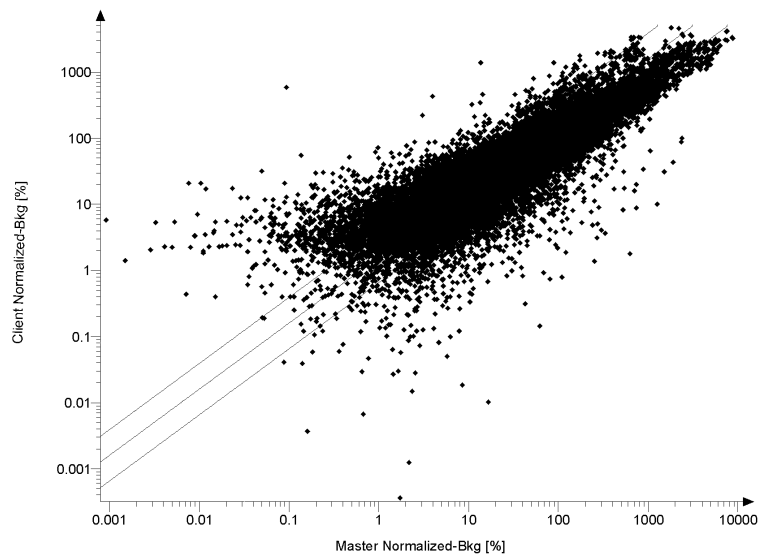


Abbildung 37: Korrelation: AID-/- II („Master“) und DT40 Cre1 II („Client“) (SD=3,9/-1,5)

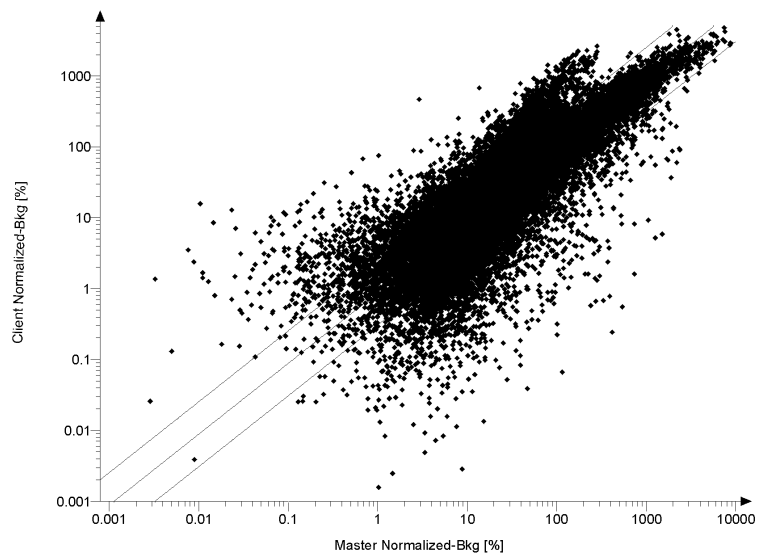


Abbildung 38: Korrelation: AID-/- II („Master“) und DT40 Cre1 III („Client“) (SD=2,5/-3,2)

4.3 Statistik und Datenbanksuche

In diesem Abschnitt werden die Gene, die nach Anwendung der zwei statistischen Methoden zwischen AID^{-/-} und DT40 Cre1 als differentiell exprimiert definiert werden, aufgeführt:

- Die „Normalized-background“-Werte, die das AIDA-Matrix-Programm erzeugte, wurden zunächst in Tabellen einander gegenübergestellt. Dabei verglichen beide Statistiken einzelne Expressionsdifferenzwerte mit der Variabilität dieser Werte aus wiederholten Hybridisierungen. Auch die 2 x 2 Werte aus den Oligo-Hybridisierungen wurden einbezogen. Die inhomogenen Werte und die Werte, die zu stark von ihrem Duplikat abwichen, wurden ausgenommen.
- t-Test-Statistik: Hier wurden 52 Gene, die in AID^{-/-} schwächer exprimiert sind als in DT40 Cre1 beschrieben, die mit einer Wahrscheinlichkeit P von 0,01 als zufällig identifiziert gelten, also definitionsgemäß mit einer hohen Wahrscheinlichkeit differentiell exprimiert sind. Die Basis für diese Werte bildeten die „Normalized-background-Werte“ aus AIDA Matrix, die in Tabellen einander gegenübergestellt und verglichen wurden (siehe auch 3.2.18 auf Seite 36).
- „Significance analysis of microarrays“ (SAM)-Statistik: Je höher der SAM-Score, der durch SAM jedem Gen zugeordnet wird, desto wahrscheinlicher ist, dass es sich um eine differentielle Expression handelt. Hier wurde vom Statistiker ein Score von $+/- 6,8$ als Schwellenwert festgelegt. Alle Werte, die zwischen $-6,8$ und $+6,8$ lagen, fielen in die schon beschriebene „false discovery rate“. Hiernach wurden 61 Gene beschrieben, die in AID^{-/-} schwächer exprimiert wurden, als in DT40 Cre1 (siehe dazu Abb. 8, Seite 38). Die in der Tabelle aufgelisteten Kandidaten wurden weiterhin nach ihrem „fold change“ beurteilt, der die differentielle Expression als Verhältnis beschreibt. Je niedriger der Wert ist, desto größer ist der Expressionsunterschied (ein Wert von 0,1 entspricht einem 10fachen Expressionsunterschied zugunsten der Wildtypzelle). Expressionsunterschied und SAM-Score verhalten sich proportional zueinander.
- Die meisten Rohwerte, die durch AIDA gewonnen wurden, fielen aus der Auswertung hinaus: Von den ursprünglich über 14 000 aufgetragenen ESTs konnten schließlich etwas über 200 ESTs durch die statistischen Methoden als signifikant differentiell exprimiert definiert werden, also ca. 1,5 %. Es fiel sofort auf, dass es keine signifikanten ESTs gab, die durch beide Methoden (t-Test- und SAM-Statistik) detektiert wurden. Nach beiden Statistiken wurden etwa eben so viele hoch- wie herunterregulierte Gene in der Knockoutzelle gefunden. Nach SAM wurden 110 differentiell exprimierte Gene beschrieben, deren Score über einem willkürlich festgelegten Wert

von +/- 6,8 lag, von denen 61 in AID-/- schwächer exprimiert waren. Mittels t-Test-Statistik wurden 100 ESTs benannt, von denen 52 ESTs in AID^{-/-} schwächer exprimiert waren als in DT40 Cre1. Hier werden nur die schwächer exprimierten Gene aufgeführt, also die Gene, deren Expression vermutlich durch AID induziert wird.

Auch wenn es keine gemeinsamen, von beiden Statistiken als signifikant bezeichneten Gene gab, wurde zu den ESTs, die nach den Statistiken in AID^{-/-} niedriger exprimiert waren als in DT40 Cre1, eine Homologensuche in der „bursal EST Database“ durchgeführt, um festzustellen, ob die potentiell differentiell exprimierten ESTs bekannte Teilsequenzen eines Gens oder Proteins sind, die an Reparatur oder Rekombination beteiligt sind. Zunächst wurden sie in tabellarischer Form alle aufgeführt und benannt, soweit dies nach der „bursal EST Database“ möglich war (siehe Tabellen 2 und 3 auf S. 60-68). Im nächsten Abschnitt 4.4 werden einige ESTs auf ihre Funktion hin detaillierter beschrieben.

Tabelle 2: ESTs, die nach t-Teststatistik in AID-/- Zellen niedrig, und in DT 40 Zellen hoch exprimiert werden. ESTs zu denen Annotationen gefunden wurden, sind *hervorgehoben*. Ihr „Score“ (nicht zu verwechseln mit dem SAM-Score), der ein Maß für die Ähnlichkeit zur annotierten Sequenz darstellt, sollte nach der „swiss protein“ Datenbank („blsp“) mindestens 50 und nach der „gene bank“ („blge“) mindestens 100 betragen, um mit hoher Wahrscheinlichkeit signifikant zu sein. „*“ kennzeichnet Sequenzen, die nur einen niedrigen Score haben, so dass Annotationen hier nicht aufgeführt wurden. Wie man sehen kann, gibt es in der Liste auch eine Reihe von nicht sequenzierten ESTs. Aus der Datenbank lässt sich auch entnehmen, ob ein EST zu einem „Cluster“ aus mehreren EST-Sequenzen gehört. In dem Fall kann auch die Annotation eines Cluster-Mitgliedes interessant sein. Datenbank: <http://swallow.gsf.de/dt40Est.html>.

