

Medizinische Fakultät Charité – Universitätsmedizin Berlin  
Campus Benjamin Franklin  
aus dem Institut für Allgemeine Pathologie  
Direktor: Prof. Dr. med. Harald Stein

## **THEMA**

# **Molekulare Abgrenzung des Burkitt-Lymphoms vom diffusen großzelligen B-Zell-Lymphom durch Microarray- basierte Genexpressionsanalysen**

Inaugural-Dissertation  
zur Erlangung der  
medizinischen Doktorwürde  
Charité – Universitätsmedizin Berlin  
Campus Benjamin Franklin

vorgelegt von  
Michael Kühn  
aus Hannover

Referent: Priv.- Doz. Dr. rer. nat. Michael Hummel

Koreferent: Prof. Dr. med. Peter Daniel

Gedruckt mit Genehmigung der Charité - Universitätsmedizin Berlin  
Campus Benjamin Franklin

Promoviert am: 01.06.2008

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	- 7 -
1.1	Burkitt-Lymphom (BL).....	- 7 -
1.2	Diffuses großzelliges B-Zell-Lymphom (englisch: <i>diffuse large B-cell lymphoma</i> )-	10 -
1.3	Unterschiede und Gemeinsamkeiten des Burkitt-Lymphoms (BL) und des diffusen großzelligen B-Zell-Lymphoms (DLBCL) .....	- 12 -
1.4	Anwendung der transkriptomweiten Genexpressionsanalyse in der molekularen Hämatopathologie .....	- 13 -
1.5	Genexpressionsanalysen des diffusen großzelligen B-Zell-Lymphoms (DLBCL) und des Burkitt-Lymphoms (BL).....	- 15 -
1.6	Fragestellung .....	- 16 -
1.6.1	Microarray-basierte transkriptomweite Genexpressionsanalysen von Burkitt-Lymphomen (BL) und diffusen großzelligen B-Zell-Lymphomen (DLBCL). -	16 -
1.6.2	Vergleich der Genexpressionsanalysen dieser Arbeit und der darüber hinausgehenden Ergebnisse eines Verbundprojektes der Deutschen Krebshilfe.....	- 16 -
1.6.3	Vergleich der Genexpressionsanalysen dieser Arbeit mit dem vormals beschriebenen Subgruppenklassifikator des diffusen großzelligen B-Zell-Lymphoms.....	- 16 -
<b>2</b>	<b>Material und Methoden</b> .....	- 17 -
2.1	Reagenzien .....	- 17 -
2.1.1	RNA-Isolierung.....	- 17 -
2.1.2	cDNA-Synthese.....	- 17 -
2.1.3	In-vitro-Transkription.....	- 17 -
2.1.4	cDNA- und cRNA-Aufreinigung.....	- 17 -
2.1.5	Hybridisierung.....	- 17 -
2.1.6	Färbereagenzien .....	- 18 -
2.1.7	Puffer .....	- 18 -
2.1.8	Microarrays .....	- 19 -
2.1.9	Software .....	- 19 -
2.1.10	Affymetrix Gerätekomponenten.....	- 19 -

2.1.11	Reagenzien für Immunhistologie .....	- 19 -
2.2	Lymphomgewebe .....	- 21 -
2.2.1	Auswahl des Lymphomgewebes .....	- 21 -
2.3	Immunhistochemische Färbungen.....	- 22 -
2.3.1	Anfertigen von Gewebeschnitten .....	- 22 -
2.3.2	Vorbehandlung der Gewebeschnitte .....	- 22 -
2.3.3	Alkalische Phosphatase Anti-Alkalische Phosphatase (APAAP)-Methode ....	- 23 -
2.3.4	EnVision-System.....	- 24 -
2.4	Präparation der Nukleinsäuren für die Genexpressionsanalyse .....	- 25 -
2.4.1	Aufreinigung von Gesamt-RNA aus Tumorgeweben.....	- 25 -
2.4.2	Quantifizierung der aufgereinigten Ribonukleinsäuren .....	- 25 -
2.4.3	Qualitätsanalyse der Gesamt-RNA im Bioanalyzer.....	- 26 -
2.4.4	Erststrangsynthese der cDNA .....	- 26 -
2.4.5	Zweitstrangsynthese .....	- 27 -
2.4.6	In-vitro-Transkription.....	- 27 -
2.4.7	Aufreinigung von cDNA und fluoreszenz-markierter cRNA .....	- 27 -
2.4.8	Fragmentierung der cRNA .....	- 28 -
2.5	Affymetrix GeneChip Instrument .....	- 28 -
2.5.1	Hochdichte Oligonukleotid-Microarrays .....	- 28 -
2.5.2	Human Genome Array U133 A von Affymetrix .....	- 30 -
2.5.3	Hybridisierung.....	- 30 -
2.5.4	Waschen und Färben .....	- 30 -
2.5.5	Scannen der GeneChips .....	- 31 -
2.6	Daten-Analyse .....	- 31 -
2.6.1	Analyse des einzelnen Microarrays und Qualitätskontrolle.....	- 31 -
2.6.2	GeneSpring-Software .....	- 32 -
2.6.3	Normalisierung.....	- 32 -
2.6.4	Filtern von Genen.....	- 33 -
2.6.5	One-way-analysis of variance (ANOVA) .....	- 34 -
2.6.6	Kreuzvalidierung und Klassenprädiktion.....	- 35 -
<b>3</b>	<b>Ergebnisse .....</b>	<b>- 37 -</b>

3.1	Charakterisierung der verwendeten Patientenproben.....	- 37 -
3.1.1	Tumorzellreichtum im Lymphomgewebe.....	- 37 -
3.1.2	Diagnose anhand morphologischer und immunologischer Kriterien der WHO-Klassifikation der malignen Lymphome .....	- 38 -
3.1.3	Genetische Analysen der Lymphomproben .....	- 39 -
3.1.4	Diagnostische Einordnung unter Verwendung aller WHO-Diagnosekriterien: Histologie, Immunologie und Genetik .....	- 40 -
3.2	Qualitätskontrolle der Probenaufarbeitung und der <i>Microarrays</i> .....	- 43 -
3.2.1	Analyse der Gesamt-RNA- und cRNA-Qualität der Lymphomproben .....	- 43 -
3.2.2	Qualitätskontrollen der <i>Microarrays</i> .....	- 44 -
3.3	Entwicklung eines 100 Markergene umfassenden Prädiktors und Vorhersage einer molekularen Diagnose.....	- 45 -
3.3.1	Einteilung der Lymphomproben in ein Training- und Test-Set.....	- 45 -
3.3.2	Identifikation von 2115 differentiell exprimierten Genen zwischen BL/BLL und DLBCL innerhalb des Training-Sets.....	- 46 -
3.3.3	Diagnosevorhersage der Lymphomproben mittels 100 differentiell exprimierten Genen (Markergenprädiktor).....	- 49 -
3.4	Der molekulare Burkitt-Lymphom (mBL)-Prädiktor – entwickelt in der Zusammenarbeit eines Verbundprojektes der Deutschen Krebshilfe .....	- 53 -
3.4.1	Der im Rahmen eines Verbundprojektes der Deutschen Krebshilfe entwickelte mBL-Prädiktor und Überschneidungen mit dem Markergenprädiktor der vorliegenden Arbeit .....	- 53 -
3.5	Der von Wright et. al. entwickelte Prädiktor zur Subgruppeneinteilung von DLBCL und sein Bezug zum Markergenprädiktor .....	- 54 -
3.6	Vergleich der drei <i>Microarray</i> -basierten Prädiktoren und Eigenschaften der diagnostisch vorhergesagten Lymphomproben.....	- 56 -
3.6.1	Markergenprädiktion: Eigenschaften der als Burkitt-Lymphom (BL) vorhergesagten Gewebeproben.....	- 56 -
3.6.2	Markergenprädiktion: Eigenschaften der als diffuses großzelliges B-Zell-Lymphom vorhergesagten Gewebeproben.....	- 57 -
3.6.3	Markergenprädiktion: Eigenschaften der als „nicht klassifizierbar“ vorhergesagten Gewebeproben .....	- 57 -

<b>4</b>	<b>Diskussion .....</b>	<b>- 60 -</b>
4.1	Unterschiede in der Entwicklung des Markergenprädiktors und anderer molekularer Prädiktoren .....	- 61 -
4.2	Molekular definierte Burkitt Lymphome (Markergen-BL) .....	- 62 -
4.3	Molekular definierte diffuse großzellige B-Zell-Lymphome (Markergen-DLBCL)-	64 -
4.4	Molekular nicht-klassifizierbare aggressive B-Zell-Non-Hodgkin Lymphome (nicht klassifizierbare B-NHL) .....	- 66 -
<b>5</b>	<b>Zusammenfassung.....</b>	<b>- 70 -</b>
<b>6</b>	<b>Abkürzungsverzeichnis.....</b>	<b>- 72 -</b>
<b>7</b>	<b>Literaturverzeichnis.....</b>	<b>- 74 -</b>
<b>8</b>	<b>Danksagung.....</b>	<b>- 80 -</b>
<b>9</b>	<b>Lebenslauf.....</b>	<b>- 81 -</b>
<b>10</b>	<b>Publikation.....</b>	<b>- 82 -</b>
<b>11</b>	<b>Erklärung.....</b>	<b>- 83 -</b>

# 1 Einleitung

## 1.1 Burkitt-Lymphom (BL)

Das Burkitt-Lymphom (BL) zählt zur Gruppe der aggressiven B-Zell-Non-Hodgkin-Lymphome (B-NHL) und wurde im Jahre 1958 erstmalig von Denis Burkitt als klinisches Syndrom beschrieben: Er beobachtete damals bei afrikanischen Kindern ein Krankheitsbild, das mit Tumoren im Bereich des Unterkiefers einherging und eine extrem hohe Letalität aufwies <sup>1</sup>. Nachdem O'Connors drei Jahre später die lymphatische Herkunft der Tumorzellen nachgewiesen hatte, wurde dieses klinische Syndrom erstmalig als BL bezeichnet <sup>2</sup>. Wie schon in früheren Klassifikationen wurde das BL auch in der WHO-Klassifikation lymphatischer Tumoren aus dem Jahre 2001 als eigene Tumorentität definiert. Dabei wurden drei epidemiologische Varianten unterschieden: Die endemische Form, die in etwa 95% der Fälle mit dem Epstein-Barr-Virus (EBV) assoziiert ist, stellt in den Äquatorregionen Afrikas und Südamerikas bei Kindern die häufigste neoplastische Erkrankung dar. Unabhängig von geographischen Faktoren hingegen ist die so genannte sporadische Form des BL in weniger als 30% der Fälle mit EBV assoziiert. An einer dritten Variante erkranken fast ausschließlich immunsupprimierte Menschen, weshalb auch von der Immundefizienz-assoziierten Form des BL gesprochen wird. Eine EBV-Assoziation liegt bei dieser Variante in ca. 25-40% der Fälle vor <sup>3</sup>. Unter Erwachsenen ist das BL insgesamt eine eher seltene Erkrankung und macht weniger als 1%, bei Kindern hingegen etwa 30% der B-NHL aus <sup>4</sup>.

Im lichtmikroskopischen Bild sind die Burkitt-Lymphomzellen etwa mittelgroß und zeigen typischerweise ein diffuses monomorphes Infiltrationsmuster. Im Zellkern findet man multiple basophile zentralständige Nukleoli. Das Zytoplasma ist stark basophil und weist in der Regel Lipidvakuolen auf. Charakteristisch ist weiterhin eine hohe Proliferationsrate (mit vielen mitotischen Figuren) und gleichzeitig finden sich zahlreiche apoptotische Zellen. Ein weiteres morphologisches Charakteristikum für das BL ist das so genannte Sternenhimmelbild: Dieses entsteht durch eine Vielzahl von zwischen den Tumorzellen gelegenen Makrophagen <sup>3</sup>. Beim Vorliegen dieser klassischen Morphologie wird häufig auch vom klassischen BL gesprochen. Demgegenüber existiert jedoch eine morphologische Variante des BL, bei der die Tumorzellen bezüglich Zellkerngröße und Form einen höheren Pleomorphismus aufweisen (meist als so genanntes atypisches oder Burkitt-like-Lymphom bezeichnet, kurz: BLL). Das BLL zeigt im morphologischen Bild häufig Überschneidungen mit dem diffusen großzelligen B-Zell-Lymphom <sup>3,5</sup> (vgl. auch Kapitel 1.3.).

Immunhistologisch exprimieren BL-Zellen die B-Zell-assoziierten Antigene CD19, CD20, CD22 und CD79a. Oberflächen-Immunglobuline (in den meisten Fällen IgM, seltener IgG oder IgA) sind in nahezu allen Fällen vorhanden. Die Expression der für Keimzentrums-B-Zellen typischen Marker CD10 und Bcl-6 weist auf eine Keimzentrumsherkunft der Tumorzellen hin. Die Wachstumsfraktion der BL-Zellen ist außerordentlich hoch: Nahezu 100% der Tumorzellen exprimieren das proliferationspezifische Ki-67-Antigen <sup>6</sup>.

Nahezu alle BL weisen charakteristischerweise eine reziproke chromosomale Translokation auf, bei der das C-MYC Onkogen (auch kurz MYC genannt) durch Juxtaposition zu einem der Immunglobulinloki transloziert wird. Dadurch gerät das MYC-Gen unter den Einfluss der Promotoren der jeweiligen Immunglobulingene auf den Chromosomen 14 (Lokus der Immunglobulin-Schwerkette, englisch *immunoglobulin heavy chain*, kurz: IgH), 22 (Lokus der Immunglobulin-Leichtkette lambda, kurz: Igλ) und 2 (Lokus der Immunglobulin-Leichtkette kappa, kurz: Igκ) und wird auf diese Weise konstitutiv überexprimiert. In 80% aller BL stellt der IgH-Lokus den Translokationspartner dar, was zu einer t(8;14)(q24;q34)-Formation führt <sup>7</sup>. In 15% der Fälle ist der Translokationspartner der Igκ-Lokus auf Chromosom 2p11, in den restlichen 5% der Fälle der Igλ-Lokus von Chromosom 22q11 <sup>8</sup>. MYC ist ein Helix-Loop-Helix Leucin-Zipper-Transkriptionsfaktor, der die Transkription einer Reihe von Genen reguliert, die in Zellzyklusregulation, Apoptose, Zelladhäsion und Zelldifferenzierung involviert sind. Hierbei ist die präzise Funktion von MYC komplex und letztlich unvollständig verstanden. Jedoch konnte in einer Vielzahl von Arbeiten gezeigt werden, dass die Expression von MYC in den meisten Zelltypen - eingenommen lymphatische Zellen - zu einer Zellzyklusprogression und Zellproliferation führt <sup>9-13</sup>. Adams et al. zeigten darüber hinaus anhand einer Serie von Experimenten, dass die Dysregulation von MYC eine wichtige Rolle bei der Lymphomgenese spielt: Das Einbringen eines MYC-Transgens (durch Kopplung des Gens mit dem μ Immunglobulin-Enhancer E) in das B-Zell-Kompartiment von Mäusen führte in den betroffenen Zellen zu einer hohen Inzidenz von B-Zell und Prä-B-Zell-Lymphomen <sup>14</sup>.

Das BL zeigt, wie alle anderen aggressiven B-NHL, einen rapiden klinischen Verlauf und führt unbehandelt innerhalb von Monaten zum Tode des Betroffenen. Erkrankte präsentieren sich häufig in fortgeschrittenen Stadien, was auf die hohe Wachstumsfraktion und Disseminierung der Tumorzellen zurückzuführen ist. Im Stagingssystem nach Murphy <sup>15</sup> befinden sich bei Diagnosestellung etwa 30% der Patienten in den Stadien 1 und 2 während 70% schon Stadium 3 und 4 erreicht haben. Das klinische Bild sowie die Tumorlokalisation variiert je nach klinischer Variante: Während sich das endemische BL initial in 50% der Fälle mit einem (extranodalen) Befall des Kiefers bzw. anderer Gesichtsschädelknochen manifestiert <sup>16</sup>, sind



beim sporadischen BL abdominale Lokalisationen - insbesondere im Bereich der Ileo-Zökal-Klappe - häufiger zu finden<sup>17</sup>. Die Immundefizienz-assoziierten BL manifestieren sich häufig nodal. Darüber hinaus gibt es eine Vielzahl anderer möglicher Manifestationsorte, die prinzipiell bei allen drei Varianten vorkommen können (Zökum, Omentum majus, Nieren u.v.a.)<sup>18, 19</sup>. Bei allen drei klinischen Varianten besteht die Gefahr einer ZNS-Beteiligung. Ein Befall der Ovarien bzw. der weiblichen Mammæ kommt insbesondere während der Schwangerschaft oder dem Beginn der Pubertät vor und könnte darauf zurückzuführen sein, dass BL Prolaktinrezeptoren exprimieren<sup>3</sup>. Leukämische Phasen können in fortgeschrittenen Stadien beobachtet werden: Die Diagnose einer Burkitt Leukämie (L3/ALL) mit Knochenmarkbefall und im Blut zirkulierenden Blasten ist insgesamt jedoch selten und mit einer schlechten Prognose assoziiert<sup>20</sup>.

Auf Grund der sehr hohen Wachstumsfraktion und der Aggressivität des Tumors kommt der schnellen Einleitung einer zytostatischen Therapie große Bedeutung zu: Die Chemotherapie des BL konnte in den letzten Jahren durch Intensivierung und Einführung neuer Medikamente deutlich verbessert werden. Insbesondere durch die Anwendung von hochdosiertem Methotrexat und Cytarabin in der Kombinationschemotherapie konnte in verschiedenen Studien bei der Mehrzahl der Patienten eine Heilung erreicht werden<sup>20, 21</sup>. Die Heilungsraten des sporadischen BL betragen bei Kindern bis 90% und bei Erwachsenen bis zu 70%. Falls Rezidive auftreten, werden sie in der Regel innerhalb des ersten Jahres nach Diagnosestellung beobachtet. Patienten, die nach zwei Jahren kein Rezidiv entwickelt haben, können als geheilt eingestuft werden<sup>20</sup>. Erste Ergebnisse einer Studie aus Texas zeigen, dass die zusätzliche Gabe von Rituximab - einem monoklonalen humanisierten Anti-CD20 Antikörper - zusammen mit einem Standardchemotherapie regime, die Prognose der erkrankten Patienten signifikant verbessern könnte<sup>22</sup>. Ergänzend zur intensiven Polychemotherapie ist die intrathekale Gabe von Chemotherapeutika zur Prophylaxe und Therapie eines ZNS-Befalls essentiell. Supportive Therapie ist für das Management von akuten Therapie-assoziierten Komplikationen von entscheidender Bedeutung<sup>23</sup>.

## 1.2 Diffuses großzelliges B-Zell-Lymphom (englisch: *diffuse large B-cell lymphoma*)

Die diffusen großzelligen B-Zell-Lymphome (englisch: *diffuse large B-cell lymphoma*; *DLBCL*) gehören wie die BL zu den aggressiven reifen B-NHL. Sie stellen eine klonale Proliferation großer neoplastischer B-Lymphozyten dar und sind unter den T- und B-Zell-NHL mit einem Anteil von 30-40% die häufigste Entität<sup>4</sup>. Diese Gruppe von Tumoren beinhaltet stark inhomogene morphologische, klinische, biologische, immunphänotypische, zytogenetische und molekulargenetische Gesichter. Das heterogene Erscheinungsbild der DLBCL wurde in der WHO-Klassifikation der malignen Lymphome von 2001 berücksichtigt, indem fünf verschiedene morphologische Varianten und drei Subtypen definiert wurden<sup>24</sup>: Die verschiedenen Varianten und Subtypen (morphologische Varianten: zentroblastisch, immunoblastisch, anaplastisch, T-Zell/Histiozytenreich, plasmoblastisch; Subtypen: mediastinal DLBCL, intravaskuläres DLBCL, primäres Ergusslymphom) entsprechen Lymphomen, deren Unterscheidung häufig nicht genügend reproduzierbar ist und die keine eigenständigen bzw. überlappende klinische und biologische Merkmale aufweisen. Trotz der Heterogenität gibt es zwischen den verschiedenen Varianten Gemeinsamkeiten, die sie als DLBCL kennzeichnen: Es ist allen DLBCL gemeinsam, dass ihre Zellkerne etwa doppelt so groß (oder größer) sind wie die Kerne normaler Lymphozyten. Immunphänotypisch exprimieren die Mehrzahl der DLBCL verschiedene Pan-B-Marker, wie CD19, CD20, CD22, und CD79a, von denen jedoch nicht immer alle vorhanden sind. Zellmembranständige und/oder zytoplasmatische Immunglobuline (IgM > IgG > IgA) können in 75% der Fälle nachgewiesen werden. Manche DLBCL exprimieren CD5 (10%) oder Keimzentrumsmarker wie CD10 und BCL6 (25-50%). BCL2-Expression findet man in etwa 30-50% der Fälle, während die Expression von Plasmazell-assoziierten Markern wie Syndecan (CD138) selten ist. Die Wachstumsfraktion, gemessen durch Ki-67-Positivität ist in der Regel höher als 40% und in einigen Fällen sogar höher als 90%<sup>24</sup>.