| EST Name aus DKFZ 426 _ Bibliothek | Annotation | Score | Quelle | Cluster Sequenzen |
|------------------------------------|-------------------------------------------------------|----------|-----------|-------------------|
| 27g7 | <i>mitochondrial</i> | 479/1588 | blsp/blge | 18 |
| 8m5 | <i>mitochondrial</i> | 240/472 | blsp/blge | 4 |
| 17f12 | <i>β-Aktin</i> | 416/1376 | blsp/blge | > 20 |
| 12o12 | <i>Hühner Tandem PH- Domäne beinhaltendes Protein</i> | 352/1285 | blsp/blge | 18 |
| 4f10 | <i>PAP1 m RNA und Protein</i> | 407/591 | blsp/blge | 6 |
| 26d21 | <i>untranslatierte Region bindendes Protein</i> | 348/999 | blsp/blge | 3 |
| 26a7 | <i>ladungsabhängiger Ionenkanal</i> | 137/579 | blsp/blge | 7 |
| 15l9 | <i>ribosomales Protein</i> | 347/636 | blsp/blge | 7 |
| 7d11 | <i>Ribonuklease H1</i> | 147/547 | blsp/blge | |
| 3f15 | <i>Lipoprotein bindendes Protein/Vigilin</i> | 278/924 | blsp/blge | 1 |
| 30b22 | <i>Phosphatidylethanolamin bindendes Protein</i> | 333/127 | blsp/blge | 2 |
| 13b21 | G10 Transkript und Protein | 312/280 | blsp/blge | |
| 25k13 | humanes KIAA0368 Genprodukt | 321/113 | blsp/blge | |
| 36c14 | <i>Caspase 3 mRNA</i> | 270 | blge | 5 |
| 10m2 | <i>Translationselongationsfaktor</i> | 186/101 | blsp/blsp | |
| 9l15 | <i>mRNA bindendes PAI 1 Protein</i> | 133 | blge | |
| 21d8 | <i>Ig μSchwerkette</i> | 81/121 | blsp/blge | > 20 |

| EST Name aus DKFZ 426_ Bibliothek | Annotation | Score | Quelle | Cluster Sequen- zen |
|-----------------------------------------|------------------------------------------------------------------------------------------------------------------|-------|-----------|---------------------------|
| 12b6 | <i>Cytochrom C Oxydase Un- tereinheit</i> | 90 | blsp | |
| 17o1 | <i>v-myb Onkogen Homolog mRNA</i> | 99 | blge | > 20 |
| 12l20 | <i>Homolog zu „human 54 kDa nuclear non POU-domain- containing octamer binding protein“ („NonO“)</i> | 82 | blsp | 8 |
| 8h8 | <i>Hsp 70 bindendes Protein</i> | 81 | blsp | |
| 4f18 | * | 30/46 | blsp/blge | |
| 31f5 | * | 30/50 | blsp/blge | |
| 17h10 | * | | | 2 |
| 10g15 | * | | | > 20 |
| 30n2 | * | 0/42 | blsp/blge | 16 |
| 6b21 | * | 32/46 | blsp/blge | |
| 30f12 | * | 33/44 | blsp/blge | |
| 25d17 | * | 33/44 | blsp/blge | 1 |
| 2a12 | * | 32/44 | blsp/blge | |
| 18c8 | * | 31/42 | blsp/blge | 2 |
| 30f10 | * | 35/44 | blsp/blge | |
| 36p11 | * | 34/42 | blsp/blge | |
| 3h19 | * | 0/42 | blsp/blge | |
| 12l1 | keine Sequenzinformation | | | |
| 32a15 | keine Sequenzinformation | | | |
| 31n7 | keine Sequenzinformation | | | |
| 37d17 | keine Sequenzinformation | | | |
| 7i12 | keine Sequenzinformation | | | |
| 18e8 | keine Sequenzinformation | | | |
| 18n11 | keine Sequenzinformation | | | |
| 5a1 | keine Sequenzinformation | | | |
| 33d13 | keine Sequenzinformation | | | |
| 34a8 | keine Sequenzinformation | | | |
| 33l3 | keine Sequenzinformation | | | |
| 3i6 | keine Sequenzinformation | | | |
| 5o4 | keine Sequenzinformation | | | |
| 20e10 | keine Sequenzinformation | | | |
| 34n15 | keine Sequenzinformation | | | |
| 30d5 | keine Sequenzinformation | | | |
| 1i13 | keine Sequenzinformation | | | |

| EST Name aus DKFZ 426_ Bibliothek | Annotation | Score | Quelle | Cluster Sequen- zen |
|-----------------------------------------|--------------------------|-------|--------|---------------------------|
| 7c11 | keine Sequenzinformation | | | |

Tabelle 3: ESTs, die nach der SAM Methode in AID -/- Zellen niedrig, und in DT40 Cre1 Zellen hoch exprimiert werden. Die Tabelle ist nach dem *Fold change* geordnet, der das Maß der differentiellen Expression beschreibt. Je kleiner das Verhältnis, desto größer der Unterschied zwischen Wildtyp- und k.o.- Zelle. ESTs zu denen Annotationen gefunden wurden, sind *hervorgehoben*. Ihr „Score“(nicht zu verwechseln mit dem SAM-Score), der hier ein Maß für die Ähnlichkeit zur annotierten Sequenz darstellt, sollte nach der „swiss protein“ Datenbank („blsp“) mindestens 50 und nach der „gene bank“ („blge“) mindestens 100 betragen, um mit hoher Wahrscheinlichkeit signifikant zu sein. Es gibt viele nicht sequenzierte ESTs. „*“ bedeutet, daß die Annotation einen zu niedrigen Score hat und daher nicht aussagekräftig ist. Aus der Datenbank lässt sich auch entnehmen, ob ein EST zu einem „Cluster“ aus mehreren EST-Sequenzen gehört. In dem Fall kann auch die Annotation eines Cluster-Mitgliedes interessant sein.

| EST aus 426_ Bibli- thek | Name DKFZ Biblio- thek | fold change | Annotation | Score | Quelle | Cluster Sequen- zen |
|--------------------------------------|---------------------------------|----------------|-----------------------------------------------------------------|---------|-----------|---------------------------|
| 14l11 | | 0,1 | <i>Aminopeptidase B</i> | 204/75 | blsp/blge | |
| 24m4 | | 0,1 | * | 32/42 | blsp/blge | |
| 38l8 | | 0,1 | keine Sequenzinformation | | | |
| 27o3 | | 0,2 | <i>Hühner Leichtkettengen</i> | 166/529 | blsp/blge | > 20 |
| 12f18 | | 0,2 | <i>Hühner Ig Schwerketten μ konstante Region</i> | 283/577 | blsp/blge | > 20 |
| 6h7 | | 0,2 | <i>Hühner Aktinin</i> | 72/210 | blsp/blge | |
| 34g24 | | 0,2 | <i>Ähnlichkeit zu Hydrocarbonrezeptortranslocator</i> | 254 | blge | |
| 24f19 | | 0,2 | <i>IL-2 Rezeptor γ Vorläufer</i> | 175 | blsp | 6 |
| 24l6 | | 0,2 | * | 32/44 | blsp/blge | |
| 32a11 | | 0,2 | * | 35/44 | blsp/blge | 1 |
| 38a19 | | 0,2 | keine Sequenzinformation | | | |
| 4g19 | | 0,2 | keine Sequenzinformation | | | |
| 34e23 | | 0,2 | keine Sequenzinformation | | | |
| 4c21 | | 0,2 | keine Sequenzinformation | | | |
| 4c23 | | 0,2 | keine Sequenzinformation | | | |
| 12n1 | | 0,3 | <i>Hühner Gen für Zentromer Protein</i> | 278 | blge | 1 |
| 4d21 | | 0,3 | <i>ähnlich humanem hypothetischen Protein</i> | 97/129 | blsp/blge | 1 |
| 4g24 | | 0,3 | <i>Homolog zu humanem Karyopherin/Importin</i> | 317/363 | blsp/blge | |
| 4f19 | | 0,3 | <i>hypothetisches Zink Finger Protein</i> | 32 | blsp | |
| 30i22 | | 0,3 | <i>Hühner μ konstante Region</i> | 315/890 | blsp/blge | > 20 |

| EST aus 426_ thek | Name DKFZ Biblio- thek | fold change | Annotation | Score | Quelle | Cluster Sequen- zen |
|----------------------------|---------------------------------|----------------|------------------------------------------------------------------------|----------|-----------|---------------------------|
| 6k19 | | 0,3 | <i>dbpB/B-1, DNA binden- des Protein</i> | 237/1106 | blsp/blge | 13 |
| 18o1 | | 0,3 | homolog zu humanem KIAA0349 Protein | 119/153 | blsp/blge | |
| 34h20 | | 0,3 | * | 32/42 | blsp/blge | |
| 18n8 | | 0,3 | * | 32/50 | blsp/blge | > 20 |
| 4f18 | | 0,3 | * | 30/46 | blsp/blge | |
| 22c16 | | 0,3 | * | 32/44 | blsp/blge | 1 |
| 12p11 | | 0,3 | * | 35/46 | blsp/blge | |
| 2a18 | | 0,3 | keine Sequenzinformation | | | |
| 38110 | | 0,3 | keine Sequenzinformation | | | |
| 3d18 | | 0,3 | keine Sequenzinformation | | | |
| 12o9 | | 0,3 | keine Sequenzinformation | | | |
| 32a16 | | 0,3 | keine Sequenzinformation | | | |
| 12o14 | | 0,3 | keine Sequenzinformation | | | |
| 18i2 | | 0,4 | <i>Hämoflavoprotein</i> | 267 | blsp | 1 |
| 4c22 | | 0,4 | <i>BRCA1 assoziiertes Pro- tein</i> | 165 | blge | 1 |
| 30i18 | | 0,4 | <i>humanes hypothetisches Gen</i> | 61 | blge | 1 |
| 2g1 | | 0,4 | <i>humanes Transformer2α Protein</i> | 31 | blsp | 15 |
| 21m7 | | 0,4 | <i>homolog zu muriner Prolyl-4-hydroxylaseα 2 Untereinheit</i> | 121 | blsp | 2 |
| 18p15 | | 0,4 | homolog zu cDNA aus Ri- ckenbücherei | 144 | blsp | |
| 24a1 | | 0,4 | * | 0/42 | blsp/blge | |
| 18o11 | | 0,4 | * | 31/46 | blsp/blge | 1 |
| 18n1 | | 0,4 | * | 35/50 | blsp/blge | |
| 36g7 | | 0,4 | * | 0/40 | blsp/blge | |
| 33e19 | | 0,4 | keine Sequenzinformation | | | |
| 6i4 | | 0,4 | keine Sequenzinformation | | | |
| 6a1 | | 0,4 | keine Sequenzinformation | | | |
| 32a18 | | 0,4 | keine Sequenzinformation | | | |
| 34k24 | | 0,5 | <i>Hühner Y-Box Protein</i> | 662 | blge | 13 |
| 8j7 | | 0,5 | <i>homolog zu hu- maner Protein- Tyrosinphosphatase</i> | 75 | blsp | 1 |

| EST aus 426_ thek | Name DKFZ Biblio- thek | fold change | Annotation | Score | Quelle | Cluster Sequen- zen |
|----------------------------|---------------------------------|----------------|-------------------------------------|---------|-----------|---------------------------|
| 25n24 | | 0,5 | <i>KDEL Rezeptor</i> | 370 | blsp | 8 |
| 18p9 | | 0,5 | <i>hypothetisches Protein</i> | 150/184 | blsp/blge | 1 |
| 26k13 | | 0,5 | Hühner-Klon 126p17 | 484 | blge | 1 |
| 18o9 | | 0,5 | * | 31/42 | blsp/blge | 1 |
| 36p2 | | 0,5 | keine Sequenzinformation | | | |
| 30i5 | | 0,5 | keine Sequenzinformation | | | |
| 18b1 | | 0,6 | * | 30/44 | blsp/blge | |
| 30e7 | | 0,6 | keine Sequenzinformation | | | |
| 36n3 | | 0,6 | keine Sequenzinformation | | | |
| 30b21 | | 0,7 | Homologien zu PNAS und Proteinen | 77/329 | blsp/blge | 11 |
| 18j16 | | 0,8 | Mitochondrium | 1467 | blge | > 20 |
| 18n15 | | 0,8 | keine Sequenzinformation | | | |

4.4 Analyse der ESTs mit Hilfe der Datenbank „dkfz 426“

In diesem Abschnitt werden die durch Statistik ermittelten differentiell exprimierten ESTs, die mit Hilfe der Datenbank dkfz 426 im vorigen Abschnitt 4.3 benannt werden konnten und die möglicherweise an Rekombination und Reparatur beteiligt sind, ausführlicher beschrieben. Dabei werden zunächst die nach t-Test und anschließend die nach SAM-Statistik als differentiell exprimiert definierten ESTs besprochen.

1. t-Test:

- (26d21) Untranslatierte Region „UTR“ bindende Proteine können in 5'-UTR bindende und 3'UTR bindende Proteine unterteilt werden, die durch ihre Bindung an die entsprechende Region an der Regulation der Genexpression beteiligt sind [32]. Ein 3' UTR bindendes Protein kann z.B. die Poly(A) Polymerase binden und damit die Translation beeinflussen. Auch hier könnten Reparatur- und Rekombinationsproteine betroffen sein.
- (12l20) „human 54 kDa nuclear non POU-domain- containing octamer binding protein“. „NonO“ ist ein nukleinsäurebindendes Protein mit DNA- und RNA- bindender Eigenschaft. Außerdem hat es eine Sequenzhomologie mit PSF, einem humanem Splicingfaktor. Desweiteren ist es an Protein-Protein- Interaktionen beteiligt und kann die Bindung von Transkriptionsfaktoren an ihre Bindungsstellen erleichtern, z.B. an den Ig V_H Promotor [85]. „NonO“ induziert die Bindung des Ku-Komplexes an die DNA und ist damit an der NHEJ-Reparatur von Doppelstrangbrüchen, sowie V(D)J-Rekombination und CSR beteiligt. Siehe auch Abschnitt 2.1.1 auf Seite 7, Abb. 4 auf Seite 11 und Abschnitt 2.1.5 auf Seite 11. Die geringere Expression von NonO in AID defizienten Zellen kann auf eine geringere NHEJ-Reparaturaktivität schließen lassen. Somatische Hypermutation und Genkonversion werden wahrscheinlich vor allem durch homologe Rekombination repariert, so dass sich nicht eindeutig sagen läßt, ob eine geringere Bindung des Ku-Komplexes sich auch hier auswirkt. Allerdings zeigte sich in einer neueren Arbeit von [25], dass somatische Hypermutation in der G₂-Phase stattfindet, der Phase des Zellzyklus, in der bevorzugt NHEJ-Reparatur stattfindet. Da NonO zusätzlich die Bindung verschiedener sequenzspezifischer Transkriptionsfaktoren an die DNA erhöht, kann es damit auch an anderen B-zellspezifischen Rekombinationsmechanismen beteiligt sein.
- (7d11) Ribonukleasen sind RNA schneidende Enzyme. Ribonuklease H1 ist an DNA-Synthese und Reparatur beteiligt und schneidet insbesondere RNA, die mit einem komplementären DNA-Strang gepaart ist.

2. SAM-Statistik

- (34g24) ARNT („aryl hydrocarbon receptor nuclear translocator,“) ist ein Mitglied der Helix-loop-Helix-Transkriptionsfaktoren.
- (4f19) Zinkfinger sind Transkriptionsfaktor-Untereinheiten, die einen stabilen Komplex zwischen Faktor und DNA bilden können. Damit sind sie in die Transkriptionskontrolle eingebunden.
- (6k19) ist ein unbekanntes DNA bindendes Protein.
- (4c22) BRCA1 assoziiertes Protein. BRCA1 ist an DNA Reparatur beteiligt. Die Mutation dieses Gens erhöht in Menschen die Wahrscheinlichkeit an Brusttumoren zu erkranken.
- (34k24) Y-Box Proteine modellieren die Expression von verschiedenen Genen und gehören zur Familie der Transkriptionsfaktoren.

Nach dieser Analyse wurden keine Gene gefunden, deren Funktion eindeutig mit Rekombination oder Reparatur in Zusammenhang gebracht werden kann. Ein DNA-bindendes Protein, oder ein Transkriptionsfaktor könnten jedoch eine Rolle in dem gesuchten Prozeß spielen. Um ihre Funktion weiter zu verfolgen, sollten die Ergebnisse jedoch erst durch ein weiteres Experiment bestätigt werden.

4.5 Semiquantitative RT-PCR

Um die bisherigen Ergebnisse zu bestätigen, wurde die differentielle Expression von 10 Kandidaten mittels RT-PCR überprüft. Die Expressionsunterschiede von 7 ESTs mit hoher Differenz (0,1=10fach; 0,2=5fach) aus den durch SAM detektierten Kandidaten (siehe Tab. 4), sowie 3 ESTs mit bestimmten Annotationen aus der t-Test-Analyse (siehe Tab. 5) sollten mittels semiquantitativer RT-PCR bestätigt werden.

| EST-Name | fold change | Annotation |
|----------|-------------|--------------------------------|
| 14l11 | (0,1) | Aminopeptidase B |
| 24m4 | (0,1) | * |
| 6h7 | (0,2) | Aktinin |
| 24f19 | (0,2) | IL-2 Rezeptor-Voläufner |
| 24l6 | (0,2) | * |
| 32a11 | (0,2) | * |
| 6k19 | (0,3) | dbpB/B-1,DNA bindendes Protein |

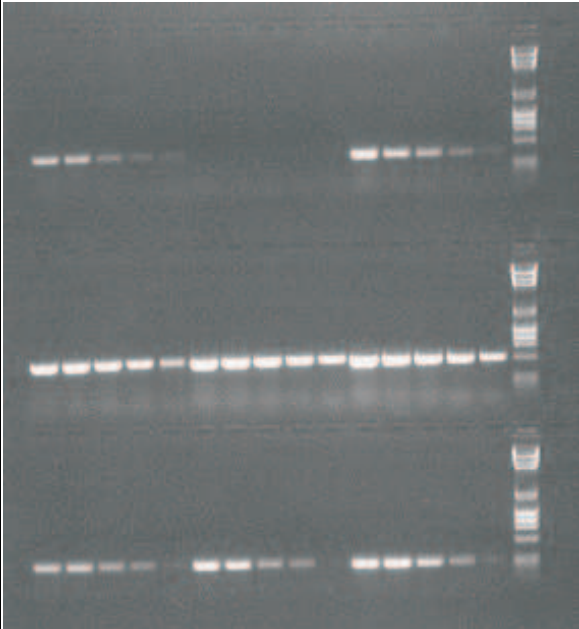
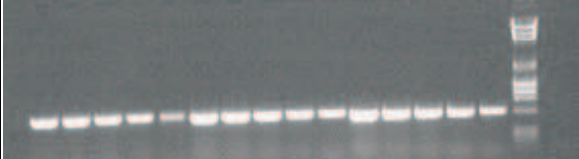
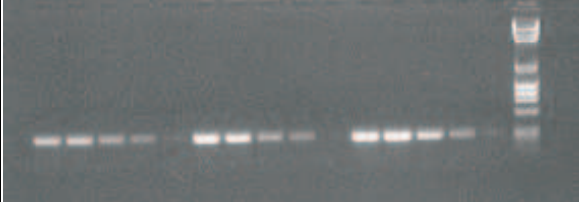