Genetisch zeichnen sich die DLBCL dadurch aus, dass sie in der Regel eine Vielzahl von (genetischen) Aberrationen aufweisen. Eine häufige Veränderung stellt die in etwa 20-30% der Fälle vorkommende Translokation des BCL2-Gens vom Chromosom 18 auf das Chromosom 14 dar, die charakteristischerweise im Follikulären Lymphom auftritt<sup>25</sup>. Es ist daher nahe liegend, dass es sich bei einem Teil der DLBCL mit t(14;18) um eine Transformation eines Follikulären Lymphoms in ein DLBCL handeln könnte. Bis zu 30% der DLBCL zeigen Aberrationen der 3q27-Region, die in der Regel zu Veränderungen des BCL6-Protoonkogens führt. In wenigen

Fällen findet man eine Translokation des MYC-Gens mit den Immunglobulin-Schwer- oder Leichtketten-Genen<sup>24, 26</sup>.

Patienten mit einem DLBCL präsentieren sich charakteristischerweise mit rapid progressiven Tumorherden in nodaler oder extranodaler Lokalisation, die häufig zu einer lokalen Verdrängungssymptomatik führen. Extranodale Erstmanifestationen des Tumors sind im Gegensatz zu vielen andern Lymphomen häufig und kommen in etwa 40% der Fälle vor<sup>6</sup>. Am häufigsten betroffen ist der Gastrointestinaltrakt. Lokalisationen wie Haut, ZNS, Knochen, Hoden, Weichteilgewebe, weiblicher Genitaltrakt, Lunge, Niere, Leber, Waldeyer'scher Rachenring und Milz kommen aber ebenfalls vor. Eine Primärpräsentation mit Knochenmarkbefall und Ausschwemmung von Tumorzellen ins periphere Blut ist dennoch selten<sup>24</sup>. DLBCL führen wie die BL unbehandelt schnell zum Tode, sind mit Polychemotherapie jedoch potentiell heilbar. Der therapeutische Goldstandard besteht in der Chemotherapie nach dem CHOP-Schema, die um die Gabe von Rituximab (vgl. Kapitel 1.1) erweitert wird. Hierdurch kann eine Langzeitremission bei der Mehrzahl der Patienten induziert werden, wobei sich jedoch bei etwa 30% der Erkrankten der Tumor resistent gegenüber der Therapie erweist oder innerhalb kurzer Zeit nach Therapieende ein Rezidiv auftritt<sup>27</sup>. Da die Ansprechraten auf die Therapie stark unterschiedlich sind, wird in der klinischen Praxis eine Risikostratifizierung mit dem *International Prognostic Index (IPI-Score)* durchgeführt. Basierend auf klinischen Parametern (Alter, Laktatdehydrogenase-Wert, Allgemeinzustand, Stadium nach Ann-Arbor, Anzahl der extranodalen Manifestationen) lässt sich hierdurch das Risiko des Patienten einschätzen<sup>28</sup>. Patienten mit hoch intermediärem Risiko (gemäß altersadaptiertem *IPI-Score*<sup>28</sup>) scheinen insgesamt von einer Hochdosischemotherapie mit anschließender autologer Stammzelltransplantation zu profitieren. In diesem Risikokollektiv konnten 5-Jahres-Überlebensraten von bis zu 74% erreicht werden<sup>29</sup>.

Aufgrund der morphologischen und klinischen Heterogenität dieser Erkrankung wurde vielfach postuliert, dass unter dem „Überbegriff“ DLBCL im Grunde mehrere biologisch und klinisch eigenständige Lymphomentitäten subsumiert sind<sup>24</sup>. Durch Genexpressionsanalysen von DLBCL-Gewebeproben mittels einer DNA-Chiptechnologie (*cdNA-Microarrays*) konnten in der Tat zwei bis drei große molekulare Gruppen identifiziert werden, die sich offenbar von unterschiedlichen Stadien der B-Zellentwicklung ableiten und sich auch prognostisch zu unterscheiden scheinen (Details siehe Kapitel 1.5).

### 1.3 Unterschiede und Gemeinsamkeiten des Burkitt-Lymphoms (BL) und des diffusen großzelligen B-Zell-Lymphoms (DLBCL)

BL und DLBCL bilden zusammen die Gruppe der aggressiven reifen B-NHL, die unbehandelt einen rapiden klinischen Verlauf zeigen. Bei geeigneter Therapie lassen sich aber beim BL möglicherweise bessere Erfolge erzielen als beim DLBCL. Gleichzeitig stellt die pathologisch-differentialdiagnostische Einordnung eine erhebliche Herausforderung dar: Obwohl in der WHO-Klassifikation der malignen Lymphome klare Diagnosekriterien für das BL und das DLBCL definiert worden sind, ist die Unterscheidung dieser beiden Entitäten nicht in allen Fällen möglich, da sich beide Lymphome in morphologischen, immunphänotypischen und genetischen Merkmalen überschneiden können. Wie in Kapitel 1.1 dargestellt ist die Abgrenzung anhand der Morphologie kein sicheres Unterscheidungskriterium: Während die Morphologie des BL in einigen Fällen charakteristisch ist (so genanntes klassisches BL), stellt das BLL (morphologische Variante des BL, vgl. Kapitel 1.1) im morphologischen Bild häufig einen fließenden Übergang zum DLBCL dar.

Die zusätzliche Einbeziehung des Immunphänotyps lässt eine sichere Unterscheidung der beiden Lymphome nicht immer zu, da die beiden Entitäten sich auch in dieser Hinsicht stark ähneln können (vgl. Tabelle-1). Gemäß der WHO-Klassifikation sollen BL/BLL gegenüber DLBCL im Mittel eine höhere Wachstumsfraktion aufweisen (gemessen mit dem Proliferationsmarker Ki-67 nahezu 100% vs. 40-90%), jedoch sind diese Angaben statistische Werte und können in Einzelfällen stark differieren. Auch die Expression von BCL2 wird immer wieder als Unterscheidungskriterium herangezogen, weil klassische BL dieses Molekül laut WHO-Klassifikation nicht, DLBCL jedoch in etwa 40% exprimieren sollen. Jedoch ist die Expression von BCL2 in den letzten Jahren immer häufiger auch in BLL beschrieben worden. Eine Arbeit suggeriert sogar, dass bis zu 15% der BLL das BCL2-Protein exprimieren<sup>30</sup>. Teilweise ist die Expression von BCL2 – im Widerspruch zur WHO-Klassifikation – sogar in klassischen BL beschrieben worden. Bislang ist kein immunphänotypisches Merkmal bekannt, das als spezifisch für eine Unterscheidung der beiden Lymphomentitäten gelten kann. Vielmehr können viele in der Routinediagnostik untersuchten Moleküle prinzipiell von beiden Lymphomentitäten exprimiert werden (vgl. Tabelle-1)<sup>30, 31</sup>. Obwohl bei der Differentialdiagnose maligner Lymphome zytogenetische Untersuchungen hilfreich sein können, stellen diese ebenfalls kein spezifisches Unterscheidungskriterium dar: Die MYC-Umlagerung gilt als charakteristisch für das BL, allerdings ist sie auch bei etwa 5-10% der DLBCL nachweisbar<sup>26, 32-34</sup>.

**Tabelle-1:** Häufigkeitsverteilung in der Expression von immunphänotypischen Markermolekülen zwischen BL/BLL und DLBCL

Markermolekül	BL/BLL (in %)	DLBCL (in %)
Ki-67	88-100*	53-100*
CD10	62-100	27-100
BCL6	70-88	75-86
BCL2	4-38	3-53

Angegeben ist die Anzahl positiver Fälle in Prozent jeweils als Variationsbreite der berechneten Mittelwerte aus den Arbeiten von Brazier, Frost und Cogliatti et. al.<sup>30, 31, 35</sup>. \*Mittlere Prozentzahl positiver Tumorzellen pro Fall.

Die oben ausgeführten Überschneidungen der diagnostischen Kriterien führen dazu, dass die Unterscheidung zwischen BL/BLL und den DLBCL zu den größten Herausforderungen in der Diagnostik von B-NHL zählt. Dies wurde eindrucksvoll durch eine Studie im Rahmen des *Non-Hodgkin Lymphoma Classification Project* belegt, in der eine internationale Gruppe auf Lymphomdiagnostik spezialisierter Hämatopathologen aggressive B-NHL-Proben mit Burkitt-ähnlicher Morphologie untersuchte: Die Experten kamen unter Berücksichtigung morphologischer, immunphänotypischer und genetischer Merkmale in lediglich 53% der Fälle dieser Kategorie zu einer übereinstimmenden Diagnosestellung<sup>4</sup>. Da die beiden Erkrankungen – wie oben dargestellt – jedoch unterschiedliche Behandlungsstrategien erfordern, könnte die falsche Diagnose zu inadäquater Behandlung von Patienten mit reifem aggressiven B-NHL führen.

#### **1.4 Anwendung der transkriptomweiten Genexpressionsanalyse in der molekularen Hämatopathologie**

Mit Hilfe so genannter *Microarrays* lässt sich der Expressionsstatus zehntausender Gene in Zellen oder Gewebeverbänden simultan ermitteln. *Microarray*-Analysen ermöglichen

inzwischen sogar die Untersuchung des gesamten Transkriptom (d.h. alle aktiven Gene) einer Zelle oder eines Zellverbandes in einem Versuchsansatz.

Dabei besteht das technische Prinzip der *Microarray*-Technologie darin, dass auf der festen *Microarray*-Oberfläche Sonden vorhanden sind, welche die Transkripte aller humanen Gene (ca. 33 000) repräsentieren. An diese kann eine fluoreszenzmarkierte Zell-/Gewebe-RNA hochspezifisch binden (Hybridisierung). Anschließend wird die Fluoreszenzintensität an jeder Sondenlokalisierung des *Microarrays* detektiert. Die gemessene Fluoreszenzintensität ist ein relatives Maß für die Expressionsstärke des entsprechenden Transkriptes.

Die Relevanz von Transkriptomanalysen für biologische Prozesse ergibt sich aus folgenden Überlegungen: In jeder Zelle wird nur ein Teil der etwa 33 000 humanen Gene aktiv in messenger-RNA (mRNA) transkribiert. Die Anzahl und die Art dieser mRNA-Transkripte für jedes Gen einer Zelle hängen von vielen Faktoren ab. Dazu gehören die Zellart, das Differenzierungsstadium, die Aktivität intrazellulärer Transduktionskaskaden, der Proliferationsstatus, der Einfluss extrazellulärer Stimuli u.v.a. Insgesamt spiegeln sich nahezu alle Unterschiede von Zelltypen und Zellzuständen in der Genexpression wider <sup>36</sup>. Außerdem führt der Weg zur Proteinexpression über die Transkription von Genen. Somit stellt die Genexpression eine Hauptdeterminante der biologischen Eigenschaften von gesunden und von malignen Zellen dar. Diese lässt sich – wie oben ausgeführt - durch *Microarray*-Messungen ermitteln.

Eine besondere Bedeutung kommt der *Microarray*-Technologie in der molekularen onkologischen Forschung zu: Zur malignen Entartung von Zellen reichen in der Regel Veränderungen einzelner Gene nicht aus. Vielmehr führen erst mehrere unterschiedliche genetische Aberrationen zu einer Dysregulation komplexer Transduktionskaskaden und damit zur Entwicklung eines malignen Tumors <sup>37</sup>. Da *Microarrays* eine simultane Analyse nahezu aller transkribierten Gene einer Zelle ermöglichen, lassen sich hierdurch neue Einblicke in die Biologie von Tumoren gewinnen. In verschiedenen Studien konnten mit Hilfe dieser Analysetechnik Gene identifiziert werden, die in der Pathogenese von Tumoren eine entscheidende Rolle spielen und mögliche Ziele für eine molekulare Therapie darstellen <sup>38-44</sup>. Des Weiteren ließ sich mehrfach zeigen, dass Gengruppen gefunden werden konnten, die mit der Krankheitsprognose stark korrelierten <sup>38, 40, 45-47</sup>. In Zukunft könnte auf diese Weise eine genaue Risikostratifizierung von Tumorpatienten möglich sein. Darüber hinaus wäre es möglich, neue biologisch relevante Klassifizierungen von Tumoren zu definieren, welche die klinische Heterogenität histologisch ähnlicher Tumoren erklären.

## 1.5 Genexpressionsanalysen des diffusen großzelligen B-Zell-Lymphoms (DLBCL) und des Burkitt-Lymphoms (BL)

Wie bereits in Kapitel 1.2 beschrieben, stellt die Gruppe der diffusen großzelligen B-Zell-Lymphome (DLBCL) eine in Morphologie, Immunphänotyp, Genetik und klinischem Verhalten heterogene Gruppe dar <sup>24</sup>. Diese Beobachtungen gaben vielfach Anlass zu der Frage, inwieweit diese Gruppe von Lymphomen unter Berücksichtigung biologischen und klinischen Verhaltens präziser subklassifiziert werden könnte <sup>6</sup>. Mittels *Microarray*-basierter transkriptomweiter Genexpressionsstudien wurden molekulare Subgruppen der *DLBCL* identifiziert, die sich im biologischen und klinischen Verhalten deutlich voneinander unterscheiden <sup>45, 47-53</sup>. Diese beiden neu identifizierten molekularen Subgruppen leiten sich offenbar von unterschiedlichen Stadien der B-Zellentwicklung ab. In der einen Gruppe fanden sich hohe Aktivitäten von Genen, die auch in Keimzentrums-B-Zellen typischerweise aktiv sind (GCB-Gruppe für englisch: *germinal center B-cell*). Die Genaktivitäten der anderen Gruppe ähnelten den von in-vitro aktivierten B-Zellen (ABC-Gruppe für englisch: *activated B-cell*). In beiden Gruppen scheinen außerdem unterschiedliche onkogene Mechanismen zur Tumorgenese zu führen: In der ABC-Gruppe hingegen fanden die Autoren eine Aktivierung von Genen des NFκB-Transduktionsweges <sup>39</sup>. Zwei in DLBCL häufige genetischen Aberrationen - die Translokation des BCL2-Protoonkogens (t(14;18) sowie die Amplifikation des c-rel-Lokus auf Chromosom 2p - wurden hingegen lediglich in der GCB-Gruppe nachgewiesen <sup>48</sup>. Darüber hinaus scheinen sich die beiden molekularen Gruppen auch klinisch zu unterscheiden: in der retrospektiven Analyse zeigten die Patienten in der GCB-Gruppe gegenüber den Patienten in der ABC-Gruppe ein signifikant besseres Überleben nach Chemotherapie (Gesamtüberleben nach CHOP-Therapie: GCB: 62% versus ABC: 26%; vgl. auch Kapitel 1.4) <sup>48</sup>. Einige DLBCL in der zitierten Studie ließen sich keiner der beiden Gruppen eindeutig zuordnen, da sie sowohl Merkmale der GCB- als auch der ABC-Gruppe trugen. Diese Fällen wurden deshalb als Gruppe-3-DLBCL bezeichnet <sup>48</sup>.

Bisher existieren keine transkriptomweiten Analysen der Genexpression von klassischen BL-Patientenproben. Lediglich die Untersuchung von Tumorzelllinien (Zellkultur) suggeriert, dass es einen engen biologischen Verwandtschaftsgrad zwischen BL-Zelllinien und DLBCL-Zelllinien vom GCB-Typ gibt. ABC-DLBCL-Zelllinien hingegen scheinen sich von den BL-Zelllinien deutlich zu unterscheiden <sup>54</sup>. Dieser Untersuchung widerspricht jedoch ein Nebenbefund einer Arbeit von Rosenwald et al., in der von 230 DLBCL- und 10 BLL-Gewebeproben *Microarray*-Analysen durchgeführt worden waren <sup>48</sup>: Die BLL bildeten hier

weder eine eigene Gruppe, noch ließen sie sich einer der neuen molekularen DLBCL-Subgruppen zuordnen. Sie verteilten sich diffus über die ABC- und die GCB-Gruppe<sup>48</sup>. Insbesondere wegen der geringen Fallzahl und der nicht durchgeführten Untersuchung von klassischen BL in dieser Studie bleibt die Abgrenzung der BL von den DLBCL unklar. Auch die biologische Beziehung zwischen BL und den neu identifizierten Subtypen des DLBCL bedarf weiterer Klärung. Diese Fragen sollten im Rahmen des Verbundprojektes der Deutschen Krebshilfe („Molekulare Mechanismen bei malignen Lymphomen“) – als dessen Teil diese Arbeit entstanden ist – untersucht werden.

## **1.6 Fragestellung**

### *1.6.1 Microarray-basierte transkriptomweite Genexpressionsanalysen von Burkitt-Lymphomen (BL) und diffusen großzelligen B-Zell-Lymphomen (DLBCL)*

Die zentrale (experimentelle) Aufgabe dieser Arbeit bestand in der Klärung der Frage, ob sich durch transkriptomweite *Microarray*-basierte Analysen von BL- und DLBCL-Gewebeproben differentiell exprimierte Gene identifizieren lassen, mit deren Hilfe die diagnostische Abgrenzung dieser beiden Lymphomentitäten möglich ist.

### *1.6.2 Vergleich der Genexpressionsanalysen dieser Arbeit und der darüber hinausgehenden Ergebnisse eines Verbundprojektes der Deutschen Krebshilfe*

Daneben sollten die - als Teil eines Verbundprojektes der Deutschen Krebshilfe - in dieser Arbeit generierten Ergebnisse aus Genexpressionsanalysen an 40 Lymphomproben mit weiterführenden Analysen des Gesamtverbundprojektes an zusätzlichen 180 Lymphomproben verglichen werden.

### *1.6.3 Vergleich der Genexpressionsanalysen dieser Arbeit mit dem vormals beschriebenen Subgruppenklassifikator des diffusen großzelligen B-Zell-Lymphoms*

Schließlich sollten die im Rahmen dieser Arbeiten generierten Daten mit denen von Wright et al. publizierten Genexpressionsanalysen zur Definition der molekularen Subgruppen des DLBCL verglichen werden. Dabei sollte insbesondere der Frage nachgegangen werden, in welcher Beziehung die vormals identifizierten molekularen Subgruppen des DLBCL zum BL stehen.



## 2 Material und Methoden

### 2.1 Reagenzien

#### 2.1.1 RNA-Isolierung

*RNeasy Mini Kit*

QIAGEN

#### 2.1.2 cDNA-Synthese

T7-(dT) 24-Primer

Operon, *HPLC purified DNA*

DEPC-H<sub>2</sub>O

Ambion

*5x First-Strand cDNA Reaction Buffer*

Invitrogen Life Technologies

0,1 M DTT

Invitrogen Life Technologies

10 mM dNTP-Mix

Invitrogen Life Technologies

*Reverse Transcriptase Superscript II*

Invitrogen Life Technologies

*5x Second Strand Reaction Buffer*

Invitrogen Life Technologies

E.Coli-DNA-Ligase

Invitrogen Life Technologies

E.Coli-DNA-Polymerase I

Invitrogen Life Technologies

E.Coli RNase H

Invitrogen Life Technologies

#### 2.1.3 In-vitro-Transkription

*Bio Array RNA Transcript Labeling Kit*

Enzo Life Sciences

*GeneChip Sample Cleanup Module*

Affymetrix

10x TBE

BioWhittaker

#### 2.1.4 cDNA- und cRNA-Aufreinigung

*GeneChip Sample Cleanup Module*

Affymetrix

#### 2.1.5 Hybridisierung

Kontrolloligonukleotid B2

Affymetrix

20x *GeneChip Eukaryotic Hybridization Control Kit*  
*Hering Sperm DNA*  
Acetyliertes bovines Serum-Albumin (50 mg/ml)  
2x Hybridisierungspuffer  
Aqua dest.