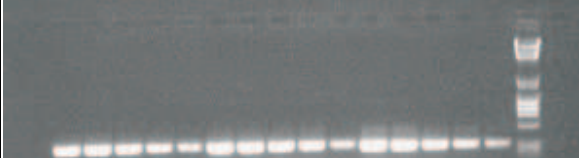
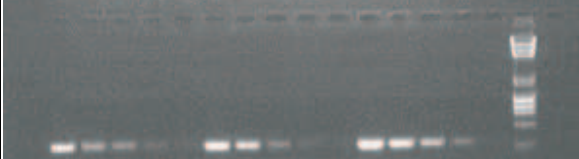
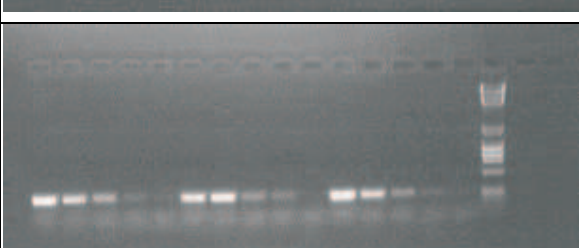
Tabelle 4: 7 SAM-Kandidaten wurden nach Expressionsdifferenz gewählt. „*“ bedeutet, dass der „blsp“- oder „blge“-Score niedrig war, so dass keine eindeutige Annotation gefunden wurde (Tab. 3.)

| EST-Name | Annotation |
|----------|-----------------------|
| 26d21 | UTR bindendes Protein |
| 12l20 | „NonO“ |
| 17o7 | v-myb Onkogen Homolog |

Tabelle 5: 3 t-Test-Kandidaten wurden nach ihren Annotationen gewählt. 26d21 und 12l20 sind durch ihre mutmaßliche DNA- oder RNA-Bindung interessant (siehe 4.4).

Die RNA von DT40 Cre1, DT40 Cre1 AID^{-/-} und Cl 18 wurde isoliert und in cDNA nach den Protokollen 3.2.9-3.2.10 (ohne radioaktive Markierung) umgeschrieben. Es wurde nacheinander cDNA aus 50; 16; 5; 1,6; 0,5 ng Gesamt-RNA eingesetzt, deren Amplifikationsprodukt nach PCR auf jedem Bild von links nach rechts in jeweils 5 Spuren pro Zellklon (DT40 Cre1, DT40 Cre1 AID^{-/-}, Cl 18) zu sehen ist (siehe Tab. 6). Die Abwesenheit der Amplifikation von AID in DT40 Cre1 AID^{-/-} wurde als negative Kontrolle eingesetzt. β -Aktin zeigt als positive Kontrolle überall gleich starke Banden. Der eingesetzte Längenmarker ist λ Hind III Φ X174 HaeIII. Da es keine Unterschiede in der Intensität des Bandenmusters der einzelnen verglichenen Zelllinien gibt, scheint die Expression der betreffenden

ESTs gleich zu sein. Damit konnte die differentielle Expression nicht bestätigt werden.

| | RT-PCR | | | | | | | | | | | | | | | |
|----------------|--------------------------------------------------------------------------------------|----|-----|-----|-----|---------|----|-----|-----|-----|------|----|-----|-----|-----|----|
| | DT40 Cre 1 | | | | | Aid -/- | | | | | C118 | | | | | SM |
| ng | 50 | 16 | 5.0 | 1.6 | 0.5 | 50 | 16 | 5.0 | 1.6 | 0.5 | 50 | 16 | 5.0 | 1.6 | 0.5 | |
| AID |  | | | | | | | | | | | | | | | |
| β -Aktin |  | | | | | | | | | | | | | | | |
| 6h7 |  | | | | | | | | | | | | | | | |
| 1411 |  | | | | | | | | | | | | | | | |
| 24f19 |  | | | | | | | | | | | | | | | |
| 2416 |  | | | | | | | | | | | | | | | |
| 24m4 |  | | | | | | | | | | | | | | | |

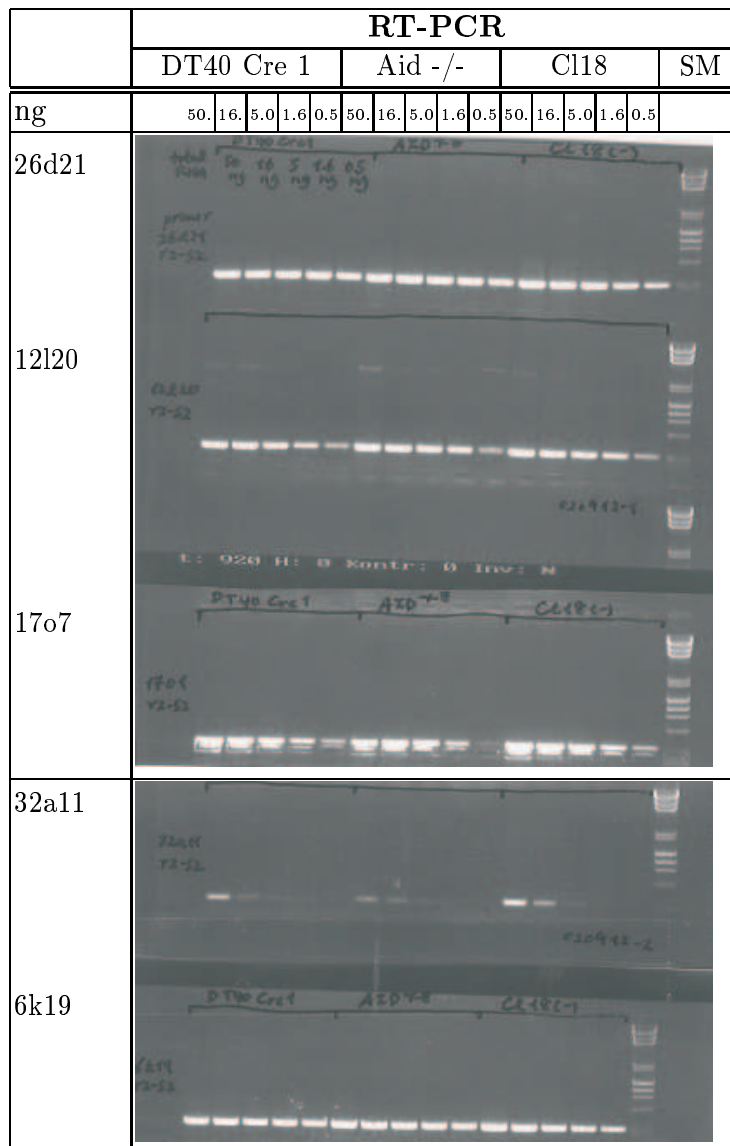


Tabelle 6: Das Bandenmuster dieser RT-PCR-Bilder zeigt, dass die untersuchten Gene, bzw. ESTs in allen drei Zelllinien, DT40 Cre1, AID -/- und Cl18 in gleicher Quantität vorliegen. Damit kann eine differentielle Expression dieser Gene in den einzelnen Zelllinien ausgeschlossen werden. Die ESTs sind links untereinander aufgeführt. Von links nach rechts sind die Mengen an RNA in ng aufgetragen. Der Längenmarker am rechten Rand wird mit SM beschrieben.

5 Diskussion

Das Ziel der vorliegenden Arbeit war es, einen Beitrag zur molekularen Wirkungsweise von AID zu liefern. Nach den Ergebnissen der hier durchgeführten Experimente wirkt AID nicht als globaler Transkriptionsfaktor in antikörperproduzierenden B-Zellen. Durch Hybridisierung der Filter konnten keine, durch ein weiteres Experiment bestätigte, Transkriptionsunterschiede zwischen Wildtyp- und Knockoutzelle festgestellt werden. Die hier detektierten scheinbaren Expressionsunterschiede waren nicht signifikant.

In der Arbeitshypothese wurden drei mögliche Szenarien zur Wirkung von AID angeführt (siehe Abschnitt 1 auf S. 4). Die Ergebnisse der Experimente zeigten, dass sich die Wirkung von AID nicht, wie im zweiten Szenario dargestellt, auf Transkriptionsebene niederschlägt. Vermutlich editiert AID, wie im dritten Szenario vorgeschlagen, die DNA und ruft durch die dadurch erzeugte Mutation sekundär Rekombinationsprozesse hervor. Die drei Szenarien zur Wirkung von AID stellen sich unter Einbeziehung der Ergebnisse folgendermassen dar:

1. AID als RNA-editierendes Enzym hätte vielleicht zu einer Veränderung von einzelnen mRNAs geführt. Ein dadurch geringgradig verändertes Transkriptionsprofil würde sehr wahrscheinlich nicht detektiert. Nach diesen Ergebnissen kann eine Editierung der mRNA nicht ausgeschlossen werden. Die alternative mRNA war vielleicht nur in wenigen Kopien vorhanden. Im niedrigen Expressionsbereich gibt es viele unspezifische Signale, so dass das richtige Signal darin hätte untergehen können.
2. Folgende Ergebnisse zeigten, dass es sich mit AID nicht um einen globalen Transkriptionsfaktor handelt, der primär die Expression vieler Gene induziert. Dabei wird davon ausgegangen, dass die durchgeführten Kontrollen, die während aller Zwischenschritte der Experimente durchgeführt wurden, ausreichend waren, um folgende Aussagen treffen zu können:
 - (a) Schon die von AIDA Compare erzeugten Korrelationsgraphen zeigten keine deutlichen Unterschiede im Vergleich zwischen Wildtypzelle vs. Wildtypzelle (Abb. 32 auf Seite 54) und Wildtypzelle vs. Knockoutzelle (Abb. 33 und 34 auf Seite 55). Dies wäre bei einer, durch einen globalen Transkriptionsfaktor hervorgerufenen, deutlichen Transkriptionsveränderung erwartet worden.
 - (b) Die Auswertung der Primärdaten mit dem AIDA-Programm ergab viele scheinbar differentiell exprimierte Werte. Die Werte der einzelnen Hybridisierungen unterschieden sich (siehe Korrelationsgraphen S.54-57). Wildtypzelle III und Knockoutzelle II lieferten mehr unspezifische Werte. Die entsprechenden Duplikatkorrelationsgraphen (Abb. 27 und 29 auf S. 51- 52) zeigten entsprechend eine stärkere Streuung.

- (c) Die Expressionswerte der einzelnen Hybridisierungen, die das Programm erzeugte, wurden anschliessend statistisch ausgewertet. Dadurch sollten die Werte gefunden werden, deren Expressionsdifferenz im Vergleich zwischen Wildtyp- und Knockouthybridisierung in allen Fällen gleichermassen hoch war. Die t-Test-Statistik ist dabei eine Standardmethode, um die Signifikanz von Werten wiederholter Messungen zu überprüfen. Sie ist nicht speziell auf die Anforderungen von Mikroarrays ausgerichtet. Die SAM-Statistik ist eine statistische Methode speziell für Mikroarrays [31]. So wurde davon ausgegangen, dass mit dieser Methode signifikante Gene detektiert würden. Beide Methoden wurden angewandt und erzeugten signifikant differentiell exprimierte Werte: Aus ursprünglich 14 000 aufgetragenen ESTs konnten etwas über 100 Kandidaten als in der Knockoutzelle herunterreguliert formuliert werden. Beide Methoden bezeichneten unterschiedliche Gene als signifikant. Dies liegt daran, dass keine wirklichen Expressionsunterschiede vorlagen. Die Ergebnisse ergaben sich rein rechnerisch aus den unterschiedlichen Signifikanzkriterien der Methoden.
 - (d) Es wurde erwartet, dass für Genkonversion spezifische Gene mit Rekombination oder Reparatur zu tun haben sollten. Bei der Suche der von der Statistik als signifikant bezeichneten Gene in den Datenbanken, wurden keine Bezeichnungen gefunden, die eindeutig mit diesen Prozessen in Verbindung gebracht werden konnten.
 - (e) Eine Auswahl der durch t-test oder SAM-Statistik gewonnenen Kandidatengene konnte durch das RT-PCR-Experiment nicht bestätigt werden. Das Mikroarray-Experiment konnte also durch ein weiteres Experiment nicht bestätigt werden.
3. Nach der neueren Literatur wirkt AID als DNA-editierendes Enzym, was nicht primär zu einer Transkriptionsveränderung führt.

AID deaminiert Cytosine zu Uracilen an der DNA der V- bzw. C-Region. Die so erzeugten Mutationen führen zu unterschiedlichen Reparaturmechanismen, die Genkonversion, somatische Hypermutation und „Class-switch-recombination“ nach sich ziehen [58], [22]. Dass es durch die ektopische Expression von AID in B-Zellhybridomen und Fibroblasten zu induzierten Mutationen auf der DNA kommt, unterstützt diese Vermutung und zeigt, dass somatische Hypermutation nicht nur auf die V-Region beschränkt ist [43], [86]. Durch AID induzierte Mutationen spielen vermutlich nicht nur bei der Rekombination der Immunglobulingensegmente eine Rolle, sondern können die Kanzerogenese initiieren. Da sich, durch Ausschaltung von RAD51-Paralogen in der DT40-Zelllinie, Genkonversion zu somatischer Hypermutation umwandeln liess, wurde vermutet, dass diese beiden Rekombinationsmechanismen von einer gemeinsamen Läsion hervorgerufen werden könnten

[68]. Wenn AID Uracile auf der DNA erzeugt, müssen diese entfernt werden. Normalerweise wirken Uracile auf der DNA als Substrate für „BER“-Reparatur, dem Hauptmechanismus zur Eliminierung von Uracilen auf der DNA [42]. Dabei wird zunächst eine Uracil-DNA-Glycosylase aktiviert, die die Zucker-Basenbindung löst und damit Uracil aus der DNA ausschneidet. Darauf wird das Phosphodiester-Rückgrat von einer AP-Endonuklease/Lyase eingeschnitten und so eine abasische Stelle auf der DNA zurückgelassen. Polymerase β inseriert anschliessend ein Cytosin, worauf die Mutation korrigiert wäre. Wenn nun statt einer korrekten Reparatur eine Mutation resultieren soll, kommt es im Weiteren darauf an, welche Reparaturenzyme vorhanden sind. Durch Aktivierung spezifischer Polymerasen, wie Polymerase ι , Polymerase η oder Polymerase ζ kommt es zu einem fehlerhaften Baseneinbau, der zu somatischer Hypermutation führen kann [7], [88], [20]. Bei „Class-switch-recombination“ und somatischer Hypermutation sind MSH-Proteine beteiligt [24], [59], [70]. Genkonversion kann nur stattfinden, wenn sich homologe Pseudogensequenzen in nächster Nähe befinden, die als Matrize dienen können. In DT40-Zellen findet in hoher Rate Genkonversion, und gleichzeitig in einer niedrigeren Rate somatische Hypermutation statt. Wenn ein für Genkonversion essentielles Enzym der RAD51-Gruppe ausgeschaltet wird, wird der Hypermutationsmechanismus aktiviert [68]. Dies deutet darauf hin, dass die Wahl des Mutationsmechanismus nicht nur von vorhandenen Enzymen, sondern auch von der Möglichkeit zur homologen Rekombination abhängig ist, und dass diese, wenn möglich bevorzugt wird.

Transkriptionsprofile können dazu beitragen, einen ersten Überblick darüber zu geben, welche Gene von bestimmten Faktoren abhängig sind. Zu Beginn dieser Untersuchungen wusste man noch wenig über die Wirkungsweise von AID. Im Verlauf konnten dann, durch die Veröffentlichungen anderer Arbeitsgruppen, AID-abhängige Faktoren weitgehend eingeschränkt werden.

Vielleicht hat sich die angewandte Methode als nicht hinreichend empfindlich erwiesen, einzelne, differentiell exprimierte Gene, zwischen anderen, unspezifischen Signalen zu detektieren. Die Transkripte der beteiligten Polymerasen und Reparaturenzyme könnten andererseits sowohl in der Wildtyp- als auch in der Knockoutzelle vorhanden sein, wenn die Enzyme, neben der spezifischen Reparaturfunktion, auch andere Funktionen erfüllen. Die Aktivität dieser Enzyme würde dann nicht auf Transkriptionsebene reguliert und würde daher auch nicht zu einer differentiellen Expression führen.

Der gezielte Nachweis von Transkripten der betreffenden Reparaturenzyme in Wildtyp- und Knockoutzelle, könnte in Zukunft aufklären, in welcher Weise die Regulation stattfindet.

6 Zusammenfassung

Um mit einer wirksamen Immunantwort auf möglichst viele in der Umwelt vorhandene Antigene reagieren zu können, haben B-Zellen die Fähigkeit, ihr genetisches Material zu rekombinieren und dadurch aus einer relativ geringen Anzahl von Gensegmenten auf der DNA ein sehr grosses Repertoire an Genkombinationen zu bilden. Dabei kommt es temporär zu einer Destabilisierung der DNA, die gleichzeitig immer auch eine Gefahr für die Zelle bedeutet. Der Rekombinationsmechanismus sollte also sehr fein reguliert sein, damit die Zelle von dieser Fähigkeit profitieren kann. Eine Schlüsselstellung in der Regulation dieser Rekombinationsvorgänge spielt die Aktivitätsinduzierte Cytidin Deaminase (AID). Angesichts der grundlegenden Bedeutung von AID für die Erzeugung der Antikörpervariabilität ist es erforderlich, die bisher wenig bekannte molekulare Wirkungsweise dieses Enzyms aufzuklären. In der vorliegenden Arbeit sollte untersucht werden, ob AID als Transkriptionsfaktor wirkt, der die Expression von Genen induziert, die an Rekombinationsvorgängen beteiligt sind. Als Modellsystem für die Untersuchungen wurde die B-Zelllinie DT40 benutzt, da in dieser Zelllinie Genkonversion, als einer der beteiligten Rekombinationsmechanismen, abläuft. Frühere Arbeiten hatten gezeigt, dass nach Ausschalten des AID-Gens in DT40-Zellen keine Genkonversion mehr stattfindet.