Affymetrix  
Promega Corporation  
Invitrogen Life Technologies  
siehe 1.1.7  
BioWhittaker

### 2.1.6 Färbereagenzien

2x MES Färbe-Puffer  
Acetyliertes bovines Serum-Albumin  
Ziegen-IgG  
Biotinylierter anti-Streptavidin-Antikörper (Ziege)  
Streptavidin-Phycoerythrin  
Aqua dest.

siehe 1.1.7  
Invitrogen Life Technologies  
Sigma-Aldrich  
Vector Laboratories  
Molecular Probes  
BioWhittaker

### 2.1.7 Puffer

2x Hybridisierungspuffer:

12x MES Stock  
5 M NaCl  
0,5 M EDTA  
10% Tween 20

Stringenter Wasch-Puffer:

12x MES Stock  
5 M NaCl  
10% Tween 20  
dest. H<sub>2</sub>O

Nicht-Stringenter Wasch-Puffer:

20x SSPE  
10% Tween 20  
dest. H<sub>2</sub>O

12x MES *Stock*-Puffer:

MES-freies, saures Monohydrat  
MES-Natrium-Salz  
Aqua dest.

### 2.1.8 *Microarrays*

*Human Genome Chip* U133A

Affymetrix

### 2.1.9 *Software*

*Microarray Suite* (MAS) Version 5.0

Affymetrix

*GeneSpring* Version 7.0

Silicon Genetics

### 2.1.10 *Affymetrix Gerätekomponenten*

Steuerungsrechner

Dell

Hybridisierungsöfen

Affymetrix

*Fluidics*-Station

Affymetrix

*Argon-Laser-Scanner*

Agilent

### 2.1.11 *Reagenzien für Immunhistologie*

Tris-Puffer:

50 mM Tris, pH 7,4-7,6  
150 mM NaCl (Roth)  
(bei Paraffinschnitten zusätzlich:  
0,1% Tween 20)

Citrat-Puffer:

0,1 M Zitronensäure-  
Monohydrat (Merck) pH=6,0

HuSe:

100 ml Biseko Humanserum  
(Biotest Pharma GmbH)

RPMI-Verdünnungsmedium:

50 ml 10x RPMI 1640 (Biochrom)

50 ml inaktiviertes Rinderserum  
(Biochrom oder Invitrogen)  
0,5 g Natriumazid (Merck)  
400 ml Aqua dest.  
pH 7,4-7,6 (Einstellung mit HCl oder NaOH)

Entwicklerlösung für Gefrierschnitte:

Lösung A:  
175 ml Entwicklungspuffer  
62,5 ml Propandiol.Lösung  
100 mg Levimasol (Sigma)

Lösung B:  
125 mg Naphthol-As-Biphosphat  
(Sigma)  
1,5 ml Dimethylformamid (Merck)

Lösung C:  
50 mg Natriumnitrit (Merck)  
1,25 ml Aqua dest.  
0,5 ml Neufuchsin-Lösung

Nach Lösung des Natriumnitrits wird das Neufuchsin hinzugegeben und eine Minute gemischt. Zum Ansetzen der Entwicklungslösung wird erst Lösung C, dann Lösung B zu Lösung A gegeben. Der pH-Wert wird mit HCl auf 8,8 eingestellt. Die Entwicklungslösung wird filtriert und die Präparate in der Küvette anschließend für 20 Minuten auf einem Schüttler in der Entwicklungslösung inkubiert. Die Entwicklungslösung darf nach dem Filtrieren höchstens 10 min stehen, bevor sie auf die Präparate gegeben wird.

Entwicklungspuffer:

0,15 M NaCl (Merck)  
0,1 M Tris-HCl (Sigma)  
0,04 M Tris-Base (Roth)  
Ad 1000 ml Aqua dest.

Propandiol-Lösung:	21 g Amino-2-methyl-1,3-propandiol (Merck) Ad 1000 ml Aqua dest.
Neufuchsin-Lösung:	5 g Neufuchsin (Merck) 100 ml 2 N HCl (Merck) pH 8,0
Hämalaun-Lösung:	50 g Kalilaun 1 g Hämatoxylin 0,2 NaJO <sub>3</sub> ( Merck)

In einem Liter Aqua dest. werden Kalilaun, Hämatoxylin und NaJO<sub>3</sub> gelöst und über Nacht verrührt. Danach werden 50 g Chloralhydrat (Merck) und 1 g Zitronensäure (Merck) dazugeben und erneut über Nacht verrührt.

## 2.2 Lymphomgewebe

Für diese Arbeit habe ich 40 Gewebeproben ausgewählt, welche Patienten im Zeitraum von 1984 bis 2003 entnommen worden waren und bei denen die Diagnose eines Burkitt-Lymphoms (BL), Burkitt-like-Lymphoms (BLL) oder diffusen großzelligen B-Zell-Lymphoms gestellt worden war und die eine ausreichende RNA-Qualität für eine *Microarray*-Hybridisierung aufwiesen. Für dieses Projekt lag ein positives Votum der zuständigen Ethikkommission vor. Die Gewebeauswahl erfolgte entsprechend dem unten dargestellten Algorithmus in drei Schritten.

### 2.2.1 Auswahl des Lymphomgewebes

In einem ersten Auswahlschritt habe ich von 294 archivierten Proben aggressiver B-NHL histologische Giemsa- und HE-Färbungen mikroskopiert und unter Anleitung erfahrener Pathologen das Gewebe bezüglich des Tumorzellreichtums beurteilt. Proben, die nach histomorphologischen Kriterien einen Tumorzellreichtum von < 70 Prozent aufwiesen, sortierte ich aus. In einem zweiten Auswahlschritt habe ich das kryokonservierte Gewebe von Fällen, die im Paraffinschnitt einen Tumorzellgehalt > 70% aufwiesen, untersucht: Zur exakten Evaluation des Tumorzellanteils habe ich Gefrierschnitte des Gewebes histologisch (Hämatoxylin + Eosin,

im Folgenden kurz „H&E“ genannt) und immunhistochemisch (Ki-67, BerMac3, CD3 und CD20) gefärbt. Zeigte sich im Gefriergewebe ein Tumorzellanteil < 70 Prozent, habe ich die Proben ebenfalls aussortiert. Von den übrig gebliebenen Gewebeproben habe ich anschließend die RNA-Qualität im *Agilent Bioanalyzer* überprüft (vgl. Kapitel 2.4.3). Nur Gewebeproben, die eine ausreichende RNA-Qualität und -Menge für die *Microarray*-Analyse aufwiesen, habe ich in die Studie eingeschlossen.

Der dritte Auswahlschritt umfasste die Re-Evaluierung der ursprünglich erstellten Diagnose in einer gemeinsamen Sitzung der Direktoren der deutschen Referenzzentren für Lymphknotenpathologie aller in die Studie eingeschlossenen Fälle. Für die verwendeten Gewebeproben wurde eine Konsensusdiagnose festgelegt, der alle beteiligten Pathologen zustimmten (detaillierte Darstellung im Ergebnisteil Kapitel 1.1.2). Die diagnostische Einordnung erfolgte entsprechend den histomorphologischen und immunphänotypischen Kriterien der aktuellen WHO-Klassifikation der malignen Lymphome aus dem Jahre 2001. Zur Evaluation des Immunphänotyps wurde ein Panel von immunhistochemischen Färbungen durchgeführt (Färbepanel siehe Tabelle-2, vgl. auch Kapitel 3.1).

## **2.3 Immunhistochemische Färbungen**

### *2.3.1 Anfertigen von Gewebeschnitten*

Zur Herstellung von Gefrierschnitten wurden Gefrierblöcke bei  $-30^{\circ}\text{C}$  mit Hilfe eines Kryostaten (Reichert-Jung, 2800 Frigocut E) geschnitten. Die Schnittdicke betrug jeweils  $10\ \mu\text{m}$ . Das Erstellen von Paraffinschnitten erfolgte bei Raumtemperatur unter Verwendung eines Microtoms (Microm GmbH, Heidelberg, Typ 4800). Alle Gewebeschnitte wurden auf beschichtete Objektträger (DAKO ChemMate Code No.S2024) aufgebracht.

### *2.3.2 Vorbehandlung der Gewebeschnitte*

Gefrierschnitte stellte ich zur Dehydrierung 10 min in Aceton und trocknete sie anschließend drei Minuten unter einer Abzugshaube. Bis zur weiteren Bearbeitung bewahrte ich die Schnitte in Kartellboxen bei  $-80^{\circ}\text{C}$  auf. Um die Entstehung von Kondensationsniederschlägen auf den Schnitten beim Einfrieren und Auftauen zu vermeiden, gab ich in Filterpapier gewickeltes kristallines Calciumchlorid-2-Hydrat (Merck 2382) in die Kartellbox.

Paraffinschnitte habe ich nach dem Schneiden für 12 Stunden bei 68°C getrocknet. Zur Entparaffinierung behandelte ich die Paraffinschnitte zunächst 10 min mit Xylol und brachte sie danach in eine absteigenden Alkoholreihe (96% Ethanol, 90% Ethanol, 70% Ethanol, 0% Ethanol in reinem Tris-Puffer) ein. Für die Proteindemaskierung, die erst eine Bindung der Antikörper an entsprechende Zielepitope ermöglicht, kochte ich die Schnitte in 0,1 M Citratpuffer (pH = 6,0, Kochdauer siehe Tabelle-2).

**Tabelle-2:** Verwendete Antikörper und Vorbehandlung der Gewebeschnitte

Antigen/Antikörper	Herkunft	Ausgangsmaterial	Verdünnung	Vorbehandlung der Schnitte
CD3	DAKO	Paraffinmaterial	1:40	2 Minuten Citrat
CD20/L26	DAKO	Paraffinmaterial	1:100	2 Minuten Citrat
Ki-67	DAKO	Gefriermaterial	1:100	30 Minuten Aceton
CD68/PGM1	DAKO	Paraffinmaterial	1:10	2 Minuten Citrat
CD10	Novocastra	Paraffinmaterial	1:25	2 Minuten Citrat
CALLA	DAKO	Gefriermaterial	1:25	30 Minuten Aceton
Bcl-6	DAKO	Paraffinmaterial	1:2	5 Minuten Citrat
Bcl-2	DAKO	Paraffinmaterial	1:25	2 Minuten Citrat
BerMac3	Eigenprodukt des Instituts	Gefriermaterial	1:20	10 Minuten Aceton

### 2.3.3 Alkalische Phosphatase Anti-Alkalische Phosphatase (APAAP)-Methode

Immunhistologische Färbungen mit der APAAP-Methode ermöglichen den einzelzellbezogenen Nachweis eines Antigens in Gewebeschnitten. Hierbei wird die spezifische Bindung eines Antikörpers durch eine Farbreaktion nachgewiesen. Der Nachweis des Zielepitops erfolgt dabei in drei Schritten: Ein Primärantikörper bindet spezifisch an ein Zielantigen auf den Zellen im Gewebe. Danach wird ein Sekundärantikörper (Brückenantikörper) appliziert, der einerseits an die Fc-Teile des Primärantikörpers und andererseits auch an den danach applizierten Enzym-Anti-Enzymkomplex (APAAP-Komplex) bindet. Primärantikörper und APAAP-Komplex stammen dabei von dergleichen Spezies (Maus), während der Brückenantikörper von einer

anderen Spezies (Kaninchen) stammt. Nach Substratzugabe wird ein Farbstoff enzymatisch umgesetzt, der im Gewebeschnitt zu einem lokalen Farbniederschlag führt.

Zur immunhistochemischen Färbung der Paraffinschnitte löste ich Primär- und Brückenantikörper in einem speziellen Antikörperverdünnungsmedium (*ChemMate Antibody Diluent*, DAKO) (Verdünnungsverhältnis siehe Tabelle-2). Den Brückenantikörper (Kaninchen anti-Maus, DAKO) verdünnte ich danach zusätzlich in einem Verhältnis von 1:20 mit einem RPMI-Humanserum (HuSe)-Gemisch (100 ml HuSe + 900 ml RPMI), den APAAP-Komplex im Verhältnis 1:10 mit RPMI. Die Bindung des Primärantikörpers erreichte ich durch 30-minütige Inkubation von 200-300 µl der Verdünnungslösung auf dem Schnitt. Anschließend spülte ich die Schnitte mit Tris-Puffer, um überschüssigen nicht gebundenen Primärantikörper zu entfernen. Die Applikation des Brückenantikörpers führte ich ebenfalls 30 min mit 200-300 µl der entsprechenden Lösung und nachfolgendem Waschschrift durch. Die Schnitte inkubierte ich anschließend 10 min mit 300 µl des APAAP-Komplexes und spülte sie gründlich ab. Brückenantikörper und APAAP-Komplex inkubierte ich nachfolgend mehrmals jeweils 10 Minuten zur Signalverstärkung. Auf diese Weise binden mehrere Brückenantikörper an einen Primärantikörper und vermitteln dadurch eine stärkere Farbreaktion. Die Entwicklerlösung für die Farbreaktion verwendete ich in dem vom Hersteller empfohlenen Ansatz (*ChemMate Detection Kit APAAP Mouse*, DAKO). Zur besseren Kontrastdarstellung im mikroskopischen Bild führte ich eine Gegenfärbung der Kernstrukturen mit Hämatoxylin durch. Unter Anwendung von lauwarmem Wasser wusch ich den überschüssigen Farbstoff ab. Zur Eindeckung der Gewebeschnitte verwendete ich auf 50-60°C erwärmte Kaisers Glycerin-Gelatine (Merck).

Da Gefriermaterial im Unterschied zu Paraffingewebe nicht fixiert wird, dürfen für die immunhistochemischen Färbungen keine detergentshaltigen Lösungen verwendet werden. Daher habe ich für die Färbungen des Gefriergewebes die Antikörper in RPMI-Verdünnungsmedium verdünnt und auf die aufgetauten Schnitte gemäß den oben beschriebenen Inkubationszeiten aufgebracht. Für die Waschschriffe nutzte ich ebenfalls Tween-freien Tris-Puffer. Darüber hinaus wurde eine spezielle Entwicklerlösung verwendet (vgl. Kapitel 2.1.11)

#### 2.3.4 *EnVision-System*

Die Färbungen mit den Primärantikörpern gegen die Antigene IgM, IgD und IgG habe ich mit der von der Firma DAKO patentierten Polymer-Konjugat-Methode *EnVision* durchgeführt. Diese basiert auf der Kopplung von Sekundärantikörpern und Detektionskomponenten an einen



Polymertorso: An ein Dextranpolymer sind sowohl ein Sekundärantikörper, als auch eine große Anzahl an Peroxidaseenzymen gekoppelt. Nach Applikation eines Primärantikörpers (Maus) können auf diese Weise Sekundärantikörper (Kaninchen, anti-Maus) und Detektionssystem gleichzeitig aufgetragen werden. Zur Visualisierung der Farbreaktion im Gewebe muss anschließend eine entsprechende Substrat-Chromogen-Lösung appliziert werden.

Für die Färbungen mit dem *EnVision*-System behandelte ich die Schnitte nach entsprechender Epitopdemaskierung (vgl. Kapitel 2.3.2 sowie Tabelle-2) mit einer speziellen Lösung (*Peroxidase Blocking Solution*, DAKO) zur Blockade der endogenen Peroxidase gemäß den Anweisungen des Herstellers. Danach inkubierte ich die Schnitte 10 Minuten mit dem Primärantikörper. Die Applikation des Dextranpolymers sowie der Chromogen-Substratlösung führte ich mit Hilfe des *ChemMate* DAKO *EnVision Detection Kit*, Peroxidase/DAB, Kaninchen/Maus (K5007) unter Verwendung des *TechMate 500 Plus Autostainers* der Firma DAKO gemäß den Anweisungen des Herstellers durch.

## **2.4 Präparation der Nukleinsäuren für die Genexpressionsanalyse**

### *2.4.1 Aufreinigung von Gesamt-RNA aus Tumorgeweben*

Aus Gefrierschnitten gewann ich Gesamt-RNA mit Hilfe des *RNeasy Mini Kits* der Firma QIAGEN entsprechend den Anweisungen des Herstellers. Pro Patientenfall brachte ich 5-7 Schnitte von jeweils 20 µm Dicke aus kryokonserviertem Lymphomgewebe in Lysepuffer ein und homogenisierte mit Hilfe des *QIAshredder-Kits* (QIAGEN). Um Kontaminationen mit RNasen zu vermeiden, trug ich während der Versuchsdurchführung Handschuhe, dekontaminierte den Arbeitsplatz und alle verwendeten Geräte mit *RNase Zap* (Ambion) und verwendete ausschließlich RNase-freie Gefäße.

### *2.4.2 Quantifizierung der aufgereinigten Ribonukleinsäuren*

Die Quantifizierung der aufgereinigten Gesamt-RNA erreichte ich durch photometrische Messung anhand der Absorption bei 260 nm. Eine optische Dichte von 1 ( $OD_{260nm}$ ) entspricht dabei einer RNA-Konzentration von 40 µg/ml bei einer Schichtdicke der Küvette von 1 cm.

Um die Reinheit der Nukleinsäurepräparationen abzuschätzen maß ich zusätzlich die Absorption bei 280 nm. Der Quotient  $OD_{260}/OD_{280}$  gibt Aufschluss über die Verunreinigung mit Proteinen und sollte größer als 1,6 sein.

### 2.4.3 Qualitätsanalyse der Gesamt-RNA im Bioanalyser

Mit Hilfe des *Agilent 2100 Bioanalyser* unterzog ich die aufgereinigte RNA einer Qualitätsanalyse. Das Prinzip dieser Messung beruht auf einer kapillarelektrophoretischen Auftrennung von Nukleinsäuren, wobei hierfür erforderliche elektrokinetische Kräfte mit Hilfe von Pin-Elektroden aufgebaut werden. Qualitativ gute RNA zeigt sich im elektronischen Elektropherogramm in Form zwei deutlich abgrenzbarer Spitzen (*Peaks*), welche die ribosomale 18S- und 28S-RNA widerspiegeln. Die *Peak*-Flächenratio von 28S- zu 18S-*Peak* sollte für qualitativ hochwertige RNA ungefähr bei 2 liegen. Degradation der Gesamt-RNA durch potentiell vorhandene RNasen zeigt sich durch Anwesenheit von kleineren Fragmenten und vermindertem Fluoreszenzsignal der ribosomalen *Peaks*. Da die messenger-RNA (mRNA) nur einen sehr kleinen Teil der Gesamt-RNA einer Zelle ausmacht, wird ihr Signal im Bioanalyser durch das Signal der ribosomalen RNA überlagert und kann deshalb nicht direkt beurteilt werden. Die Qualität der ribosomalen RNA ist jedoch ein gutes Maß für die Qualität der mRNA.

### 2.4.4 Erststrangsynthese der cDNA

Als Ausgangsmaterial für die cDNA-Synthese dienten 5-8 µg Gesamt-RNA. Zur Erststrangsynthese, bei der ein RNA-DNA-Hybrid entsteht, verwendete ich das *Superscript Kit* von Invitrogen entsprechend den Anweisungen des Herstellers. Für die *Primer*-Hybridisierung gab ich 1 µl der T7-(dT)<sub>24</sub>-Primers (100 pmol/µl; Sequenz: 5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT)<sub>24</sub>-3') zu 11 µl RNA-Lösung, inkubierte für 10 min bei 70° C und stellte den Ansatz anschließend auf Eis. Nach Zugabe von 4 µl *First Strand Reaction Buffer* (Endkonzentration 1x) sowie 2 µl DTT (Endkonzentration 10 mM) und 1 µl dNTP-Mix (Endkonzentration 500 µM) zu den 11 µl der mit dem Primer hybridisierten RNA habe ich den Ansatz bei 42° C 2 min auf die gewünschte Temperatur gebracht. Die *Reverse Transkriptase Superscript II* pipettierte ich danach hinzu und inkubierte den Reaktionsansatz bei 42° C für eine Stunde.