Um mögliche Expressionsunterschiede darzustellen, wurden Transkriptionsprofile von AID-positiven und AID-negativen Zellen angefertigt. Mittels dieser für die Klinik bisher relativ neuen molekulardiagnostischen Methode kann die Transkriptionsaktivität von Zellen in Form von Mikroarrays sichtbar gemacht werden. Die Mikroarrays waren in diesem Falle Nylonfilter, auf denen B-zellspezifische cDNAs aufgetragen sind. Die Mikroarray-Filter wurden dann mit radioaktiv markierter cDNA aus mRNA AID-positiver bzw. AID-negativer DT40-Zellen hybridisiert. Falls AID die Transkription von Genen induziert, sollten sich die Hybridisierungsmuster unterscheiden. Nach Analyse der Primärdaten der Profile schienen einzelne Gene unterschiedlich hoch exprimiert zu sein. Die Primärdaten sind die Daten, die von einem speziell für Mikroarrays konzipierten Computerprogramm erzeugt werden. Danach kann jedem Punkt gemäß seiner Intensität auf dem Filter ein Zahlenwert zugeordnet werden. Da bei Anfertigung der Arrays viele nichtsignifikante Werte erzeugt werden, wurde durch statistische Methoden angestrebt, diese aus der Auswertung herauszufiltern. Die meisten der durch das Programm berechneten Werte fielen damit hinaus. Einige Gene wurden gefunden, deren Expression in der AID-negativen Zelle herunterreguliert zu sein schien. Dieses Ergebnis wurde überprüft, indem die RNA-Mengen dieser Kandidatengene direkt mit RT-PCR verglichen wurden. Die detektierten, statistisch signifikanten Expressionsunterschiede konnten mit dieser Methode nicht bestätigt werden.

Zusammenfassend zeigen die Ergebnisse, dass AID nicht als Transkriptionsfaktor wirkt, der das Expressionsmuster der Zelle global beeinflusst. Allerdings

kann es sein, dass einzelne differentiell exprimierte Gene mit der hier verwendeten Methode nicht erfasst wurden. Dieses Ergebnis stimmt mit kürzlich veröffentlichten Arbeiten überein, wonach das AID-Genprodukt nicht primär die Transkription von Genen induziert, sondern Cytosin zu Uracil auf der DNA deaminiert. Die so erzeugte Mutation ruft sekundär Reparaturprozesse hervor.

7 Abkürzungen

| | |
|-----------------|--------------------------------------------------------------|
| AID | Aktivitätsinduzierte Cytidin Deaminase |
| AIDA | Advanced Image Data Analysis |
| ALV | Avian leukosis virus |
| APOBEC-1 | Apolipoprotein B editing catalytic subunit 1 |
| AP-Endonuklease | aprimidinische/ apurinische Endonuklease |
| ATP | Adenosintriphosphat |
| BER | base excision repair |
| blge | Gene bank Datenbank |
| blsp | Swiss Protein Datenbank |
| bp | Basenpaare |
| B-Zelle | Antikörperproduzierende Zelle, bei Vögeln bursalen Ursprungs |
| C | Celsius |
| cDNA | complementary DNA (mit RT umgeschriebene mRNA) |
| cpm | counts per minute |
| C-Region | Konstante Region des Antikörpers |
| C_H | konstante Schwereketten |
| CSR | Class switch recombination |
| CTP | Cytidintriphosphat |
| CVID | Common variable immunodeficiency |
| dA | Desoxyadenosin |
| dCTP | Desoxycytidintriphosphat |
| DEPC | Diethylenpyrocarbonat |
| dkfz | Deutsches Krebsforschungszentrum |
| DLBCL | Diffuse large B-cell lymphoma |
| DMSO | Dimethylsulfoxide |
| DNA | Desoxyribonukleinsäure |
| DTT | Dithiothreitol |
| dUTP | Desoxyuridintriphosphat |
| EDTA | Ethylenediamintetraessigsäure |
| EST | Expressed sequence tag |
| FACS | Fluorescence activated cell sorting |
| FBS | Fetales Kälberserum |
| GC | Genkonversion |
| GTP | Guanosintriphosphat |
| HCl | Salzsäure |
| HIGM | Hyper-IgM-Syndrom |
| H_2O | Wasser |

| | |
|------------------------|---------------------------------------------------|
| HR | Homologe Rekombination |
| Ig | Immunglobulin |
| IgA | Immunglobulin A |
| IgE | Immunglobulin E |
| IgH | Immunglobulin heavy chain |
| IgL | Immunglobulin light chain |
| IgM | Immunglobulin M |
| J | Junction |
| kb | Kilobase |
| kbp | kilobasepair |
| LTR | Long terminal repeat |
| μ g | Mikrogramm |
| MHC | Major histocompatibility complex |
| min. | Minuten |
| ml | Mililiter |
| mRNA | messenger RNA |
| NaOH | Natriumhydroxyd |
| NaPi | Natrium-(Hydrogen-)phosphat |
| ng | Nanogramm |
| NHEJ | Non-homologous end joining (Reparaturmechanismus) |
| Oligos | Oligonukleotide |
| PAP1 | Poly(A)Polymerase1 |
| PCR | Polymerase chain reaction |
| pg | Picogramm |
| pH | Wasserstoffionenkonzentration |
| PNK | Polynukleotidkinase |
| PolyA ⁺ RNA | Polyadenylierte RNA |
| RACE | Rapid amplification of cDNA ends |
| RAG | Recombination activating gene |
| REAL | Revised european american lymphoma |
| RNA | Ribonukleinsäure |
| RNase | Ribonuklease |
| RNasin | RNaseinhibitor |
| rpm | Umdrehungen pro Minute |
| rRNA | ribosomale RNA |
| RT | Reverse Transcriptase |
| RT-PCR | Reverse transcription polymerase chain reaction |
| RZPD | Ressourcenzentrum für Genomforschung |
| SAM | Significance analysis of microarrays |
| S Region | Signal sequence region |
| SsDNA | Salmon sperm DNA |

| | |
|----------|-------------------------------------|
| SCID | Severe combined immunodeficiency |
| SDS | Natriumdodecylsulfat |
| SH | somatische Hypermutation |
| SSC | Standard Saline Citrate |
| TAE | Tris-Acetat-EDTA-Puffer |
| TTP | Thymidintriphosphat |
| tRNA | transfer RNA |
| UTP | Uridintriphosphat |
| UTR | untranslatierte Region |
| V(D)J | variable (diversity) joining |
| v-myb | avian myeloblastosis viral oncogene |
| V-Region | variable Region des Antikörpers |
| vs. | versus |

Literatur

- [1] Abdrakhmanov I, Lodygin D, Geroth P, Arakawa H, Law A, Plachy J, Korn B, Buerstedde J-M (2000) A large database of chicken bursal ESTs as a resource for the analysis of vertebrate gene function. *Genome Research* 10:2062-2069
- [2] Alberts B (1994) *Molecular Biology of the Cell*, 3rd ed. New York: Garland Publishing, Inc.
- [3] Alizadeh A A, Eisen M B, Davis R E, Ma C, Lossos I S, Rosenwald A, Boldrick J C, Sabet H, Tran T, Yu X, Powell J I, Yang L, Marti G E, Moore T, Hudson Jr J, Lu L, Lewis D B, Tibshirani R, Sherlock G, Chan W C, Greiner T C, Weisenburger D D, Armitage J O, Warnke R, Levy R, Wilson W, Grever M R, Byrd J C, Botstein D, brown P O, Staudt L M (2000) Distinct types of diffuse large B- cell lymphoma identified by gene expression profiling. *Nature* 403:503-511.
- [4] Allen R C, Armitage R J, Conley M E, Edelhoff S, Disteché C M, Simoneau D K (1993) CD40 ligand gene defects responsible for X-linked hyper IgM-Syndrome. *Science* 259:990-993
- [5] Arakawa H, Hauschild J, Buerstedde J M (2002) Requirement of the activation- induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* 295:1301-1306
- [6] Baba T W, Giroir B P, Humphries E H (1985) Cell lines derived from avian lymphomas exhibit two different phenotypes. *Virology* 144: 139-144
- [7] Bebenek K, Tissier A, Frank E G, Mc Donald J P, Prasad R, Wilson S H, Woodgate R, Kunkel T A (2001) 5'-Desoxyribose phosphate lyase activity of human DNA polymerase ι in vitro. *Science* 291:2156-2159
- [8] Berek C, Berger A, Apel M (1991) Maturation of the immune response in germinal centers. *Cell* 67:1121-1129
- [9] Betz A G, Milstein C, Gonzalez-Fernandez A, Pannel R, Larson T, Neuberger M S (1994) Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell* 77:239-248
- [10] Bezzubova O Y, Buerstedde J M (1994) Gene conversion in the chicken immunoglobulin locus: A paradigm of homologous recombination in higher eukaryotes. *Experientia* 50:270-276
- [11] Bonifaci N, Moroianu J, Radu A, Blobel G (1997) Karyopherin $\beta 2$ mediates nuclear import of a mRNA binding protein. *Proc. Nat. Acad. Sci.* 94:5055-5060

- [12] Bosma G C, Custer R P, Bosma M J (1983) A severe combined immunodeficiency mutation in the mouse. *Nature* 301:527-530
- [13] Brazma A, Hingamp P, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansong W, Ball C A, Causton H C, Gaasterland T, Glenisson P, Holstege F C P, Kim I F, Markowitz V, Matese J C, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M (2001) Minimum information about a microarray experiment (MIAME)- toward standards for microarray data. *Nature Genetics* 29:365-371
- [14] Bross L, Fukita Y, McBlane F, Demolliere C, Rajewski K, Jacobs H (2000) DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. *Immunity* 13:589-597
- [15] Bross L, Muramatsu M, Kinoshita K, Honjo T, Jacobs H (2002) DNA double-strand breaks: Prior to but not sufficient in targeting hypermutation. *J. Exp. Med.* 195:1187-1192
- [16] Brown P O, Botstein D (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genetics* 21: 33-37
- [17] Buerstedde J M, Reynaud C A, Humphries E H, Olson W, Ewert D L, Weill J C (1990) Light chain gene conversion continues at high rate in an ALV-induced cell line. *EMBO J.* 9:921-927
- [18] Buerstedde J M, Takeda S (1991) Increased ratio of targeted to random integration after transfection of chicken B- cell lines. *Cell* 67:179-188
- [19] Chua K F, Alt F W, Manis J P (2002) The function of AID in somatic mutation and class switch recombination: Upstream or downstream DNA breaks. *J. Exp. Med.* 195:F37-F41
- [20] Diaz M, Flajnik M F (1998) Evolution of somatic hypermutation and gene conversion in adaptive immunity. *Immunological Reviews* 162:13-24
- [21] Diaz M, Verkoczy L K, Flajnik M F, Klinman N R (2001) decreased frequency of somatic hypermutation and impaired affinity maturation but intact germinal center formation in mice expressing antisense RNA to DNA polymerase zeta. *J. Immunol.* 167:327-335
- [22] Di Noia J, Neuberger M S (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419:43-48
- [23] Eisen M B, Spellman P T, Brown P O, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Nat. Acad. Sci.* 95: 14863-14868

- [24] Ehrenstein M R, Neuberger M S (1999) deficiency in MSH2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *EMBO J.* 18:3484-3490
- [25] Faili A, Aoufouchi S, Gueranger Q, Zober C, Leon A, Bertocci B, Weill J-C, Reynaud C-A (2002) AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. *Nature Immunology* 3:815-821
- [26] Fugmann S D (2001) RAG1 and RAG2 in V(D)J recombination and transposition. *Immunol. Res.* 23:23-39
- [27] Fukita Y, Jacobs H, Rajewsky K (1998) Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* 9:105-114
- [28] Fulop G M, Philips R A (1990) The SCID mutation causes a general defect in DNA repair. *Nature* 347:479-482
- [29] Gellert M (1992) Molecular analysis of V(D)J recombination. *Annu. Rev. Genet.* 22:425-446
- [30] Goossens T, Klein U, Küppers R (1998) Frequent occurrence of deletions and duplications during somatic hypermutation: Implication for oncogene translocations and heavy chain disease. *Proc. Nat. Acad. Sci.* 95: 2463-2468
- [31] Goss Tusher V, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Nat. Acad. Sci.* 98:5116-5121.
- [32] Gray NK, Wickens M (1998) Control of translation initiation in animals. *Annu. Rev. Cell. Dev. Biol.* 14:399-458
- [33] Hacia J C (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nature Genetics* 21:42-47.
- [34] Harris N L, Jaffe E S, Stein H, Banks P M, Chan J K, Cleary M L, Delsol G, De Wolf- Peeters C, Falini B, Gatter K C (1994) A revised European- American classification of lymphoid neoplasms: a proposal from the international lymphoma study group. *Blood* 84:1361-1392
- [35] Hayward W S, Neel B G, Astrin S M (1981) Activation of a cellular onc gene by promotor insertion in ALV- induced lymphoid leukemia. *Nature* 290:475-479
- [36] Hendrickson E A, Qin X-Q, Schatz D G, Oettinger M, Weaver D T (1991) A link between double- strand break- repair and V(D)J recombination: the scid mutation. *Proc. Natl. Acad. Sci.* 88:4061-4065

- [37] Honjo T, Kinoshita K, Muramatsu M (2002) Molecular mechanism of Class switch recombination- linkage with somatic hypermutation. *Annu. Rev. Immunol.* 20: 165-196
- [38] Hughes T R, Marton M J, Jones A R, Roberts C J, Stoughton R, Armour C D, Bennett H A, Coffey E, Dai H, He Y D, Kidd M J, King A M, Meyer M R, Slade D, Lum P Y, Stepaniants S B, Shoemaker D D, Gachotte D, Chakraburttty K, Simon J, Bard M, Friend S H (2000)Functional discovery via a compendium of expression profiles. *Cell* 102:109-126
- [39] Jacobs H, Bross L (2001) Towards an understanding of somatic hypermutation. *Curr. Op. Immunology* 13:208-218
- [40] Janeway C A, Travers P, Walport M, Shlomchik M (2001) The generation of lymphocyte antigen receptors. In: *Immunobiology*, Garland Publishing, fifth edition, New York, p.123
- [41] Levy Y, Gupta N, De Deist F, Garcia C, Fischer A, Weill J-C, Reynaud C-A (1998) Defect in IgV gene somatic hypermutation in Common Variable Immuno-Deficiency syndrome. *Proc. Nat. Acad. Sci.* 95:13135-13140
- [42] Lindahl T (2000) Suppression of spontaneous mutagenesis in human cells by DNA base excision-repair. *Mutat. Res.* 462: 129-135
- [43] Martin A, Bardwell P D, Woo C J, Fan M, Shulman M J, Scharff M D (2002) Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* 415:802-806
- [44] Mayer B J, Ren R, Clark K L, Baltimore B (1993) A putative modular domain present in diverse signaling molecules. *Cell* 73:629-630
- [45] Mc Cormack W T, Thompson C B (1990)Chicken Ig light variable region gene conversion display pseudogene donor preference and 5'to 3' polarity. *Genes Dev.* 4:548-558
- [46] Mc Cormack W T, Tjoelker L W, Thompson C B (1991) Avian B-cell development: Generation of an Immunoglobulin repertoire by gene conversion. *Annu. Rev. Immunol.* 9: 219-227
- [47] Mehta A, Kinter M T, Sherman N E, Driscoll D M (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding-protein involved in the editing of apolipoprotein B mRNA. *Mol. Cell. Biol.* 20:1846-1854
- [48] Muramatsu M, Sankaranand V S, Anant S, Sugai M, Kinoshita K, Davidson N O, Honjo T (1999)Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA- editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274:18470-18476.

- [49] Muramatsu M, Kinoshita S, Fagarasan S, Yamada S, Shinkai Y, Honjo T (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102:553-563
- [50] Muramatsu M, Honjo T (2001) Complex layers of genetic alteration in the generation of antibody diversity. *Trends in Immunology* 22:66-68
- [51] Neuberger M S, Milstein C (1995) Somatic Hypermutation. *Curr. Opin. Immunol.* 7:248-254
- [52] The Non- Hodgkin's Lymphoma Classification Project (1997) A clinical evaluation of the international lymphoma study group classification of Non-Hodgkin's lymphoma. *Blood* 89:3909-3918
- [53] Notarangelo L D, Duse M, Ugazio A G (1992) Immunodeficiency with hyper IgM. *Immunodef. Rev.*3(2):101-121
- [54] Oettinger M A, Schatz D G, Gorka C, Baltimore D (1990) RAG-1 and RAG-2 adjacent genes that synergistically activate V(D)J recombination. *Science* 248:1517-1523
- [55] Okazaki I, Kinoshita K, Muramatsu M, Yoshikawa K, Honjo T (2002) The AID enzyme induces class switch recombination in fibroblasts. *Nature* 416:340-345
- [56] Papavasiliou F N, Schatz D G (2000) Cell- cycle- regulated DNA double strand breaks in somatic hypermutation of immunoglobulin genes. *Nature* 408:216-221
- [57] Papavasiliou F N, Schatz D G (2002) The activation- induced deaminase functions in a postcleavage step of the somatic hypermutation process. *J. Exp. Med.* 195:1193-1198
- [58] Petersen-Mahrt SK, Harris RS, Neuberger MS (2002) AID mutates E.coli, suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418:99-103
- [59] Rada C, Williams G T, Nilsen H, Barnes D E, Lindahl T, Neuberger M S (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr.Biol.* 12(20):1748-55
- [60] Rajewski K (1996) Clonal selection and learning in the antibody system. *Nature* 228:751-758