#### 2.4.5 Zweitstrangsynthese

Bei der Zweitstrangsynthese wird aus dem RNA-DNA-Hybrid unter Abbau der RNA ein doppelsträngiges cDNA-Molekül synthetisiert. Dabei fügt eine RNase H kleine Strangbrüche in die RNA ein, in die sich die DNA Polymerase I einlagern und gleichzeitig mit Hilfe ihrer Polymerase- und Exonucleaseaktivität den zweiten cDNA-Strang unter Abbau des RNA-Stranges synthetisieren kann. Den Ansatz habe ich wie folgt hergestellt: Zu 91 µl DEPC-Wasser gab ich nacheinander 30 µl 5x *Second Strand Reaction Buffer* (Endkonzentration 1x), 3 µl dNTP-Mix (Endkonzentration 200 µM), 1 µl E.coli DNA-Ligase (10 U/µl), 4 µl E.coli DNA-Polymerase I (10 U/µl) und 1 µl RNase H (2 U/µl). Zu diesen 134 µl pipettierte ich den gesamten Ansatz aus der Erststrangsynthese hinzu und inkubierte bei 16° C im Wasserbad für zwei Stunden. Um sicher zu gehen, dass der zweite Strang die gesamte *T7-Promoter*-Sequenz enthält, vervollständigte ich die entsprechende Sequenz am 3'-Ende des zweiten cDNA-Stranges durch Zugabe einer T4-DNA-Polymerase. Hierzu inkubierte ich den Ansatz erneut für 5 Minuten bei 16° C und stoppte anschließend die Reaktion durch Zugabe von 10 µl 0,5 M EDTA.

#### 2.4.6 In-vitro-Transkription

Für die in-vitro-Transkription wurde der *Bio Array Transcript Labeling Kit* der Firma Enzo Life Sciences verwendet. Hierbei wird die doppelsträngige cDNA unter Einbau von Biotin markierten Nukleotiden und 30-100facher linearer Amplifikation in cRNA umgeschrieben. Der Ansatz setzt sich wie folgt zusammen: Zu 22 µl aufgereinigter cDNA habe ich nacheinander 4 µl HY Reaction Buffer (10x), 4 µl Biotin markierte Nukleotide (10x), 4 µl DTT (10x), 4 µl *RNase Inhibitor Mix* (10x) und 2 µl T7 RNA-Poymerase hinzugegeben (20x). Den Ansatz inkubierte ich danach für 5 Stunden bei 37° C im Wasserbad, wobei ich das Gemisch alle 30 Minuten für 10 Sekunden schüttelte und abzentrifugierte.

#### 2.4.7 Aufreinigung von cDNA und fluoreszenz-markierter cRNA

Die doppelsträngige cDNA sowie die cRNA reinigte ich nach den entsprechenden Syntheseschritten unter Verwendung des *GeneChip Sample Cleanup Moduls* (Affymetrix/QIAGEN) entsprechend den Anweisungen des Herstellers auf. Die Ausbeute quantifizierte ich gemäß des oben genannten photometrischen Verfahrens (vgl. Kapitel 2.4.2). Im

Anschluss unterzog ich die aufgereinigte cRNA einer Qualitätsanalyse im Agilent Bioanalyzer (vgl. Kapitel 2.4.3).

#### 2.4.8 Fragmentierung der cRNA

Bevor die markierte cRNA auf die *Microarrays* hybridisiert werden kann, muss sie fragmentiert werden. Hierfür sind Fragmente mit einer Länge von 35-200 Basenpaaren erforderlich. Diese Fragmentlänge ermöglicht eine bessere Hybridisierung der cRNA auf dem *Microarray*.

Zur Fragmentierung gab ich zu 32 µl aufgereinigter cRNA (Endkonzentration zwischen 0,5 und 2 µg/µl) 8 µl Fragmentierungspuffer (5x) der Firma QIAGEN. Den Ansatz inkubierte ich 35 Minuten bei 94°C und stellt ihn anschließend nach kurzer Zentrifugation auf Eis. Den Fragmentierungserfolg überprüfte ich anschließend durch erneute Auftragung der Proben im *Agilent 2100 Bioanalyzer*.

## 2.5 Affymetrix GeneChip Instrument

### 2.5.1 Hochdichte Oligonukleotid-Microarrays

Mit hochdichten Oligonukleotid-*Microarrays* lassen sich Genexpressionsanalysen nahezu transkriptomweit durchführen: In einem Versuchsansatz können somit die Expressionsstärken zehntausender Gene eines Gewebes gleichzeitig gemessen werden. Als Gen ist in diesem Kontext die Sequenz eines fertig gespleißten mRNA-Transkriptes zu verstehen. *Intron*- und *Promoter*-Sequenzen und andere regulatorische Einheiten der DNA, die nach heutiger Definition Bestandteile eines Gens sind, werden in diesem funktionellen Ansatz nicht erfasst. Außer gut annotierten Genen werden auch so genannte *ESTs* (*expressed sequence tags*) untersucht. Es handelt sich hierbei um kurze cDNA-Sequenzen von Transkripten, deren genaue Zuordnung zu der entsprechenden „Mutter-DNA-Sequenz“ nicht bekannt ist.

Das Grundprinzip von Oligonukleotid-*Microarrays* (synonym werden die Begriffe *GeneChip* oder *Array* verwendet) besteht in einer Umkehrung der *Nothern-Blot*-Technik: Auf einer *Chip*-Matrix sind genspezifische kurze Oligonukleotide immobilisiert, die als Sonden fungieren. Eine auf die Matrix applizierte fluoreszenzmarkierte Proben-RNA kann mit diesen Oligonukleotiden hybridisieren. Je stärker die Expression eines Gens im Gewebe ist, desto größer ist die Menge an markierter Proben-cRNA, die zur Sondenbindung (Oligonukleotide) zur Verfügung stehen und um so stärker ist – nach Anregung im Laser-Licht – auch das zu detektierende Fluoreszenzsignal.

Die Expression von Genen lässt sich auf diese Weise relativ quantifizieren. Voraussetzung für diese relative Quantifizierung ist allerdings die Linearität in der Herstellung und Amplifikation markierter cRNA aus der totalen RNA eines Gewebes. Während die Syntheseschritte zur Erst- und Zweitstrangsynthese eine 1:1-Umschreibung darstellen, ist die in-vitro-Transkription ein linearer Amplifikationsschritt<sup>55</sup>.

Die enorm hohe Sondendichte (vgl. Kapitel 2.5.2) der Oligonukleotid-*Microarrays* wird durch ein besonderes Herstellungsverfahren erreicht, das die Firma Affymetrix entwickelt hat. Hierbei wird die Sondensequenz direkt auf dem *Microarray* synthetisiert. Dies geschieht durch eine Kombination aus Nukleotidsynthese und Photolithographie. Die Technik der Photolithographie ist ursprünglich der Halbleiterindustrie entlehnt und erlaubt die schrittweise Synthese von 25-mer-Oligonukleotiden, Base für Base, an definierten Positionen auf dem Affymetrix-*Microarray* (*GeneChip*). Auf der 1,28 x 1,28 cm großen *GeneChip*-Matrix sind auf diese Weise 300 000 unterschiedliche Sondensequenzen immobilisiert, die jeweils in einer Kopienzahl von  $1 \times 10^6$  bis  $1 \times 10^7$  vorliegen. Da solch kurze Sequenzen alleine selten hochspezifisch für ein bestimmtes Transkript sind, verwendete die Firma Affymetrix pro Transkript 11 unterschiedliche 25-mer-Oligonukleotide (so genanntes *probe set*), die aus unterschiedlichen Bereichen der Zielsequenz abgeleitet sind. Diese sind in ihrer Gesamtheit hochspezifisch für das entsprechende Gen/Transkript. Die Hybridisierungsspezifität jedes der 11 Oligonukleotide eines Transkriptes kann jedoch sehr unterschiedlich sein. Deshalb entwickelte Affymetrix ein System, mit dessen Hilfe die hochspezifischen Oligonukleotide gegenüber den weniger spezifischen Oligonukleotiden bei der Ermittlung des Expressionswertes für das Gesamttranskript stärker gewichtet werden können: Für jedes der 11 Oligonukleotide existiert auf dem *Microarray* noch ein „Partner“-Oligonukleotid. Während die 11 Oligonukleotide des Paares gänzlich der abgeleiteten Sequenz entsprechen (*Perfect-Match*-Oligonukleotide, PM), tragen die 11 „Partner“-Oligonukleotide jeweils eine Austauschbase an Position 13 (*Mismatch*-Oligonukleotide, MM), die zu einer nicht optimalen Hybridisierung führt. Die Differenz aus der PM- und der MM-Intensität stellt ein Signifikanzmaß für die Hybridisierungsspezifität jedes einzelnen Oligonukleotidpaares dar. Je nach Signifikanz werden die Oligonukleotidpaare in der Berechnung des Signalwertes für das zugehörige Transkript unterschiedlich gewichtet (vgl. Kapitel 2.6.1).

### 2.5.2 Human Genome Array U133 A von Affymetrix

Die Genexpressionsanalysen in dieser Studie habe ich ausschließlich mit dem Human Genome Array U133A der Firma Affymetrix durchgeführt. Mit diesem *Microarray* können 22.238 Transkripte detektiert werden. Die Sequenzen, nach denen synthetisiert wurde, entstammen der *UniGene-DataBase* (Built 133, 20.04.2001). Diese Sequenzen wurden mit Informationen zweier weiterer Datenbanken abgeglichen und ergänzt („*Washington University EST Trace Reposistory*“ und „*University of California, Santa Cruz Golden Path Human Genome Data Base*“).

### 2.5.3 Hybridisierung

Für die Hybridisierungslösung gab ich zu 15 µg fragmentierter cRNA (Endkonzentration 0,05 µg/µl) schrittweise 5 µl B2-Kontrolloligonukleotide (Endkonzentration: 50 pM), 20x Eukaryotische Hybridisierungs-Kontrollen (*Spike*-Kontrollen bioB, bioC, bioD, cre; Endkonzentration: 1,5 pM, 5 pM, 25 pM, 100 pM), Hering Sperm DNA (Endkonzentration: 0,1 mg/ml), Acetyliertes bovines Serum-Albumin (Endkonzentration: 0,5 mg/ml) und 2x Hybridisierungspuffer (Endkonzentration: 1x). Den Ansatz füllte ich danach mit DEPC-Wasser auf 300 µl auf, mischte ihn, und inkubierte das Gemisch bei 99°C 5 min und danach direkt 5 min bei 45°C. Anschließend zentrifugierte ich den Ansatz 5 min bei maximaler Drehzahl (*Heraeus Biofuge Pico*, 1350 x g), um Schwebstoffe und unlösliche Partikel zu sedimentieren. Nach Äquilibrieren der *Microarrays* auf Raumtemperatur, füllte ich sie vor der Hybridisierung mit 200 µl des 1x Hybridisierungspuffers und inkubierte sie 10 min bei 45°C und 60 Umdrehungen pro Minute im Hybridisierungs-ofen. Nach Entnahme des 1x Hybridisierungspuffers füllte ich die *Microarrays* mit jeweils 200 µl der Hybridisierungslösung und inkubierte sie 16 Stunden bei 45°C und 60 Umdrehungen pro Minute über Nacht.

### 2.5.4 Waschen und Färben

Nach 16-stündiger Inkubation entnahm ich die Hybridisierungslösung aus den *GeneChips* und füllte sie mit jeweils 250 µl nicht-stringentem Wasch-Puffer A. Anschließend lud ich die *Microarrays* in die *Fluidics*-Station (Waschstation) und unterzog sie dem automatischen Wasch- und Fluoreszenzmarkierungsprotokoll EukGEWS2V4. Durch die dabei alternierende Inkubation mit einer Streptavidin-Phycoerithrin (SAPE)-Lösung [600 µl 2x MES-Färbe-Puffer (Endkonzentration: 1x), 48 µl acetyliertes BSA (Endkonzentration: 2 mg/ml), 12 µl Streptavidin-

Phycoerithrin (Endkonzentration: 10 µg/ml)] und einer Antikörperlösung [300 µl 2x MES-Färbepuffer (Endkonzentration: 1x), 24 µl acetyliertes bovines Serum-Albumin (Endkonzentration: 2 mg/ml), 6 µl normaler Ziege-IgG (Endkonzentration: 0,1 mg/ml), 3,6 µl biotinylierter Antikörper (Endkonzentration: 3 µg/ml), 266 µl Aqua dest.] sowie zwischengeschalteten Waschschritten mit den Puffern A und B wird eine doppelte Fluoreszenzmarkierung erreicht. Hierbei bindet das Streptavidin-Phycoerythrin sowohl an die biotinylierten Nukleotide der auf dem Chip gebundenen cRNA als auch an die applizierten biotinylierten Antikörper. Nach der Markierung wurden die *GeneChips* zum Scannen automatisch mit 250 µl Wasch-Puffer A gefüllt.

### 2.5.5 Scannen der *GeneChips*

Nach dem Waschen lud ich die Arrays in den konfokalen Argon-Laser-Scanner, der das Fluoreszenzlabel bei einer Wellenlänge von 488 nm anregt und das emittierte Licht bei einer Wellenlänge von 570 nm und einer Auflösung von 3 µm misst. Die *Scanning*-Prozedur wird automatisch zweimal durchgeführt und der mittlere Lichtintensitätswert als endgültiger Wert gespeichert.

## 2.6 Daten-Analyse

### 2.6.1 Analyse des einzelnen *Microarrays* und *Qualitätskontrolle*

Der mit Hilfe des Scanners gemessene Lichtintensitätswert jedes Oligonukleotides wird durch die *Microarray-Suite*-Software 5.0 automatisch in eine 16-bit Bildinformation übersetzt, in der die Lichtintensität in Form einer Farbkodierung dargestellt ist. Das entstehende Gesamtbild stellt die Rohdaten des jeweiligen Experimentes dar und wird in Form des so genannten CHP-Files gespeichert.

Die erste Überprüfung jedes einzelnen Arrays wird mit Hilfe der von Affymetrix selbst entwickelten Software *Microarray-Suite* 5.0 (MAS 5.0) weitgehend automatisch durchgeführt. Die fundamentalen Analyseschritte waren im folgenden die Einteilung des Bildes in ein Raster, die Korrektur des Hintergrundsignals, die Skalierung der mittleren Signalintensität des *Microarrays* auf einen Zielsignalwert zur Verringerung der artifiziellen Variabilität (*Per-Chip*-Normalisierung, siehe Kapitel 2.6.3) und die statistische Berechnung des singulären Signals für jedes Transkript auf dem Chip.

Dieses singuläre Signal (so genannter „*detection-p-value*“) ist ein statistischer Wert, der ein relatives Maß für die Expressionsstärke eines Transkriptes darstellt. Es beschreibt die Wahrscheinlichkeit, mit der ein Transkript von den zugehörigen Sonden detektiert wird. Für jedes untersuchte Transkript existiert auf dem Array ein Sonden-Set von je 11 passenden („*perfect match*“ oder kurz „PM“) und nicht passenden („*mismatch*“ oder kurz „MM“) Oligonukleotiden. Die elf PM- und MM-Sonden bilden also 11 Oligonukleotidpaare. Aus den gewichteten Signalintensitätswerten dieser 11 Oligonukleotidpaare errechnet sich in einem Mehrschritt-Algorithmus das singuläre Signal für das entsprechende Transkript: Zunächst wird durch Informationsvergleich zwischen PM- und MM-Intensität berechnet, wie präzise jedes einzelne der 11 Oligonukleotidpaare zwischen spezifischer und unspezifischer Hybridisierung unterscheiden kann. Dabei gilt: Je größer der Quotient aus PM- zu MM-Intensität, desto spezifischer ist die Hybridisierung des einzelnen Oligonukleotidpaares. Den errechneten Wert, der das Verhältnis von PM- zu MM-Intensität widerspiegelt, bezeichnet man als so genannten Diskriminationswert. Dieser stellt ein Signifikanzmaß für das entsprechende Oligonukleotidpaar dar. Anhand der Diskriminationswerte wird jedes der 11 Oligonukleotidpaare für die Berechnung des singulären Signals für das gesamte Transkript unterschiedlich gewichtet. Mit Hilfe von definierten Schwellenwerten wird anschließend noch festgelegt, ob das zugehörige Transkript in dem untersuchten Gewebe als exprimiert (bzw. P für englisch: *present*) oder nicht exprimiert (bzw. A für englisch: *absent*) eingestuft wird. Signalintensitätswerte, die genau zwischen den Schwellenwerten für P und A liegen, werden als *marginal* bezeichnet (M). Die Festlegung für P, A oder M eines Transkriptes nennt man auch *detection-call*. Da der Prozentsatz der als *present* detektierten Transkripte in Gewebe vom gleichen Typ (hier Lymphomgewebe) etwa konstant ist, beurteilte ich diesen Parameter als zusätzliche Plausibilitätskontrolle.

### 2.6.2 GeneSpring-Software

Zur statistischen Auswertung und der vergleichenden Analyse der insgesamt 40 *Microarrays* verwendete ich die Software *GeneSpring* Version 7.0 (Agilent).

### 2.6.3 Normalisierung

In biologischen Experimenten können nicht-biologische Faktoren zur unerwünschten Variabilität der Daten führen. Selbst identisches Gewebe (so genannte Replikate), das mit



unterschiedlichen Reagenzien bearbeitet und zu unterschiedlichen Zeitpunkten aufgearbeitet wurde, kann durch diese nicht-biologischen Effekte Variabilität aufweisen. Darüber hinaus können Menge und Qualität von RNA ebenso wie Färbereagenzien oder verschiedene andere Faktoren zu weiteren ungewollt relevanten Aspekten (artifizielle Variabilität) in der Vergleichbarkeit von verschiedenen *Microarrays* werden.

Um die artifizielle Variabilität zu minimieren und damit die erhobenen Daten untereinander vergleichbar zu machen, habe ich vor der vergleichenden Auswertung eine mehrstufige Normalisierung durchgeführt: Eine so genannte *Per-Chip*-Normalisierung dient dazu, die artifizielle Variabilität innerhalb jedes einzelnen *Microarrays* zu minimieren. Die zusätzlich durchgeführte *Per-Gen*-Normalisierung minimiert die artifizielle Variabilität zwischen verschiedenen *Microarrays* und hilft dadurch unterschiedliche *Microarrays* vergleichbar zu machen. Eine *Per-Chip*-Normalisierung wird automatisch durch die *Microarray-Suite*-Software der Firma Affymetrix durchgeführt und wird hier als *global scaling* bezeichnet. Die zu Grunde liegende Überlegung ist die Annahme, dass die mittlere Signalintensität eines *Microarrays* in verschiedenen Experimenten ungefähr gleich ist, wenn gleiche RNA-Mengen für die Hybridisierung eingesetzt werden: Beim *global scaling* wird für das Signal aller Sonden-Sets ein bereinigter Mittelwert (*trimmed mean signal*) kalkuliert. Hierbei werden starke Ausreißer in beide Richtungen aus den Daten herausgerechnet. Anschließend wird der bereinigte Mittelwert auf einen vom Anwender vorher festgelegten Gesamtintensitätswert (*target signal*) festgelegt. Der hierfür festgelegte Skalierungsfaktor ergibt sich aus folgender Formel:  $target\ signal = scaling\ factor \times trimmed\ mean\ signal$ . Um eine gute Vergleichbarkeit der *Microarrays* untereinander zu gewährleisten sollte der *scaling factor* allerdings bei den zu vergleichenden Proben um weniger als den Faktor 3 differieren. Um verschiedene *Microarrays* miteinander vergleichen zu können, muss ein weiterer Normalisierungsschritt durchgeführt werden. Diese zweite Normalisierung (*Per-Gene*-Normalisierung) habe ich mit Hilfe der *GeneSpring*-Software durchgeführt. Die *Per-Gene*-Normalisierung dient dem Zweck unterschiedliche Detektionseffizienzen in verschiedenen *Microarray*-Messungen zu kontrollieren. Zu diesem Zweck berechnete die Software den Quotienten aus der Signalintensität für Gen A in Probe x und der medianen Signalintensität für Gen A aus allen *Microarray*-Messungen.

#### 2.6.4 *Filtern von Genen*

Die *Microarray*-Messungen in der vorliegenden Arbeit generierten Expressionswerte von jeweils 22.283 Genen aus insgesamt 40 *Microarrays*. Bei gleichzeitiger Betrachtung aller Datenpunkte

in solch großen Datensätzen besteht die Gefahr, dass relevante Unterschiede zwischen den Datenpunkten im natürlichen Rauschen (Variabilität) der Daten verloren gehen. Aus diesem Grund ist erforderlich, die Menge der Datenpunkte sinnvoll zu reduzieren. Basierend auf der Überlegung, dass Gene, die in allen Proben gleich exprimiert sind, keinen Unterschied zwischen den Proben definieren können, habe ich diese Gene aus dem zu untersuchenden Datensatz in zwei Schritten herausgefiltert: Zunächst schloss ich Gene von der weiteren Datenanalyse aus, die in allen Proben als „nicht exprimiert“ (*absent*) eingestuft worden waren (*GeneSpring*-Funktion: *filtering genes on flags*). Im nächsten Schritt habe ich Gene auf Basis des berechneten normalisierten Expressionsstärke selektiert: Gene, deren Wert in allen Experimenten im mittleren Expressionsbereich zwischen 0,7 und 1,3 lagen, schloss ich ebenfalls von weiteren Analysen aus (*GeneSpring*-Funktion: *filtering genes on expression level*).