- [61] Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Langelouse R, Gennery A, Tezcan I, Erosoy F, Kayserili H, Ugazio A G, Brousse N, Muramatsu M, Notarangelo L D, Kinoshita K, Honjo T, Fischer A, Durandy A (2000) Activation- Induced Cytidine Deaminase (AID) Deficiency causes the autosomal recessive form of the Hyper IgM Syndrome (HIGM2) *Cell*. 102:565-575
- [62] Reynaud C-A, Anquez V, Dahan A, Weill J-C (1985) A single rearrangement event generates most of the chicken Ig light chain diversity. *Cell* 40:283-291
- [63] Reynaud C-A, Anquez V, Grimal H, Weill J-C (1987) A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell* 48:379-388
- [64] Rosenberg S A (1994) Classification of lymphoid neoplasms. *Blood* 84:1359-1360
- [65] Roth D B, Nakajima P B, Menetski J P, Bostma M J, Gellert M (1992)V(D)J recombination in mouse thymocytes: double-strand breaks near T cell receptor delta rearrangement signals. *Cell* 69:41-53
- [66] Rowlands J C, Gustafsson J A (1997) Aryl hydrocarbon receptor-mediated signal transduction. *Critical Rev. Toxicol.*27:109-134
- [67] Sale J E, Neuberger M S (1998) TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. *Immunity* 9:859-869
- [68] Sale J E, Calandrini D M, Takata M, Takeda S, Neuberger M S (2001) Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation. *Nature* 412:921-6
- [69] Schatz D G, Oettinger M A, Baltimore D (1989)The V(D)J recombination activating gene RAG-1. *Cell* 59:1035-1048
- [70] Schrader C E, Edelmann W, Kucherlapati R, Stavnezer J (1999) Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. *J. Exp. Med.* 190: 323-330
- [71] Sörlie T, Perou C M, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen M B, van de Rijn M, Jeffrey S S, Thorsen T, Quist H, Matese J C, Brown P O, Botstein D, Eystein Lonning P, Borresen-Dale A L (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Nat. Acad. Sci.* 98:10869-10874
- [72] Stavmezer J (2000) Molecular processes that regulate class switching. *Curr. Top. Microbiol. Immunol.* 245:127-168

- [73] Stevenson F, Sahota S, Zhu D, Ottensmeier C, Chapman C, Oscier D, Hamblin T (1998) Insight into the origin and clonal history of B- cell tumors as revealed by analysis of immunoglobulin variable region genes. *Immunological Reviews* 1998. Vol. 162:247-259
- [74] Storb U (1998) Progress in understanding the mechanism and consequences of somatic hypermutation. *Immunological Reviews* 162:5-11
- [75] Takata M, Sasaki M S, Sonoda E, Morrison C, hashimoto M, Utsumi H, Yamagushi-Iwai Y, Shinohara A, takeda S (1998) Homologous recombination and non- homologous end- joining pathways of DNA double- strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* 17:5497-5508
- [76] Thompson C B (1992) Creation of immunoglobulin diversity by intrachromosomal gene conversion. *Trends Genet.* 8:416-442
- [77] Thompson C B, Humphries E H, Carlson L M, chen C L, Neiman P E (1987) The effect of alterations in myc gene expression on B cell development in the bursa of Fabricius. *Cell* 51:371-377
- [78] Thompson C B, Neiman P E (1987) Somatic diversification of the chicken Ig light chain gene is limited to the rearranged variable gene segment. *Cell* 48:369-378
- [79] Ting Lee M-L, Kuo F C, Whitmore G A, Sklar J (2000) Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Nat. Acad. Sci* 97:9834-9839.
- [80] Tonegawa S (1983) Somatic generation of antibody diversity. *Nature* 302:575-81
- [81] Wiel van de M A (2002) Significance analysis of microarrays using rank scores. Technical report, Department of mathematics and computing science, Eindhoven University of Technology
- [82] Venter J C, Adams M D, Myers E W, Li P W, Mural R J, Sutton G G et al (2001) The sequence of the human genome. *Science* 291:1304-51.
- [83] Weigert M, Gattmaitan L, Loh E, Schilling J, Hood L (1978) Rearrangement of genetic information may produce immunoglobulin diversity. *Nature* 276:785-90
- [84] Weill J-C, Reynaud C-A (1987) The chicken B- cell compartment. *Science* 238:1094-1098

- [85] Yang Y-S, Yang M-C W, Tucker P W, Capra J D (1997) NonO enhances the association of many DNA-binding proteins to their targets. *Nucleic Acid. Res.* 25:2284-2292
- [86] Yoshikawa K, Okazaki I M, Kinoshita K, Muramatsu M, Nagaoka H, Honjo T (2002) AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* 296:2033-2036
- [87] Young R A (2000) Biomedical discovery with DNA arrays. *Cell* 102:9-15
- [88] Zeng X, Winter D B, Kasmer C, Kraemer K H, Lehmann A R, Gaerhart P J (2001) DNA Polymerase eta is a T-A mutator in somatic hypermutation of immunoglobulin variable genes. *Nature Immunology* 2:537-541

8 Danksagung

Mein Dank gilt insbesondere Jean-Marie Buerstedde, der mir diese Arbeit ermöglicht hat, sowie allen Mitarbeitern des Labors, die mir während der Arbeit hilfreich zur Seite standen: Hiroshi Arakawa, dessen AID-knockout-Zelllinie ich benutzen durfte, Jessica Hauschild, die immer wieder geduldig meine Arbeit gelesen und korrigiert hat, Paul Geroth, der stets mit einem guten Ratschlag dabei war. Klaus Harbers danke ich für seine Geduld und Bereitschaft, mich zu unterstützen. Außerdem danke ich Kimmo Koskela und Pekka Kohonen, sowie Olli Lassila von der Universität Turku und Anne-Marie Chaussee von der INRA in Tours für ihre Hilfe. Astrid, Nora und Ullrich danke ich für die Korrektur. Frank danke ich für die unermüdliche Unterstützung am Computer.

9 Lebenslauf

- geboren am 14.12.1971 in Hamburg
- 1991 Abitur
- 1991-1992 Aupair-Aufenthalt in Island
- 1993-2001 Studium der Humanmedizin in Hamburg und Dresden
- 2000-2001 PJ in Hamburg am UKE und im Bernhard-Nocht-Institut für Tropenmedizin
- 2001-2002 Heinrich-Pette-Institut der Universität Hamburg in der Abteilung für zelluläre Immunologie

Abbildungsverzeichnis

| | | |
|----|------------------------------------------------------------------------------------------------------|----|
| 1 | V(D)J-rekombination | 7 |
| 2 | Genkonversion | 8 |
| 3 | Somatische Hypermutation | 9 |
| 4 | „Class-switch-recombination“ | 11 |
| 5 | Abkömmlinge von DT40 | 14 |
| 6 | Mögliche Aktivität von AID | 15 |
| 7 | RZPD-Filter | 25 |
| 8 | SAM-Plot | 38 |
| 9 | Oligohybridisierung A | 42 |
| 10 | Oligohybridisierung B | 42 |
| 11 | Ausschnitt aus Oligohybridisierung A | 42 |
| 12 | Ausschnitt aus Oligohybridisierung B | 42 |
| 13 | DT40 Cre1 I | 44 |
| 14 | DT40 Cre1 AID ^{-/-} I | 44 |
| 15 | Ausschnitt aus DT40 Cre1 I | 44 |
| 16 | Ausschnitt aus DT40 Cre1 AID ^{-/-} I | 44 |
| 17 | DT40 Cre1 II | 45 |
| 18 | DT40 Cre1 AID ^{-/-} II | 45 |
| 19 | Ausschnitt aus DT40 Cre1 II | 45 |
| 20 | Ausschnitt aus DT40 Cre1 AID ^{-/-} II | 45 |
| 21 | DT40 Cre1 III | 46 |
| 22 | Ausschnitt aus DT40 Cre1 III | 47 |
| 23 | Wertetabelle AIDA Array Metrix | 48 |
| 24 | Duplikatkorrelation der Oligohybridisierung | 49 |
| 25 | Duplikatkorrelation DT40 Cre1 I | 50 |
| 26 | Duplikatkorrelation DT40 Cre1 II | 50 |
| 27 | Duplikatkorrelation DT40 Cre1 III | 51 |
| 28 | Duplikatkorrelation DT40 Cre1 AID ^{-/-} I | 51 |
| 29 | Duplikatkorrelation DT40 Cre1 AID ^{-/-} II | 52 |
| 30 | Wertetabelle AIDA Array Compare | 53 |
| 31 | Korrelation der Oligohybridisierungen | 54 |
| 32 | Korrelation: DT40 Cre1 I („Master“) und DT40 Cre1 II („Client“)(SD=1,8/- 1,7) | 54 |
| 33 | Korrelation: AID ^{-/-} I („Master“) und DT40 Cre1 I („Client“)(SD=1,9/- 2,0) | 55 |
| 34 | Korrelation: AID ^{-/-} I („Master“) und DT40 Cre1 II („Client“)(SD=1,9/- 1,7) | 55 |
| 35 | Korrelation: AID ^{-/-} I („Master“) und DT40 Cre1 III („Client“)(SD=1,9/- 6,0) | 56 |
| 36 | Korrelation: AID ^{-/-} II („Master“) und DT40 Cre1 I („Client“)(SD=3,8/- 1,8) | 56 |

| | | |
|----|---------------------------------------------------------------------------------------|----|
| 37 | Korrelation: AID-/- II („Master“) und DT40 Cre1 II („Client“)(SD=3,9/-1,5) | 57 |
| 38 | Korrelation: AID-/- II („Master“) und DT40 Cre1 III („Client“)(SD=2,5/-3,2) | 57 |

Tabellenverzeichnis

| | | |
|---|----------------------------------------------------------------------|----|
| 1 | Diversifikationsmechanismen verschiedener Spezies | 6 |
| 2 | Durch t-Test-Statistik ermittelte differentiell exprimierte ESTs . . | 60 |
| 3 | Durch SAM-Statistik ermittelte differentiell exprimierte ESTs . . | 63 |
| 4 | „SAM-Kandidaten“ | 68 |
| 5 | „t-Test-Kandidaten“ | 68 |
| 6 | RT-PCR | 71 |