### 2.6.5 One way analysis of variance (ANOVA)

Nach dem Filtern testete ich die übrig gebliebenen Gene auf ihre differentielle Gen-Expression zwischen der Gruppe der BL/BLL und der DLBCL. Dazu habe ich die 40 Lymphomproben zunächst in ein Trainings-Set (20 Proben) und ein Test-Set (20 Proben) eingeteilt: Die Kriterien für die Set-Einteilung sind im Ergebnisteil in Kapitel 1.3.1 (Ergebnisteil) dargestellt. Die Suche nach differentiell exprimierten Genen habe ich zwischen den BL/BLL-Proben und den DLBCL-Proben des Trainings-Sets vorgenommen.

Für die Analyse verwendete ich die „one way analysis of variance“ (ANOVA)-Funktion der *GeneSpring*-Software. Als spezifischen statistischen Test habe ich den *Student's t-Test* für unverbundene Stichproben eingesetzt. Getestet wird die Nullhypothese, die besagt, dass die Mittelwerte der Genexpressionsintensitäten zwischen den zwei getesteten Gruppen gleich sind. Mit Hilfe des *Student's t-Test* lässt sich überprüfen, ob sich die Differenz der beiden Mittelwerte signifikant von 0 unterscheidet. Grundvoraussetzung ist dabei die Normalverteilung der beiden untersuchten Grundgesamtheiten (parametrischer Test). Ich wählte eine Irrtumswahrscheinlichkeit  $\alpha$  von 1 Prozent (entspricht  $p = 0,01$ ). Für einen p-Wert von 0,01 ergäbe sich damit jedoch pro 10 000 getesteter Gene ein Anteil von 100 falsch positiv differentiell exprimierten Genen. Um die Rate dieser falsch positiv identifizierten Gene gering zu halten, führte ich zusätzlich eine *Multiple-Testing-Correction-Prozedur (Benjamini-Hochberg-False-Discovery-Rate)* durch<sup>56</sup>. Die *Multiple-Testing-Correction-Prozedur* -Prozedur adjustiert den p-Wert entsprechend der Menge der getesteten Gene, wobei p indirekt proportional zur Anzahl der getesteten Gene korrigiert wird. Auf diese Weise habe ich die

Irrtumswahrscheinlichkeit  $\alpha$  so gewählt, dass die Anzahl falsch positiv identifizierter Gene  $< 1$  Prozent betrug.

Da eine statistische t-Testung nicht ohne Angabe einer Standardabweichung möglich ist, jedoch keine der dazu notwendigen Replikate (so genannte doppelte Gen-Expressionsmessung derselben Lymphomprobe) vorlagen, verwendete ich das Cross-Gene-Error-Modell der GeneSpring Software. Auf diese Weise ist eine Abschätzung der Messpräzision durch Kombination der Variabilität von Genexpressionsdaten möglich und dementsprechend lässt sich auch ein Vertrauensintervall ohne Vorliegen von Replikaten berechnen. Das grundsätzliche Prinzip des Cross-Gene-Error-Modells beruht darauf, dass sich die meisten Gen-Expressionswerte in verschiedenen Experimenten nicht verändern. Die Abweichung eines Gens vom normalisierten Wert 1 in verschiedenen Experimenten wird dabei als Maß für die Variabilität innerhalb der Daten angenommen und daraus eine Standardabweichung berechnet.

Um von den mittels des *Student's t-Test* identifizierten differentiell exprimierten Genen die hundert signifikantesten zu ermitteln, wählte ich den Fisher's Exact Test. Dieser Test gehört in die Gruppe der statistischen *exact tests* und ermittelt die Signifikanz der Assoziation zwischen zwei Merkmalen in einem 2x2 Kontingenztafelsystem.

#### 2.6.6 Kreuzvalidierung und Klassenprädiktion

Die Stabilität in der Unterscheidung zwischen BL/BLL-Proben und DLBCL-Proben anhand des Expressionsstatus der 100 identifizierten Gene, ermittelte ich mittels einer Leave-One-Out-Kreuzvalidierung. Bei diesem Verfahren wird jeweils eine Probe aus dem Training-Set separiert und als Test-Probe verwendet. Anhand der Ähnlichkeit der Expressionsstärke der 100 Prädiktorgene zwischen der separierten und den restlichen Proben wird die Gruppenzugehörigkeit der Probe vorhergesagt. Diese wird danach erneut dem Trainings-Set zugeordnet und die restlichen Proben des Kollektivs nach dem gleichen Prinzip vorhergesagt.

Zur Vorhersage der Gruppenzugehörigkeit der Proben des Test-Sets verwendete ich das *K-Nearest-Neighbourhood*-Verfahren. Dieser Algorithmus klassifiziert jede Probe des Test-Sets, indem er die k ähnlichsten Proben des Training-Sets ermittelt. Hierbei wird die Ähnlichkeit durch Kalkulation der Euklidischen Distanz der normalisierten Genexpressionsintensitäten

ermittelt. Die Anzahl der  $k$  ähnlichsten Proben ist benutzerdefiniert und ich habe sie entsprechend den Empfehlung der Softwarefirma auf  $0,66 \times$  Anzahl der Proben der kleinsten Klasse des Training-Sets eingestellt ( $k=7$ ). Da die Gruppenzugehörigkeit der ermittelten  $k$  ähnlichsten Proben bekannt ist (Trainings-Set), kann die Probe gemäß der Anzahl ihrer Voten vorhergesagt werden. Dabei wird für die Zugehörigkeit zu jeder Klasse ein  $p$ -Wert berechnet und die Gewebeprobe der Gruppe mit dem kleineren  $p$ -Wert zugeordnet.

Wenn das Verhältnis der beiden  $p$ -Werte zueinander (so genannte  $p$ -Wert-Verhältnis) jedoch den Grenzwert von  $0,2$  überschreitet, wird die getestete Probe als nicht vorhersagbar eingestuft.

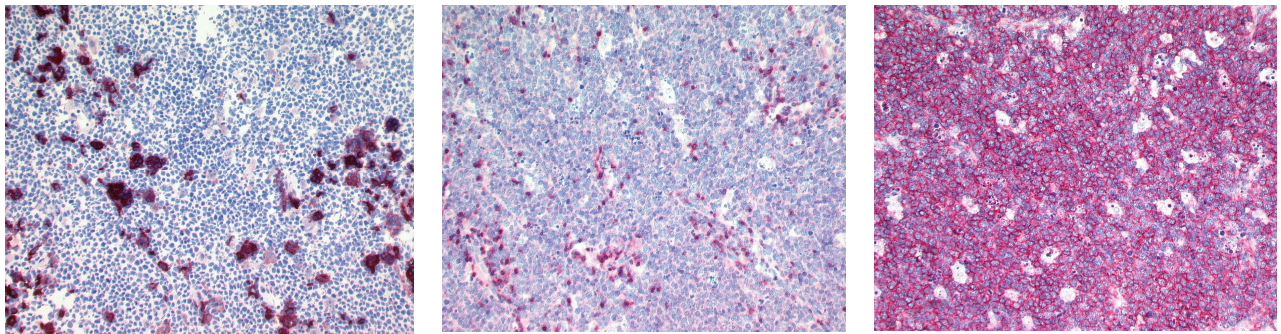
## 3 Ergebnisse

### 3.1 Charakterisierung der verwendeten Patientenproben

#### 3.1.1 Tumorzellreichtum im Lymphomgewebe

Für die Genexpressionsanalysen wurden aus einem Pool von 294 kryokonservierten BL-, BLL- und DLBCL-Gewebeproben 40 Proben ausgewählt (vgl. Kapitel 2.2.1), die - neben einer guten RNA-Qualität - einen Tumorzellgehalt von mindestens 70% aufwiesen. Auf diese Weise sollte eine mögliche Variabilität in den gemessenen Genexpressionsprofilen minimiert werden, die allein durch den unterschiedlichen Gehalt von nicht zum Tumor gehörenden *Bystander*-Zellen verursacht werden könnte. Zur Evaluation des Tumorzellgehalts wurden von allen Lymphomproben Gefrierschnitte angefertigt und mit Antikörpern gegen die Markermoleküle CD68 (Darstellung von Makrophagen), CD3 (Darstellung von T-Zellen) und CD20 (Darstellung von B-Zellen) gefärbt. Abbildung-1 zeigt exemplarisch die immunhistochemische Darstellung des Tumorzellanteils einer der verwendeten Gewebeproben.

**Abbildung-1:** Immunhistochemische Darstellung des *Bystander*-Zell-Anteils im Lymphomgewebe



Die immunhistochemische Darstellung von Makrophagen in der BerMac3-Färbung (A) und von T-Zellen in der CD3-Färbung (B) zeigt einen geringen Anteil von *Bystander*-Zellen im Lymphomgewebe. Die CD20-positiven Tumorzellen haben einen Gewebeanteil von mindestens 70%, wie die Darstellung mit Hilfe des L26-Antikörpers zeigt (C).

### 3.1.2 *Diagnose anhand morphologischer und immunologischer Kriterien der WHO-Klassifikation der malignen Lymphome*

Um für die 40 ausgewählten Gewebeproben eine histologische Diagnose nach der aktuellen WHO-Klassifikation von 2001 zu erstellen, wurde die Gewebemorphologie in Giemsa und H&E-gefärbten Paraffinschnitten von den sechs Direktoren der deutschen Referenzzentren für Lymphknotenpathologie beurteilt (Prof. Dr. H. Stein, Berlin; Prof. Dr. M.-L. Hansmann, Frankfurt; Prof. Dr. R. Parwaresch, Kiel; Prof. Dr. A. C. Feller, Lübeck; Prof. Dr. P. Möller, Ulm; Prof. Dr. H. K. Müller-Hermelink, Würzburg). Die Pathologen mikroskopierten die Gewebeproben zunächst einzeln und stellten eine Diagnose. Anschließend wurden die Gewebeproben gemeinsam mikroskopiert und eine Konsensusdiagnose festgelegt, der alle Pathologen des Panels zustimmten. Der anschließend mittels immunohistochemischer Färbungen ermittelte Immunphänotyp der Proben (CD20, CD10, BCL6, BCL2, CD5 und Ki-67; siehe Material und Methoden Kapitel 2.1) bestätigte die initiale diagnostische Einordnung in allen Fällen gemäß der aktuellen WHO-Klassifikation. Die lediglich schwache Expression von BCL2 bei 3 von 16 BL/BLL führte nicht zur Korrektur der ursprünglich gestellten histologischen Diagnose.

Als klassisches Burkitt-Lymphom (kurz: BL) wurden Gewebeproben diagnostiziert, die gemäß der WHO-Klassifikation der malignen Lymphome aus dem Jahre 2001 <sup>3</sup> folgende Merkmale besaßen:

- diffuses monotones Infiltrationsmuster mittelgroßer neoplastischer Zellen mit runden Kernen und verklumptem Chromatin, relativ klarem Parachromatin sowie multiplen basophilen mittelgroßen zentralständigen Nukleolen.
- zahlreiche Mitosen und hoher Ki-67 Index ( $\geq 90\%$ ).
- zahlreiche Apoptosen
- Sternenhimmelbild (vgl. Kapitel 1.1)
- Expression von CD20, BCL6, CD10 in Abwesenheit von CD5 und BCL2 (bei n=2 Fällen wurde eine schwache BCL2-Expression - entgegen der WHO-Klassifikation - nicht als Ausschlusskriterium betrachtet, da alle anderen diagnostischen Kriterien eindeutig erfüllt waren)

Die Diagnose atypisches BL (auch Burkitt-like-Lymphom genannt; kurz: BLL) ebenfalls gemäß der WHO-Klassifikation der malignen Lymphome aus dem Jahre 2001 <sup>3</sup> nach folgenden Kriterien vergeben:

- Mittegroße BL-Zellen mit hohem mitotischen Index und zahlreichen Apoptosen, jedoch größerer Pleomorphismus der Tumorzellen bezüglich Zellkerngröße und -form sowie einer geringeren Anzahl an Nukleolen.
- Expression von CD20, BCL6 und CD10 sowie ein Ki-67 Index von nahezu 100% bei gleichzeitiger Abwesenheit der Expression von BCL2 und CD5 (eine schwache BCL2- (n=1) bzw. CD5-Expression (n=3) wurde - entgegen der WHO-Klassifikation - nicht als Ausschlusskriterium betrachtet, da alle anderen diagnostischen Kriterien eindeutig erfüllt waren).

Als diffuses großzelliges B-Zell-Lymphom wurden gemäß der WHO-Klassifikation der malignen Lymphome aus dem Jahre 2001<sup>24</sup> Gewebeproben mit folgenden Charakteristiken diagnostiziert.

- Zellkerngröße mehr als zweimal so groß wie in normalen Lymphozyten
- Expression von CD20, sowie variable Expression von BCL6, CD10, BCL2, CD5
- Hohe Proliferationsrate ermittelt mittels Ki-67-Färbung (gewöhnlich > 40%, in einigen Fällen möglicherweise > 90).

Alle histologischen Diagnosen wurden ohne Kenntnis der zytogenetischen Befunde erstellt.

Unter Anwendung der oben genannten WHO-Kriterien wurden insgesamt 10 der 40 Lymphomproben (25%) als klassisches BL, 6 Gewebeproben als atypisches BL (15%), 23 Proben (60%) als DLBCL und eine Probe (2%) als nicht weiter klassifizierbares aggressives B-NHL diagnostiziert. Die histologische Diagnose sowie die Ergebnisse der immunhistochemischen Färbungen sind in Tabelle-3 zusammengefasst.

### *3.1.3 Genetische Analysen der Lymphomproben*

Die Ergebnisse der mittels FISH-Technik durchgeführten genetischen Analysen der 40 Gewebeproben (Analysen an paraffinfixiertem Gewebe) wurden freundlicherweise von Herrn Prof. Dr. R. Siebert aus dem Institut für Humangenetik der Universität Kiel zur Verfügung gestellt und waren nicht Gegenstand dieser Arbeit. In Kiel wurden Sonden für den IGH, IGK, IGL und den MYC-Lokus verwendet.

Von den 40 Lymphomen wiesen 19 Lymphome eine Fusion des MYC-Gens mit einem der Immunglobulinloki auf (im Folgenden IG-MYC genannt, vgl. auch Kapitel 1.1), eine der Gewebeproben zeigte eine Fusion des MYC-Gens mit einem nicht-Immunglobulinlokus (Non-

IG-MYC) und 20 Lymphome wiesen keine Translokation des MYC-Gens auf. Der MYC-Translokationsstatus der Gewebeprobe ist in Tabelle-3 den jeweiligen morphologischen und immunphänotypischen Charakteristika der Gewebeprobe gegenübergestellt.

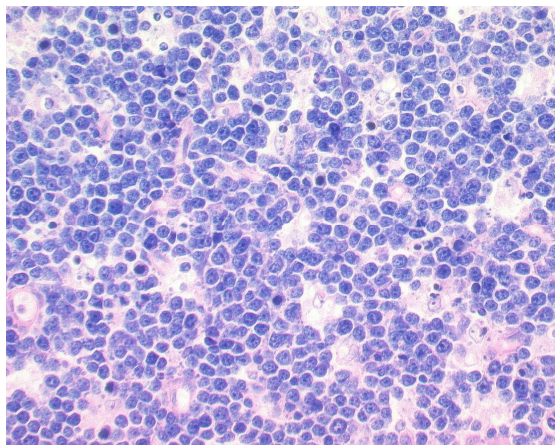
#### *3.1.4 Diagnostische Einordnung unter Verwendung aller WHO-Diagnosekriterien: Histologie, Immunologie und Genetik*

Insgesamt ergab sich eine Zahl von 10 Lymphomen, die histologisch/immunologisch einem BL und 6 Lymphome, die histologisch einem BLL entsprachen und gleichzeitig eine IG-MYC-Translokation aufwiesen (diese Gewebeprobe werden im Folgenden als MYC-positive BL (n=10) bzw. MYC-positive BLL (n=6) bezeichnet). Insgesamt 23 Lymphome erfüllten alle oben dargestellten WHO-Kriterien für ein DLBCL. Dabei wiesen 20 dieser 23 Gewebeprobe einen negativen MYC-Translokationsstatus auf (im Folgenden MYC-negative DLBCL genannt). Drei der 23 DLBCL zeigten einen positiven Nachweis einer IG-MYC-Translokation (im Folgenden MYC-positive DLBCL genannt).

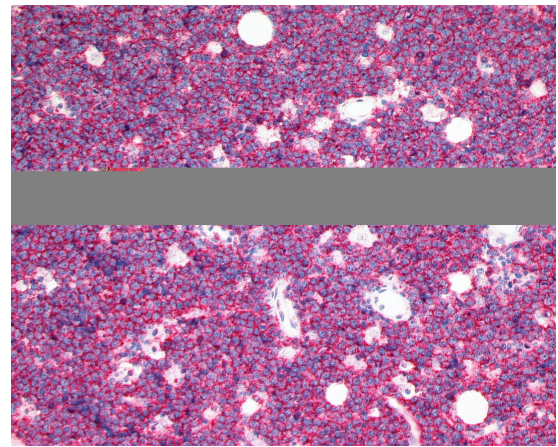
Eine der 40 untersuchten Proben zeigte das histologische Bild eines nicht genauer klassifizierbaren aggressiven B-NHL und war in der zytogenetischen Analyse positiv für eine IG-MYC-Translokation (im Folgenden MYC-positives nicht klassifizierbares aggressives B-NHL genannt).



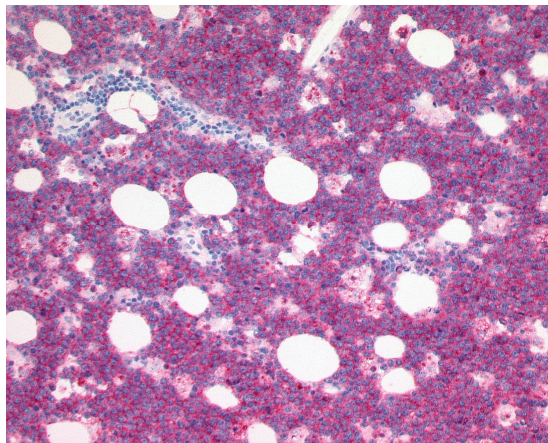
**Abbildung-2:** Histologische Diagnose eines Burkitt-Lymphoms (BL)



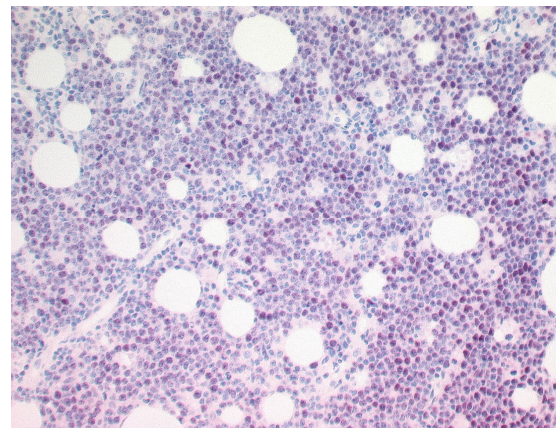
**A**



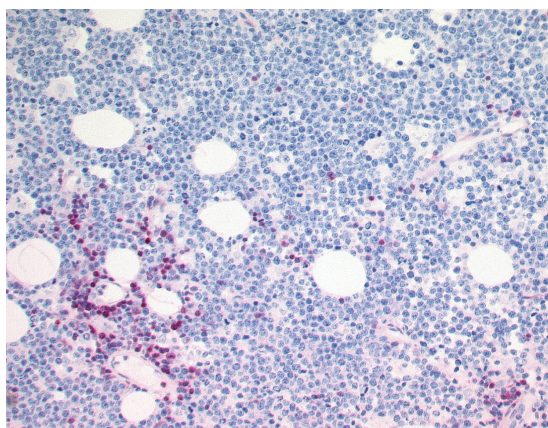
**B**



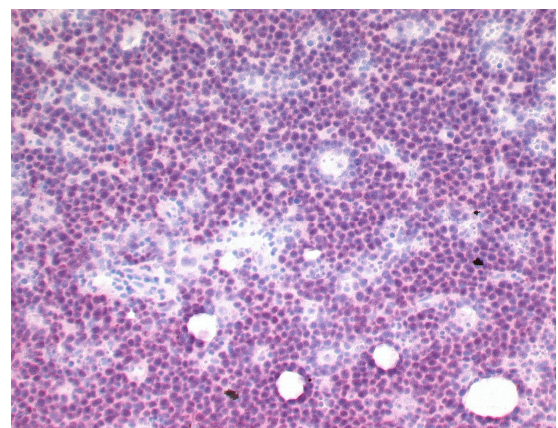
**C**



**D**



**E**



**F**

Das Burkitt-Lymphom zeigt in der Giemsa-Färbung (A, Vergrößerung) ein typisches Sternenhimmelbild und weist immunohistochemisch eine deutliche Oberflächenexpression für CD20 (B) und CD10 (C), jedoch Negativität für BCL2 (E) auf (die wenigen positiven Zellen stellen BCL2-positive T-Zellen als interne Positivkontrolle dar). Die Kerne sind positiv für BCL6 (D) und es zeigt sich eine hohe Wachstumsfraktion (95%) in der Ki-67-Färbung (F).

**Tabelle-3:** Histologische Diagnose und Ergebnisse immunhistochemischer Färbungen

Proben-ID/ morphol.- Diagnose	CD20	CD10	BCL6	BCL2	Ki-67*	MYC- Status
MPI-017 BL	pos	pos	pos	neg	99	pos
MPI-032 BL	pos	pos	pos	pos	99	pos
MPI-033 BL	pos	pos	pos	neg	97	pos
MPI-077 BL	pos	pos	pos	neg	95	pos
MPI-079 BL	pos	pos	pos	neg	90	pos
MPI-080 BL	pos	pos	n.v.	pos	90	pos
MPI-084 BL	pos	pos	pos	neg	95	pos
MPI-085 BL	pos	pos	pos	neg	90	pos
MPI-089 BL	pos	pos	pos	neg	99	pos
MPI-114 BL	pos	pos	pos	neg	90	pos
MPI-001 BLL	pos	pos	pos	neg	97	pos
MPI-002 BLL	pos	pos	pos	neg	99	pos
MPI-027 BLL	pos	pos	pos	pos	98	pos
MPI-043 BLL	pos	pos	pos	neg	99	pos
MPI-044 BLL	pos	pos	pos	neg	97	pos
MPI-082 BLL	pos	pos	pos	neg	90	pos
MPI-006 DLBCL	pos	pos	n.v.	pos	90	neg
MPI-008 DLBCL	pos	neg	pos	pos	75	neg
MPI-009 DLBCL	pos	neg	neg	pos	75	neg
MPI-010 DLBCL	pos	neg	n.v.	neg	80	neg
MPI-013 DLBCL	pos	neg	n.v.	neg	80	neg
MPI-018 DLBCL	pos	pos	pos	neg	99	pos
MPI-034 DLBCL	pos	pos	pos	pos	98	pos
MPI-039 DLBCL	pos	neg	pos	pos	60	neg
MPI-040 DLBCL	pos	neg	neg	pos	55	neg
MPI-063 DLBCL	pos	neg	n.v.	neg	n.v.	neg
MPI-065 DLBCL	pos	pos	pos	neg	75	neg
MPI-073 DLBCL	pos	neg	pos	pos	70	neg
MPI-074 DLBCL	pos	neg	pos	pos	70	pos
MPI-087 ag.B-NHL	pos	pos	pos	neg	99	pos
MPI-113 DLBCL	pos	pos	pos	pos	90	neg
MPI-118 DLBCL	pos	neg	pos	pos	80	neg
MPI-119 DLBCL	pos	pos	pos	pos	70	neg
MPI-130 DLBCL	pos	neg	neg	pos	80	neg
MPI-132 DLBCL	pos	neg	neg	pos	75	neg
MPI-133 DLBCL	pos	neg	neg	pos	99	neg
MPI-134 DLBCL	pos	neg	pos	neg	85	neg
MPI-135 DLBCL	pos	neg	neg	pos	99	neg
MPI-136 DLBCL	pos	neg	neg	pos	95	neg
MPI-006 DLBCL	pos	pos	n.v.	pos	90	neg

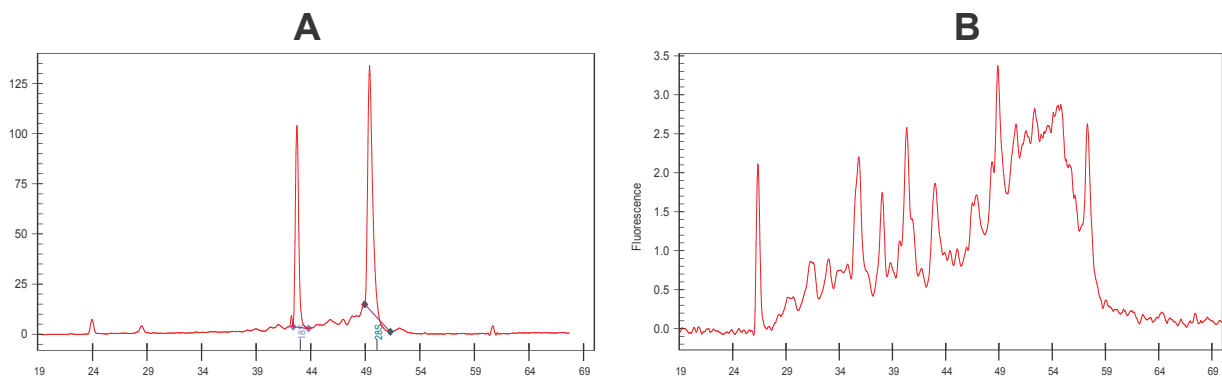
BL= Burkitt-Lymphom; BLL = Burkitt-like-Lymphom; DLBCL = diffuses großzelliges B-Zell-Lymphom; ag.B-NHL = aggressives B-Zell-Non-Hodgkin-Lymphom; pos = Tumorzellen positiv; neg = Tumorzellen negativ; n.v. = nicht vorhanden (nicht ausreichend Gewebematerial); \* Prozentzahlen positiver Tumorzellen

## 3.2 Qualitätskontrolle der Probenaufarbeitung und der *Microarrays*

### 3.2.1 Analyse der Gesamt-RNA- und cRNA-Qualität der Lymphomproben

Aus den verwendeten 40 Gewebeproben konnten jeweils mindestens 5 µg Gesamt-RNA isoliert werden (vgl. Kapitel 2.2.1). Somit stand für die verwendeten 40 Gewebeproben die erforderliche Menge Gesamt-RNA zur Einsetzung in die IVT zur Verfügung. Die kapillarelektrophoretische Auftrennung eines Aliquots der Gesamt-RNA im *Agilent Bioanalyzer* belegte die hochwertige RNA-Qualität in den verwendeten 40 Proben: Alle 40 in die Studie eingeschlossenen Gewebeproben wiesen die typische Konfiguration der 18S und 23S *RNA-Peaks* sowie zufrieden stellende *Peak-Flächenratios* um den Wert 2 auf, deren Anwesenheit einen geringen Degradationsgrad der mRNA anzeigen. In *Abbildung-3* ist das *Bioanalyzer*-Profil einer hochwertigen RNA-Probe exemplarisch einer RNA-Probe mit starker Degradation gegenübergestellt, die vormals im Rahmen des 2. Schrittes des Auswahlprozesses von der Studie ausgeschlossen worden war (vgl. Kapitel 2.2.1). Aus der isolierten Gesamt-RNA wurde in allen Fällen mindestens 15 µg markierte cRNA (photometrische Messung) als Produkt der IVT gewonnen, womit die von Affymetrix geforderte cRNA-Menge für die *Microarray*-Hybridisierung bei allen Proben zur Verfügung stand. Vor der Hybridisierung wurde eine erneute kapillarelektrophoretische Auftrennung der cRNA im *Agilent Bioanalyzer* sowohl vor als auch nach der Fragmentierung durchgeführt. Bei allen Messungen belegte das Bandenmuster die erfolgreiche IVT und Fragmentierung (vgl. Kapitel 2.4.8).

**Abbildung-3:** Gegenüberstellung von qualitativ guter versus qualitativ schlechter Total-RNA im *Agilent Bioanalyzer* (Kapillarelektrophorese)



Dargestellt ist die Fluoreszenzintensität in Abhängigkeit von der Elektrophoresezeit. Profil A zeigt gute RNA-Qualität, mit deutlich erkennbaren 18S- und 23S-RNA-Peaks. Dagegen zeigt sich in Profil B eine deutliche RNA-Degradation - erkennbar an vielen unterschiedlich großen Fragmenten (Gewebeprobe wurde nicht verwendet).

### 3.2.2 Qualitätskontrollen der Microarrays

Zwecks Qualitätskontrolle und zur Sicherung der Vergleichbarkeit der *GeneChip*-Experimente, wurde in allen 40 Experimenten eine Reihe von Kontrollen in den Versuchsablauf integriert:

Eine interne Qualitätskontrolle der verwendeten Proben-RNA und ihrer Aufarbeitung (IVT) erfolgte durch die Evaluation von Haushaltsgenen (*housekeeping genes*), die in allen humanen Zellarten weitgehend gleichmäßig exprimiert werden. Auf dem verwendeten Affymetrix GeneChip U133A werden zu diesem Zweck die Gene GAPDH (*glyceraldehyd-3-phosphate-dehydrogenase*), ACTB (*beta-actin*) und ISGF-3 (*interferon-stimulated-transcription-factor-3*) verwendet. Die SONDENSÄTZE für diese Gene sind auf dem *Microarray* für die jeweiligen Bereiche des 3'-Endes, des 5'-Endes und der Mitte des Transkriptes gestaltet, wobei das Verhältnis der 3'-Signalintensitäten zu den 5'-Signalintensitäten Auskunft über den Degradationszustand der Proben-RNA gibt. Bei allen verwendeten Proben ergaben sich vom Hersteller empfohlene 3'/5'-Ratios < 3, die eine hochwertige Qualität der ursprünglich in die IVT eingesetzten Gesamt-RNA anzeigen. Lediglich die in früheren Arbeiten beschriebene konstant >3 liegende 3'/5'-Ratio für das Haushaltsgen ISGF-3 fand sich auch in dieser Arbeit in allen Proben wieder. Eine Ursache konnte auch an dieser Stelle nicht ausgemacht werden, wobei der Wert nicht mit dem Degradationsgrad der RNA (Kapillarelektrophorese) korrelierte. Aus diesem Grund wurde diesem Befund keine weitere Beachtung geschenkt.

Die Qualitätssicherung der Arbeitsschritte Hybridisierung, Waschen und Färben erfolgte durch Zugabe definierter Mengen von Kontrollgen-RNA (prokaryotische Hybridisierungskontrollmix = *spike controls*) zur Hybridisierungslösung (vgl. Kapitel 2.5.3): Alle vier synthetisch hergestellten Transkripte BioB, BioC, BioDn (bakterielle Gene von *Escherichia Coli*), sowie das Transkript CreX (Rekombinase des Bakteriophagen P1) waren in allen 40 Experimenten detektierbar.

Zusätzlich zu den prokaryotischen Hybridisierungskontrollen werden außerdem so genannte B2-Oligonukleotide in den Hybridisierungscocktail gegeben (vgl. Kapitel 2.5.3). Auch diese stellen Positivkontrollen dar, deren entsprechende SONDENSÄTZE am äußersten Rand des Arrays angeordnet sind. Zwischen zwei *Perfect-Match*-SONDENSÄTZE (vgl. Kapitel 2.6.1) ist jeweils ein *Mismatch*-SONDENSÄTZE angeordnet, so dass sich am Rande des Arrays ein schachbrettartiges Fluoreszenzmuster ergibt. Dieses Fluoreszenzmuster fungiert als visuelle Kontrolle des *Microarrays*, und konnte in allen untersuchten Fällen nachgewiesen werden. Unter Berücksichtigung der angeführten Kontrollen genügten alle *GeneChip*-Experimente den Qualitätskriterien und konnten für die weitere Auswertung verwendet werden.

Die für die Vergleichbarkeit der *GeneChips* vom Hersteller geforderte geringe Differenz des für die verschiedenen Arrays errechneten Scaling-Faktors (vgl. Kapitel 2.6.3) war bei allen Proben erfüllt (Unterschiede  $\leq$  Faktor 3). Die Anzahl der als positiv detektierten Transkripte (*present*, vgl. Kapitel 2.6.1) belief sich in allen verwendeten Proben auf etwa 40%.

### **3.3 Entwicklung eines 100 Markergene umfassenden Prädiktors und Vorhersage einer molekularen Diagnose**

In dieser Arbeit wurde untersucht, ob sich eine Gruppe von Genen identifizieren lässt (so genannte differentiell exprimierte Gene), anhand deren Expressionsstatus sich aggressive B-NHL differentialdiagnostisch in BL/BLL bzw. DLBCL unterteilen und diagnostisch vorhersagen lassen. Dazu wurde folgendermaßen vorgegangen: Zunächst wurden die 40 Gewebeproben von aggressiven B-NHL in zwei gleich große Gruppen - ein *Training-Set* und ein *Test-Set* - eingeteilt. Für das *Training-Set* wurden Lymphomproben ausgewählt, bei denen eine differentialdiagnostische Einordnung in entweder BL/BLL oder DLBCL mit herkömmlichen diagnostischen Mitteln möglich war (Konsensusdiagnose aller sechs Lymphknotenreferenzpathologen und Erfüllung aller Kriterien der WHO-Klassifikation, siehe Kapitel 3.1.2 und 3.1.4). Innerhalb des *Training-Sets* wurden Gene identifiziert, die zwischen den beiden diagnostisch vorgegebenen Lymphomgruppen differentiell exprimiert waren. Anhand des Expressionsstatus der identifizierten Gene wurde anschließend eine „molekulare Diagnose“ für die Lymphomproben des *Test-Sets* vorhergesagt.

#### *3.3.1 Einteilung der Lymphomproben in ein Training- und Test-Set*

Für das *Training-Set* wurden ausschließlich Gewebeproben verwendet, bei denen die sechs Lymphknotenreferenzpathologen eine einheitliche Konsensusdiagnose festgelegt hatten und die alle diagnostischen Kriterien der WHO-Klassifikation erfüllten (vgl. Kapitel 3.1.2 und 3.1.4): Von den insgesamt 16 MYC-positiven BL/ MYC-positiven BLL wurden für das *Training-Set* 10 Proben zufällig ausgewählt. Ebenso wurden von den 20 MYC-negativen DLBCL 10 Gewebeproben zufällig für das *Training-Set* ausgewählt (MYC-positive DLBCL wurden nicht für das Training-Set

**Tabelle-4:** Morphologische, immunphänotypische und genetische Eigenschaften der Lymphomproben des *Training-Sets*

Proben-ID/ Morphologie	CD20	CD10	BCL6	BCL2	Ki-67*	MYC- Transl.
MPI-085 BL	pos	pos	pos	neg	90	pos
MPI-077 BL	pos	pos	pos	neg	95	pos
MPI-079 BL	pos	pos	pos	neg	90	pos
MPI-080 BL	pos	pos	n.v.	pos	90	pos
MPI-033 BL	pos	pos	pos	neg	97	pos
MPI-044 BLL	pos	pos	pos	neg	97	pos
MPI-082 BLL	pos	pos	pos	neg	90	pos
MPI-043 BLL	pos	pos	pos	neg	99	pos
MPI-001 BLL	pos	pos	pos	neg	97	pos
MPI-002 BLL	pos	pos	pos	neg	99	pos
MPI-113 DLBCL	pos	pos	pos	pos	90	neg
MPI-073 DLBCL	pos	neg	pos	pos	70	neg
MPI-118 DLBCL	pos	neg	pos	pos	80	neg
MPI-119 DLBCL	pos	pos	pos	pos	70	neg
MPI-063 DLBCL	pos	neg	n.v.	neg	n.v.	neg
MPI-040 DLBCL	pos	neg	neg	pos	55	neg
MPI-039 DLBCL	pos	neg	pos	pos	60	neg
MPI-008 DLBCL	pos	neg	pos	pos	75	neg
MPI-009 DLBCL	pos	neg	neg	pos	75	neg
MPI-136 DLBCL	pos	neg	neg	pos	95	neg

BL= Burkitt-Lymphom; BLL=Burkitt-like-Lymphom; DLBCL=diffuses großzelliges B-Zell-Lymphom; pos = Tumorzellen positiv; neg = Tumorzellen negativ; n.v. = nicht vorhanden (nicht ausreichendes Gewebematerial); \* Prozentzahlen positiver Tumorzellen; MYC-Transl. = MYC-Translokation

ausgewählt, um der Tatsache gerecht zu werden, dass die MYC-Translokation klassischerweise ein charakteristisches Merkmal des BL darstellt und nur in einer Minderheit von DLBCL gefunden wird<sup>26</sup>, vgl. Kapitel 1.3).

Die übrigen 20 Gewebeproben wurden in das *Test-Set* eingeschlossen: Dies beinhaltet dementsprechend sechs MYC-positive BL/BLL, zehn MYC-negative DLBCL, drei MYC-positive DLBCL sowie das MYC-positive nicht weiter klassifizierbare aggressive B-NHL.

### 3.3.2 Identifikation von 2115 differentiell exprimierten Genen zwischen BL/BLL und DLBCL innerhalb des *Training-Sets*

Um Gene zu identifizieren, die innerhalb des *Training-Sets* zwischen den BL/BLL und DLBCL aggressiven B-NHL differentiell exprimiert werden, wurde die GeneSpring Software verwendet: Der erste Analyseschritt umfasste den Ausschluss von Genen, die in keiner der 20

Gewebeproben exprimiert wurden (n=7180) bzw. in allen Proben einen normalisierten Expressionswert zwischen 0,7 und 1,3 aufwiesen (n=3545) (vgl. Kapitel 2.6.4). Im zweiten Analyseschritt wurden die verbliebenen 11558 Gene (ursprünglich 22 283) unter Anwendung des *Student's t-Test* auf differentielle Expression zwischen den beiden diagnostischen Gruppen getestet ( $p=0,01$ ). Durch gleichzeitige Anwendung einer *Multiple Testing* Korrektur (*Benjamini and Hochberg False Discovery Rate*) konnte der Anteil an falsch positiv differentiell exprimierten Genen auch in diesem sehr großen Datensatz unter 1 Prozent gehalten werden.

Unter Anwendung der oben genannten Kriterien konnten zwischen den beiden Lymphomgruppen des *Training-Sets* insgesamt 2115 signifikant differentiell exprimierte Gene gefunden werden. Um eine überschaubare Anzahl von Genen zu erhalten, wurden von diesen die 100 signifikantesten mittels des *Fisher's Exact-Test* ermittelt. Alle 100 identifizierten Gene zeigten einen p-Wert  $< 0,0076$  auf (vgl. Tabelle-5).

**Tabelle-5:** Top-100-Markergene, die zwischen BL- und DLBCL-Gruppe des *Training-Sets* differentiell exprimiert waren.

Gen-Name	Affymetrix-Annotation	P-Wert
tumor necrosis factor receptor superfamily, member 5	215346_at	0,00004
guanine nucleotide binding protein (G protein), gamma 10	201921_at	0,00004
major histocompatibility complex, class II, DQ alpha 1	212671_s_at	0,00005
major histocompatibility complex, class II, DR beta 3	215193_x_at	0,00005
leukocyte-associated Ig-like receptor 1	210644_s_at	0,00005
dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)	201537_s_at	0,00005
formyl peptide receptor 1	205119_s_at	0,00005
major histocompatibility complex, class I, B	208729_x_at	0,00005
Src-like-adaptor	203760_s_at	0,00005
chromosome 22 open reading frame 4	33778_at	0,00006
CD83 antigen (activated B lymphocytes, immunoglobulin superfamily)	204440_at	0,00006
myeloid cell nuclear differentiation antigen	204959_at	0,00007
zd24d07.s1 Soares_fetal_heart_NbHH19W	205298_s_at	0,00007
pim-1 oncogene	209193_at	0,00007
erythrocyte membrane protein band 4.1-like 3	211776_s_at	0,00007
leukocyte immunoglobulin-like receptor, subfamily B	211336_x_at	0,00007
formin binding protein 1	212288_at	0,00007
small nuclear ribonucleoprotein polypeptide E	203316_s_at	0,00007
TNFRSF1A-associated via death domain	1729_at	0,00007
S100 calcium binding protein A6 (calcyclin)	217728_at	0,00007
tumor necrosis factor, alpha-induced protein 3	202643_s_at	0,00007
caspase 1, apoptosis-related cysteine protease	211366_x_at	0,00008
docking protein 1, 62kDa (downstream of tyrosine kinase 1)	216835_s_at	0,00008
ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e	200096_s_at	0,00008
FN5 protein	219806_s_at	0,00008
major histocompatibility complex, class I, F	221875_x_at	0,00008
major histocompatibility complex, class I, F	221978_at	0,00008
single-stranded DNA binding protein 2	203787_at	0,00008
rabconnectin-3	212820_at	0,00009
SWI/SNF	208793_x_at	0,00009
host cell factor 2	219484_at	0,00009
tumor necrosis factor receptor superfamily, member 5	205153_s_at	0,00009
chromosome 1 open reading frame 38	210785_s_at	0,00009
transcription factor 3	209153_s_at	0,00009
Pleckstrin	203471_s_at	0,00009
Homo sapiens cDNA clone IMAGE:2776565	222281_s_at	0,00010
metallothionein-IF	210524_x_at	0,00010
chromosome 2 open reading frame 17	222129_at	0,00010
SLAM family member 8	219385_at	0,00011
tumor necrosis factor (ligand) superfamily, member 13	209500_x_at	0,00011
biliverdin reductase A	203773_x_at	0,00011
major histocompatibility complex, class I, A	213932_x_at	0,00011
echinoderm microtubule associated protein like 2	204398_s_at	0,00011
ceroid-lipofuscinosis, neuronal 2, late infantile	200743_s_at	0,00011
SLAM family member 8	219386_s_at	0,00011
neutrophil cytosolic factor 2	209949_at	0,00011
not named	201042_at	0,00011
hypothetical protein FLJ20477	51146_at	0,00011
selenophosphate synthetase 1	208939_at	0,00012
nac51f11.x1 NCI_CGAP_Brn23	213373_s_at	0,00012
heterogeneous nuclear ribonucleoprotein A3	211929_at	0,00012
caspase 1, apoptosis-related cysteine protease	209970_x_at	0,00012
Sjogren syndrome antigen A1	204804_at	0,00012
leukocyte immunoglobulin-like receptor, subfamily B	210225_x_at	0,00012
CD58 antigen, (lymphocyte function-associated antigen 3)	211744_s_at	0,00013
centaurin, alpha 1	90265_at	0,00013
KIAA0980 protein	207705_s_at	0,00013
Transcribed sequence with strong similarity to protein ref:NP_065136.1	219737_s_at	0,00013
gamma-glutamyltransferase 1	208284_x_at	0,00013
lectin, galactoside-binding, soluble, 3 (galectin 3)	208949_s_at	0,00013
WD repeat and SOCS box-containing 2	201760_s_at	0,00013
SWI/SNF	212257_s_at	0,00013
tumor rejection antigen (gp96) 1	200599_s_at	0,00013
karyopherin (importin) beta 3	211954_s_at	0,00013
complement component 1, q subcomponent binding protein	208910_s_at	0,00013
hypothetical protein FLJ23231	218810_at	0,00013

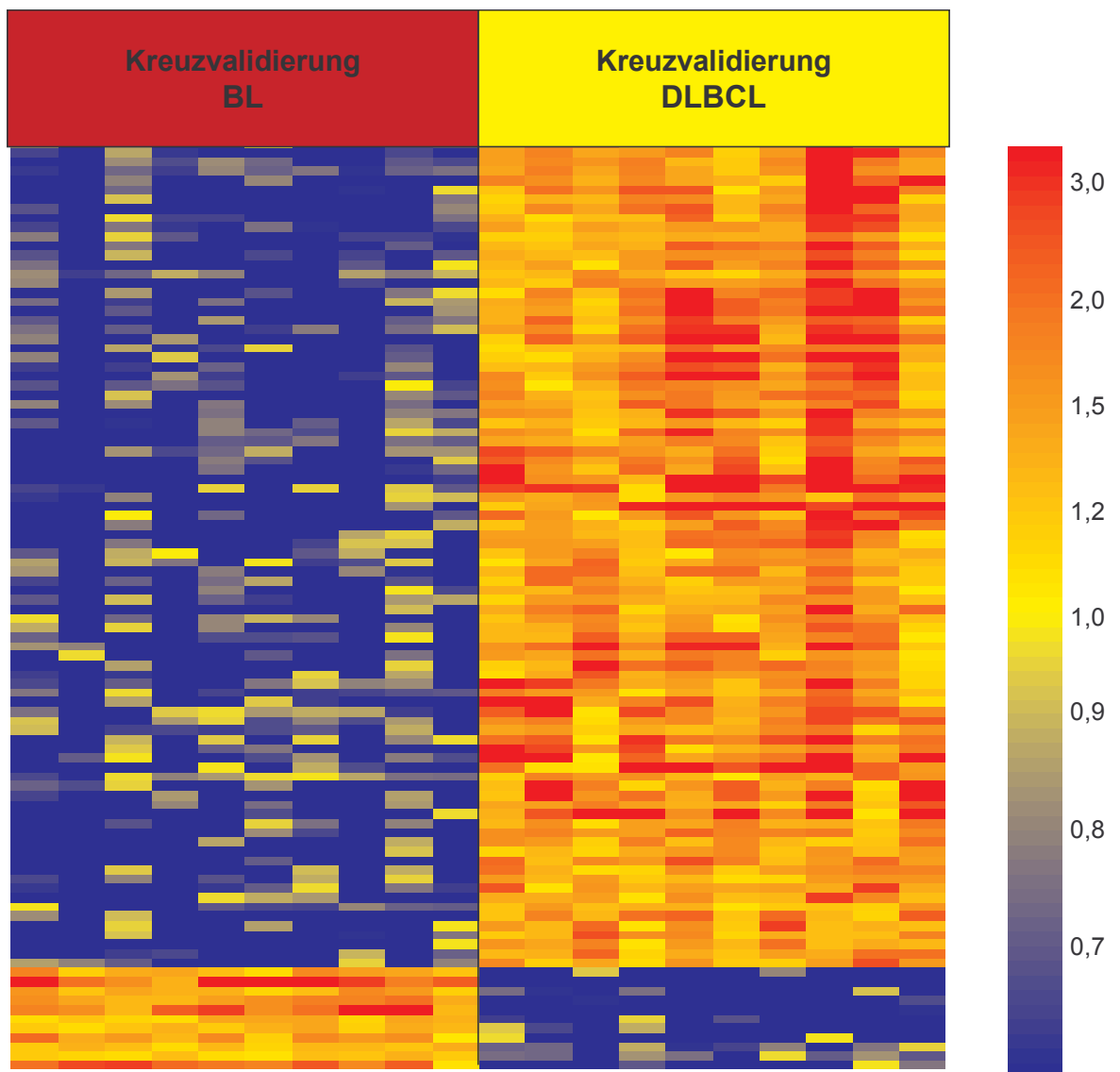


Gen-Name	Affymetrix-Annotation	P-Wert
SH3-domain binding protein 5 (BTK-associated)	201810_s_at	0,00014
KIAA1201 protein	212906_at	0,00014
homeo box C6	206858_s_at	0,00014
caspase 4, apoptosis-related cysteine protease	213596_at	0,00014
TNFAIP3 interacting protein 1	207196_s_at	0,00014
PHD finger protein 15	212660_at	0,00014
ADP-ribosylation factor GTPase activating protein 3	202211_at	0,00014
cytoskeleton related vitamin A responsive protein	200760_s_at	0,00015
gamma-glutamyltransferase 1	211417_x_at	0,00015
vacuolar protein sorting 33B (yeast)	44111_at	0,00015
tumor necrosis factor receptor superfamily, member 6	204781_s_at	0,00015
ectonucleotide pyrophosphatase/phosphodiesterase 4	204160_s_at	0,00015
myristoylated alanine-rich protein kinase C substrate	201669_s_at	0,00031
surfactant, pulmonary-associated protein B	209810_at	0,00031
KIAA0274	203656_at	0,00031
sorting nexin 11	53912_at	0,00048
microphthalmia-associated transcription factor	207233_s_at	0,00048
hepatocellularcarcinoma-associated antigen HCA557a	213861_s_at	0,00049
I-mfa domain-containing protein	217599_s_at	0,00050
TAP binding protein (tapasin)	208829_at	0,00053
iduronate 2-sulfatase (Hunter syndrome)	202439_s_at	0,00055
IL-13Ralpha	211612_s_at	0,00056
solute carrier family 23 (nucleobase transporters), member 2	209237_s_at	0,00056
galactosidase, beta 1	201576_s_at	0,00061
SP140 nuclear body protein	207777_s_at	0,00062
3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase	202772_at	0,00062
ras responsive element binding protein 1	203704_s_at	0,00062
PRO2275 mRNA, complete cds	211429_s_at	0,00063
paternally expressed 10	212094_at	0,00122
matrin 3	200626_s_at	0,00126
Epstein-Barr virus induced gene 2	205419_at	0,00132
methionine sulfoxide reductase B	218773_s_at	0,00141
hypothetical protein LOC283537	214719_at	0,00760

### 3.3.3 Diagnosevorhersage der Lymphomproben mittels 100 differentiell exprimierten Genen (Markergenprädiktor)

Mit den als differentiell exprimiert identifizierten 100 Genen hat man einen Satz von Markergenen an der Hand, mit denen sich BL/BLL von DLBCL unterscheiden lassen sollten. Um diese Hypothese zu überprüfen bzw. die Vorhersage-Präzision der 100 Gene in Bezug auf die Gruppenzugehörigkeit der Lymphomproben des *Training-Sets* zu testen, wurde eine *Leave-One-Out-Kreuzvalidierung* des *Training-Sets* durchgeführt (vgl. Kapitel 2.6.5): Jeweils eine der Proben des Datensets wird isoliert betrachtet und - basierend auf ihrer Markergenexpression - eine Klassenzuordnung zu einer der durch die restlichen Lymphomproben definierten diagnostischen Gruppe vorgenommen. Dieses Verfahren wird solange wiederholt bis für jede Gewebeprobe des *Sets* eine Klassenprädiktion vorliegt. Verglichen mit der vorgegebenen WHO-Diagnose ergab die Markergenprädiktion eine Übereinstimmung in allen 20 Fällen (100%). Die Beziehung zwischen Genexpression der hundert Markergene und Gruppenzugehörigkeit der 20 Lymphomproben des *Training-Sets* ist in Abbildung-4 dargestellt.

**Abbildung-4:** Kreuzvalidierung des *Training-Sets* an 20 Lymphomgewebeproben



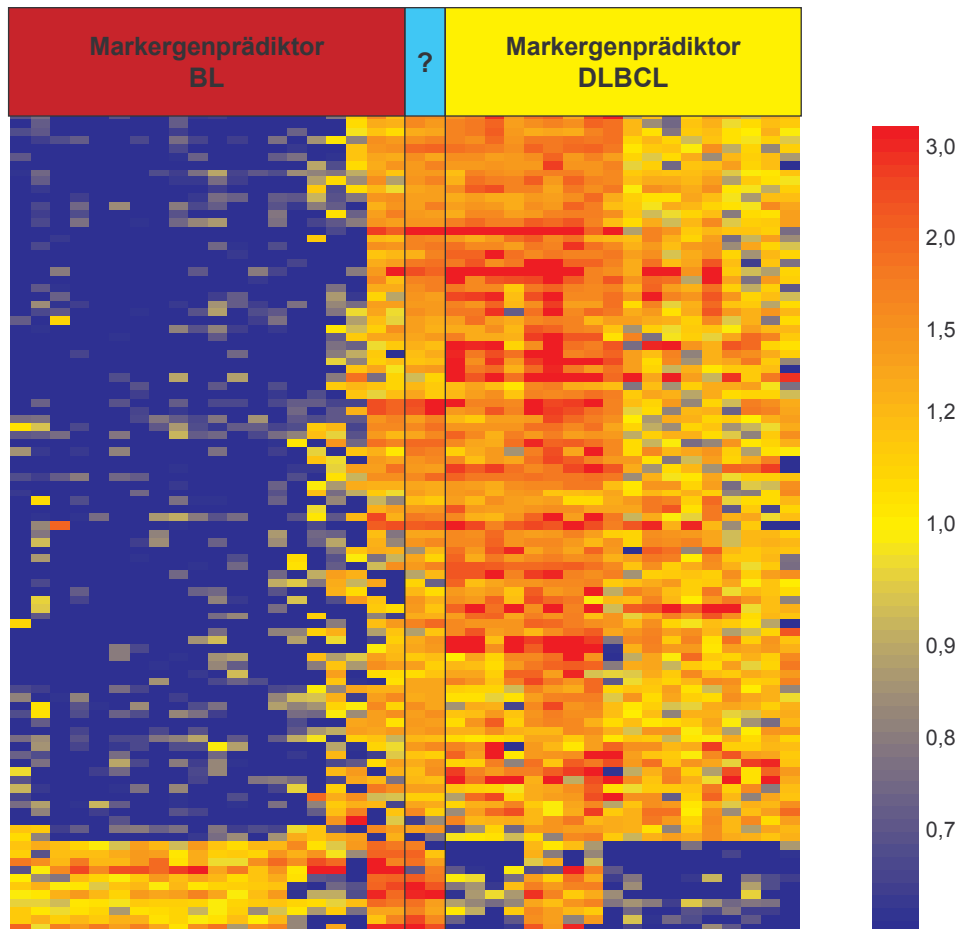
Die Abbildung zeigt die Beziehung zwischen Expressionsstatus von 100 Markergenen und Zugehörigkeit zu einer diagnostischen Gruppe. Rote Areale markieren verstärkte und blaue Areale verminderte Genexpression, wobei jede Säule eine Gewebeprobe und jede Reihe ein einzelnes Gen repräsentiert. Die Farbskala rechts im Bild zeigt die relative Genexpressionsstärke an.

Nachdem mit Hilfe der *Leave-One-Out*-Kreuzvalidierung die Vorhersage-Präzision der 100 Markergene gesichert war, wurde anschließend die Diagnose der 20 Gewebeproben des *Test-Sets* mit Hilfe des *K-Nearest-Neighbour*-Algorithmus der GeneSpring-Software vorhergesagt. Dieser Algorithmus berechnet die Ähnlichkeit (auf Basis der Euklidischen Distanz) zwischen der zu vorhersagenden Probe und den vordefinierten Gruppen des *Training-Sets* in Bezug auf die Expressionsstatus der 100 Markergene. Hierbei wird für die Zugehörigkeit der Probe zu einer der beiden Klassen je ein p-Wert kalkuliert. Nur wenn das Verhältnis der beiden p-Werte  $< 0,2$  betrug (p-Wert-Ratio  $< 0,2$ ), wurde für die Probe die Zugehörigkeit zu einer der Klassen vorhergesagt. Proben mit einer p-Wert-Ratio  $> 0,2$  wurden als nicht klassifizierbar eingestuft.

Die Ergebnisse der Markergenklassenprädiktion sind in Abbildung-5 dargestellt. Es ergab sich folgende Konstellation: Die 20 Fälle des *Test-Set* zeigten eine sehr hohe Kongruenz zwischen WHO-Diagnose und Markergenklassenprädiktion. Von den 6 der 20 Gewebeproben, bei denen alle diagnostischen Kriterien der aktuellen WHO-Klassifikation für ein BL erfüllt waren (Histomorphologie; Immunphänotyp und MYC-Translokationsstatus = MYC-positive BL) wurden alle sechs in der Markergenprädiktion als BL vorhergesagt (6/6; 100%). Von den 10 MYC-negativen DLBCL (Histologie/Immunologie: DLBCL; keine MYC-Translokation) bzw. den drei MYC-positiven DLBCL wurden insgesamt acht (8/13; 62%) als Markergen-DLBCL vorhergesagt und vier (4/13; 31%) als Markergen-BL (von denen drei eine MYC-Translokation aufwiesen) eingeordnet. Eine der 13 DLBCL-Proben des *Test-Set* sowie das MYC-positive nicht weiter klassifizierbare B-NHL zeigten bezüglich der Markergene Eigenschaften beider Lymphomentitäten und wurden deshalb vom Markergenprädiktor als nicht klassifizierbar eingestuft.

**Abbildung-5:** Differentialdiagnostische Einordnung von 40 aggressiven B-Zell-Lymphomen mittels Markergenprädiktor

A



B

**Einordnung mittels Markergenprädiktor**

		BL		DLBCL		n.k.			
		BL	DLBCL	BL	DLBCL	BL	DLBCL	n.k.	
<b>WHO-Diagnose</b>	<b>BL</b>	10	0	6	0	0			
	<b>DLBCL</b>	0	10	4	8	1			
				<b>n.k. B-NHL</b>	0	0	1		
		<b>Training-Set</b>				<b>Test-Set</b>			

Differentialdiagnostische Einordnung von 40 BL oder DLBCL mittels Markergenprädiktor. In (A) ist der Expressionsstatus von 100 Prädiktorgenen (Reihen) in 40 B-NHL (Spalten) dargestellt. Verwendet wurden Genexpressionswerte aus Affymetrix-Microarray-Analysen. Die Farbskala rechts im Bild gibt die relative Genexpressionsstärke an. Abbildung (B) zeigt den Vergleich von Diagnose mittels Genprädiktor und WHO-Diagnose in tabellarischer Darstellung. N.k.= nicht klassifizierbar

### **3.4 Der molekulare Burkitt-Lymphom (mBL)-Prädiktor – entwickelt in der Zusammenarbeit eines Verbundprojektes der Deutschen Krebshilfe**

Durch die interdisziplinäre Zusammenarbeit der Mitglieder eines Verbundprojektes der Deutschen Krebshilfe „Molekulare Mechanismen bei Malignen Lymphomen“ - in dessen Rahmen auch diese Arbeit entstanden ist – wurden neben den 40 Lymphomproben aus dieser Arbeit noch von weiteren 180 aggressiven B-NHL transkriptomweite Genexpressionsanalysen durchgeführt.

Hierbei konnte eine molekulare BL-Gensignatur (mBL-Signatur) identifiziert werden, die eine neue molekulare Definition des BL darstellt und das Spektrum der WHO-Kriterien für das BL erweitert und präzisiert <sup>57</sup>. Da die Analysen dieser Arbeit als Teil eines Verbundprojektes zur Entwicklung der mBL-Signatur beigetragen haben, gleichzeitig aber ein eigenes Projekt darstellen, wird an dieser Stelle eine Beschreibung des mBL-Prädiktors eingeführt und ein Vergleich mit dem im Rahmen der vorliegenden Arbeit erstellten Markergenprädiktor dargestellt.

#### *3.4.1 Der im Rahmen eines Verbundprojektes der Deutschen Krebshilfe entwickelte mBL-Prädiktor und Überschneidungen mit dem Markergenprädiktor der vorliegenden Arbeit*

Zur Entwicklung der mBL-Signatur wurde von der AG Bioinformatik (Leitung Dr. Rainer Spang) des Max-Planck Institutes für molekulare Genetik ein neuartiger Algorithmus (so genannter *Core-Group-Extension-Algorithmus*) entwickelt <sup>57</sup>: Basierend auf Ähnlichkeiten in der Genexpression mit hochselektierten (s.u.) BL-Proben (zusammengefasst als Kerngruppe, englisch: *core group*), identifiziert der Algorithmus andere Proben aus einem größeren Probenkollektiv, die ein ähnliches Genexpressionsverhalten zeigen. Der Algorithmus stellt anschließend eine Gruppe von Genen (Genexpressionssignatur) zusammen, die sich in dieser erweiterten Kerngruppe ähnlich verhalten.

Als „Kerngruppen-BL“ wurden acht von insgesamt 220 aggressiven B-NHL festgelegt, die strengste Kriterien für ein BL gemäß der WHO-Klassifikation der malignen Lymphome erfüllten (morphologische Einordnung eines klassischen oder atypischen BL in der *Panelreview*-Begutachtung durch die Referenzpathologen, CD20+, BCL6+, CD10+, BCL2-, CD5-, Ki-67 > 95%, IG-MYC+; Abweichungen vom Immunphänotyp wurden als absolute Ausschlusskriterien für die Kerngruppenzugehörigkeit aufgefasst). Die stringenten immunphänotypischen/

diagnostischen Anforderungen für die Kerngruppe sollten garantieren, dass die später identifizierte BL-Gensignatur nicht durch Anwesenheit von DLBCL verfälscht wird.

Der *Core-Group-Extension*-Algorithmus wurde zunächst an einem zufällig definierten *Test-Set* (n=105) eingesetzt und es konnten 64 Gene identifiziert werden, welche die molekulare Burkitt-Lymphom (mBL)-Signatur darstellten. Jeder Probe wurde ein mBL-Signatur-Index zwischen 0 und 1 zugeordnet, der die Ähnlichkeit zur Kerngruppe angab. Gewebeproben mit einem Index > 0,95 wurden als mBL bezeichnet und Gewebeproben mit einem Index < 0,05 als Non-mBL. Die verbliebenen Fälle (mBL-Index zwischen 0,05 und 0,95) erhielten die Bezeichnung „Intermediär-Fälle“.

Von den 64 Genen des mBL-Prädiktors, welche die molekulare Definition des BL darstellen, waren 59 Gene (92%) auch in der Liste der in dieser Arbeit identifizierten 2115 differentiell exprimierten Gene zu finden, welche die Grundlage für den Markergenprädiktor darstellten. Die Überschneidung zwischen dem Markergenprädiktor und dem mBL-Prädiktor betraf hingegen lediglich fünf Gene (vgl. Tabelle-6). Die Kongruenz der beiden unterschiedlichen Prädiktoren in Bezug auf die Vorhersage einer molekularen Diagnose erwies sich als sehr hoch und ist in Kapitel 3.6 dargestellt.

**Tabelle-6:** Genlistenüberschneidungen zwischen dem Markergen- und mBL-Prädiktor

<b>Schnittmenge zwischen Markergenprädiktor und mBL-Prädiktor (Gen-Name)</b>
heterogeneous nuclear ribonucleoprotein A3
formin binding protein 1
tumor necrosis factor receptor superfamily, member 5
transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
single-stranded DNA- binding protein 2

### **3.5 Der von Wright et. al. entwickelte Prädiktor zur Subgruppeneinteilung von DLBCL und sein Bezug zum Markergenprädiktor**

Im Jahre 2003 entwickelten Wright et al. auf Basis von *Microarray*-Analysen einen molekularen Klassifikator, der es ermöglichte die Subgruppen-Zugehörigkeit eines DLBCL in GCB, ABC oder Gruppe-3 anhand des Expressionsstatus von 16 Genen in 88% der Fälle richtig vorherzusagen (Wright-Klassifikator<sup>49</sup>). Um zu überprüfen ob sich diese beschriebenen molekularen Subtypen auch an dem hier verwendeten Fallkollektiv von aggressiven B-Zell-

Lymphomen vorhersagen bzw. identifizieren lassen, wurde der 16 Gene umfassende Wright-Klassifikator eingesetzt. Obwohl der Wright-Klassifikator mit Hilfe einer anderen Chip-Technik (Zwei-Farben cDNA-Chips) entwickelt worden war, ist von Wright et. al. eine Version des Prädiktors publiziert worden, der die entsprechenden Gene des cDNA-Chips den entsprechenden Genen auf dem Affymetrix U133A *GeneChip* zugeordnet hat <sup>49</sup>. Dadurch ist ein Vergleich der des Prädiktors mit anderen auf *GeneChip* U133A entwickelten Prädiktoren möglich geworden. Die Schnittmenge zwischen den 16 Genen des Wright-Prädiktors und den in dieser Arbeit initial identifizierten 2115 differentiell exprimierten Genen betrug 11 Gene. Überschneidungen zwischen dem Wright- und dem Markergenprädiktor umfassten zwei Gene (vgl. Tabelle-7/8). Ein Vergleich von Gruppen-Vorhersage des Wright- und des molekularen Prädiktors dieser Arbeit ist in Kapitel 3.6 dargestellt.

**Tabelle-7/8:** Überschneidungen zwischen den in dieser Arbeit identifizierten differentiell exprimierten Genen und dem Wright-Prädiktor

<b>Schnittmenge zwischen 2115 differentiell exprimierten Gene und Wright-Prädiktor</b> <i>(Gen-Name)</i>
B-cell linker
interferon regulatory factor 4
KIAA0870 protein
membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10)
cyclin D2
LIM domain only 2 (rhombotin-like 1)
SH3-domain binding protein 5 (BTK-associated)
pim-1 oncogene
protein tyrosine Phosphatase, non-receptor type 1
interleukin 16 (lymphocyte chemoattractant factor)
ectonucleoside triphosphate diphosphohydrolase 1

<b>Schnittmenge zwischen Markergenprädiktor und Wright-Prädiktor</b> <i>(Gen-Name)</i>
SH3-domain binding protein 5 (BTK-associated)
pim-1 oncogene

### 3.6 Vergleich der drei *Microarray*-basierten Prädiktoren und Eigenschaften der diagnostisch vorhergesagten Lymphomproben

#### 3.6.1 Markergenprädiktion: Eigenschaften der als Burkitt-Lymphom (BL) vorhergesagten Gewebeproben

Von den insgesamt 40 Gewebeproben (*Training-* und *Test-Set*) wurden mittels Markergenprädiktor insgesamt 20 als BL (Markergen-BL) vorhergesagt. Von diesen 20 Markergen-BL stellten 16 (80%) MYC-positive BL/BLL dar (Erfüllung aller WHO-Kriterien für BL oder BLL, vgl. Kapitel 3.1.4). Drei der identifizierten 20 Markergen-BL (15%) entsprachen MYC-negativen DLBCL und eine Probe einem MYC-positiven DLBCL (vgl. Tabelle-9).

Immunphänotypisch war BCL6 und CD10 in allen evaluierbaren Markergen-BL Proben konsistent exprimiert (100%). BCL2 war in den meisten Fällen negativ (15/20; 75%), auch wenn in acht Gewebeproben einzelne positive Zellen detektiert werden konnten (5/20; 15%) (vgl. Tabelle-9).

Der Vergleich der Klassifizierungsergebnisse zwischen Markergen- und mBL-Prädiktor ergab eine Übereinstimmung der molekularen Diagnose in 19 von 20 Gewebeproben (95%). Lediglich eine Gewebeprobe unterschied sich in der molekularen Diagnose der beiden Klassifikatoren: Hierbei handelte es sich um eine MYC-negative DLBCL-Probe, für die keiner der beiden Klassifikatoren die Diagnose DLBCL vorhersagte: Der mBL-Prädiktor klassifizierte die Gewebeprobe als *intermediate case* (weder der mBL- noch der Non-mBL-Gruppe zugehörig) und der Markergenprädiktor als BL (vgl. Tabelle-9).

Alle bis auf eine der 20 Markergen-BL-Proben (19/20; 95%) wurden mittels des Wright-Prädiktors als GCB-Typ eingeordnet. Lediglich eine der 20 (5%) Gewebeproben erwies sich mittels des Wright-Prädiktors als Typ-III. Diese Gewebeprobe entsprach einem MYC-positivem BL und wurde auch vom mBL-Prädiktor als mBL klassifiziert.

Der Gegenüberstellung der Diagnosevorhersage mittels der verschiedenen Prädiktoren und der histologischen Diagnose ist in Abbildung-6 und Tabelle-7 dargestellt.



### 3.6.2 *Markergenprädiktion: Eigenschaften der als diffuses großzelliges B-Zell-Lymphom vorhergesagten Gewebeproben*

18 der 40 Gewebeproben wurden mittels Markergenprädiktor als DLBCL (Markergen-DLBCL) vorhergesagt. 17 von 18 der Markergen-DLBCL (94%) erwiesen sich hierbei als MYC-negative DLBCL (histomorphologisch/immunphänotypisches Bild eines DLBCL sowie kein Nachweis einer MYC-Translokation, vgl. Kapitel 3.1.4). Eines der 18 Markergen-DLBCL (6%) entsprach einem MYC-positivem DLBCL (histomorphologisch/ immunphänotypisches Bild eines DLBCL und Nachweis einer MYC-Translokation, vgl. Kapitel 3.1.4), wobei es sich in diesem Fall um eine Non-IG-MYC-Translokation handelte.

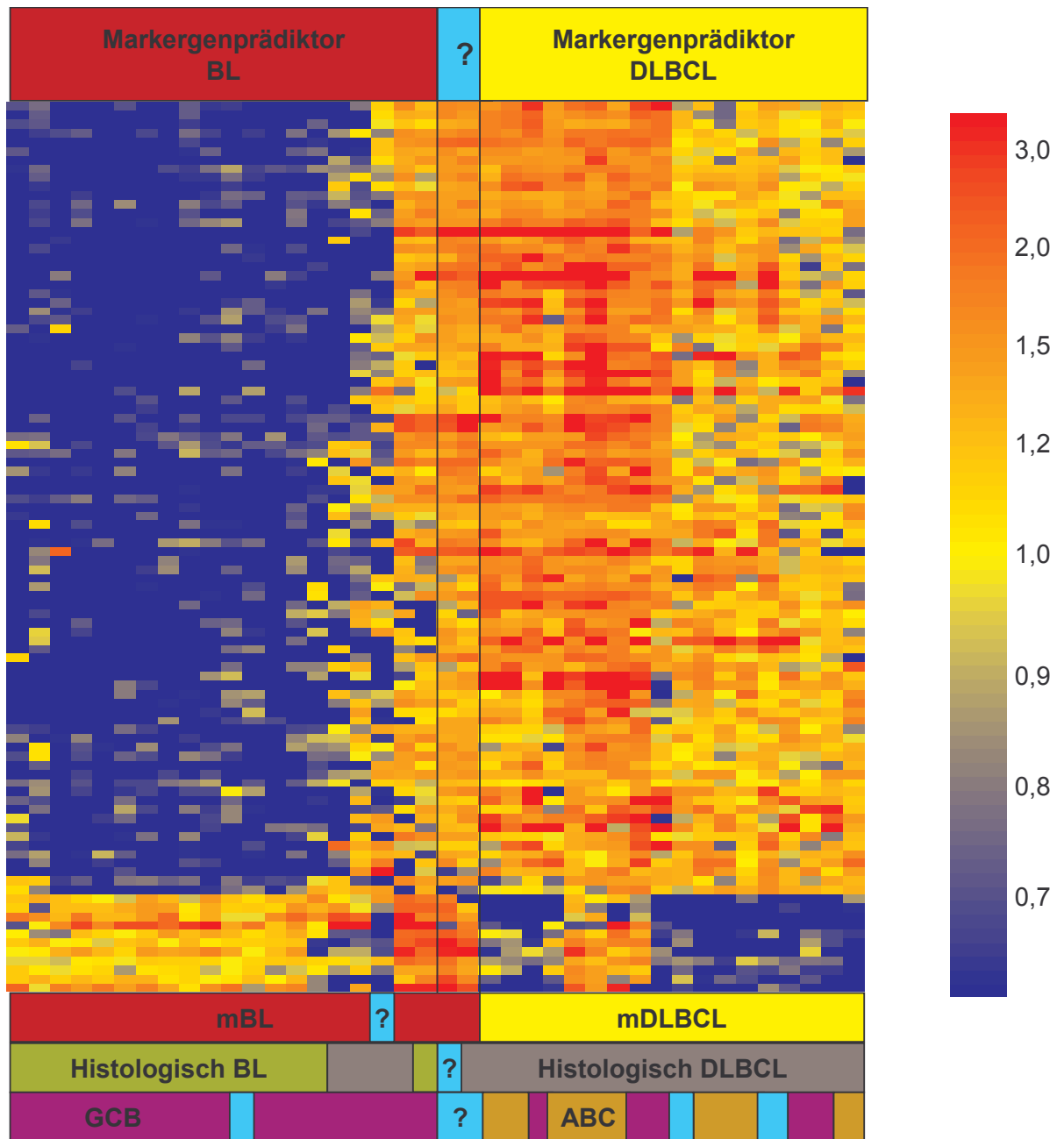
Der Vergleich der Klassenprädiktion zwischen Markergen- und mBL-Prädiktor zeigte eine hohe Kongruenz in diesem Kollektiv: Alle 18 Markergen-DLBCL-Proben wurden mittels mBL-Prädiktor übereinstimmend als Non-mBL vorhergesagt (18/18; 100%).

Die Einordnung der Markergen-DLBCL-Proben mittels Wright-Prädiktor zeigte keine eindeutige Zuordnung zu einer der vormals beschriebenen Subgruppen des DLBCL (ABC-Signatur: n=10 (56%); GCB-Signatur: n=6 (33%); Typ-III-Signatur: n=2 (11%))

### 3.6.3 *Markergenprädiktion: Eigenschaften der als „nicht klassifizierbar“ vorhergesagten Gewebeproben*

Insgesamt zwei der 40 Gewebeproben (5%) wurden mittels des Markergenprädiktors als nicht klassifizierbar eingeordnet. Die beiden Gewebeproben verhielten sich in der Expression der 100 Gene ambivalent und zeigten Merkmale beider diagnostischer Gruppen (vgl. auch Abbildungen 5 und -6). Eine der beiden Gewebeproben entsprach einem MYC-positivem DLBCL (IG-MYC-Translokation). Die zweite Gewebeprobe war interessanterweise auch in der histologischen Begutachtung durch die Referenzpathologen als nicht näher klassifizierbares aggressives B-NHL mit positivem Nachweis einer IG-MYC-Translokation beurteilt worden. Der mBL-Prädiktor vergab in diesen beiden Fällen das Merkmal mBL, der Wright-Klassifikator das Merkmal Typ-III.

**Abbildung-6:** Vergleich der Diagnosevorhersage der verschiedenen Prädiktoren



Differentialdiagnostische Einordnung von 40 BL oder DLBCL mittels Markergenprädiktor. Farblich ist der Expressionsstatus von 100 Prädiktorgenen (Reihen) in 40 B-NHL (Spalten) dargestellt (Farbskala der relativen Genexpressionsstärke rechts im Bild). Am Ober- und Unterrand ist die unterschiedliche Klassenprädiktion mit verschiedenen Prädiktoren farblich dargestellt. Oberrand: Markergenprädiktor. Unterrand (von oben nach unten): mBL-Prädiktor, histologische Diagnose, Wright-Prädiktor.

**Tabelle-9:** Morphologische, immunhistochemische, genetische und molekulare Charakteristika von aggressiven B-NHL

Charakteristika	Gesamt	Markergen- BL	Markergen- DLBCL	Markergen- nicht- klassifiziert
		<i>Anzahl (%)</i>		
<b>Total</b>	40	20	18	2
<b>Morphologie</b>				
BL	10 (25)	10 (50)	0 (0)	0 (0)
BLL	6 (15)	6 (30)	0 (0)	0 (0)
DLBCL	24 (60)	4 (20)	18 (100)	1 (50)
Reifes aggr. B-NHL/unklassifizierbar	1 (2,5)	0 (0)	0 (0)	1 (50)
<b>CD10</b>				
Pos	22 (55)	18 (90)	2 (11,1)	2 (100)
Neg	18 (45)	2 (10)	16 (88,9)	0 (0)
<b>BCL6</b>				
Pos	33 (82,5)	20 (100)	13 (65)	2 (100)
Neg	2 (5)	0 (0)	2 (11,1)	0 (0)
<b>BCL2</b>				
Pos	25 (62,5)	8 (40)	16 (88,9)	1 (50)
Neg	15 (37,5)	12 (60)	2 (11,1)	1 (50)
<b>Ki-67</b>				
≥ 90	24 (60)	20 (100)	4 (22,2)	2 (100)
≤ 90	16 (40)	0 (0)	14 (27,8)	0 (0)
<b>CD5</b>				
Pos	12 (30)	7 (35)	4 (22,2)	1 (50)
Neg	28 (70)	13 (65)	14 (27,8)	1 (50)
<b>MYC-Status</b>				
IG-MYC	19 (47,5)	17 (85)	0 (0)	2 (100)
Non-IG-MYC	1 (2,5)	0 (0)	1 (5,6)	0 (0)
MYC-Neg	20 (50)	3 (15)	17 (94,4)	0 (0)
<b>mBL-Signatur</b>				
mBL	21 (52,3)	19 (95)	0 (0)	2 (100)
Non-mBL	18 (45)	0 (0)	18 (100)	0 (0)
Intermediate	1 (2,5)	1 (5)	0 (0)	0 (0)
<b>ABC/GCB</b>				
ABC	10 (25)	0 (0)	10 (55,6)	0 (0)
GCB	25 (62,5)	19 (95)	6 (33,3)	0 (0)
Type III	5 (12,5)	1 (5)	2 (11,1)	2 (100)

## 4 Diskussion

Die präzise Unterscheidung zwischen dem Burkitt-Lymphom (BL) und dem diffusen großzelligen B-Zell-Lymphom (DLBCL) ist von großer klinischer Bedeutung, da sich die beiden Lymphomerkkrankungen prognostisch unterscheiden und darüber hinaus unterschiedlich therapiert werden<sup>5, 27, 58-60</sup>.

Obwohl in der aktuellen WHO-Klassifikation der malignen Lymphome international einheitliche Diagnosekriterien für die beiden Lymphomentitäten formuliert wurden<sup>61</sup>, stellt die Unterscheidung dieser beiden aggressiven B-NHL nach wie vor ein erhebliches differentialdiagnostisches Problem dar: Während sich ein Großteil der aggressiven B-NHL anhand der WHO-Diagnosekriterien in BL und DLBCL differenzieren lassen, ist die Unterscheidung der beiden Entitäten bei einem nicht unerheblichen Teil dieser Lymphome nicht eindeutig und daher nicht reproduzierbar<sup>62</sup>. Diesen Sachverhalt belegte unter anderem eine internationale Studie des *Lymphoma Classification Projects* recht eindrücklich: Hämatopathologen kamen in nur 53 Prozent der Fälle zu einer einheitlichen Diagnose bei der Unterscheidung von klassischen BL, atypischen BL und DLBCL<sup>63</sup>. Eine präzise und reproduzierbare Diagnose ist für die adäquate Behandlung und die Prognose der Erkrankten jedoch von entscheidender Bedeutung. Darüber hinaus ist die Reproduzierbarkeit der hämatologischen Diagnosestellung für die Vergleichbarkeit von internationalen Studien essentiell und bildet somit die Grundlage für die weitere Verbesserung der bisher etablierten Therapie und Prognose.

Es war das Ziel dieser im Rahmen eines Verbundprojektes der Deutschen Krebshilfe („Molekulare Mechanismen bei Malignen Lymphomen“) entstandenen Arbeit, durch *Microarray*-basierte Genexpressionsanalysen an Lymphomproben von BL und DLBCL eine Gruppe von Genen zu identifizieren, die eine reproduzierbare differentialdiagnostische Abgrenzung der beiden Lymphomentitäten ermöglicht.

## 4.1 Unterschiede in der Entwicklung des Markergenprädiktors und anderer molekularer Prädiktoren

Das in der vorliegenden Arbeit verwendete Probenkollektiv von 40 aggressiven B-NHL repräsentiert einen Teil des gesamten Kollektivs eines Krebshilfe-Verbundprojektes (220 Fälle). Von diesen 40 Gewebeproben wurden genomweite Genexpressionsanalysen (Affymetrix *GeneChip* U133A) erstellt. Ein *Training-Set* mit histologisch eindeutig diagnostizierbaren Gewebeproben (BL/BLL n=10; DLBCL n=10), führte zur bioinformatischen Identifizierung eines 100 Gene umfassenden molekularen Prädiktors zur differentialdiagnostischen Unterscheidung von BL und DLBCL. Die Vorhersagepräzision dieses so genannten Markergenprädiktors wurde mit Hilfe einer *Leave-One-Out*-Kreuzvalidierung überprüft und zeigte, dass alle 20 Gewebeproben des *Training-Sets* richtig diagnostisch vorhergesagt worden waren.

Dieser so entwickelte Prädiktor wurde mit den am Gesamtkollektiv des Verbundes etablierten Signaturen<sup>57, 64, 65</sup> verglichen (diskutiert in den Kapiteln 4.1.1, 4.1.2 und 4.1.3). Unterschiede in der Entwicklung der beiden Signaturen bestanden - neben der Größe des Fallkollektivs - im biostatistischen Vorgehen: Die Strategie der vorliegenden Arbeit bestand darin, zwischen histologisch eindeutigen BL und DLBCL differentiell exprimierte Gene zu identifizieren und andere aggressive B-NHL anhand dieser Gene der einen oder der anderen Entität zuzuordnen. Die Strategie des Verbundprojektes hingegen basierte auf der Idee, eine Gensignatur zu identifizieren, die charakteristisch für histologisch eindeutige BL ist und zu testen, ob andere aggressive B-NHL diese charakteristische Gensignatur ebenfalls aufweisen (mBL-Fälle) oder nicht aufweisen (Non-mBL-Fälle)<sup>57</sup>. Während in der vorliegenden Arbeit die Vorhersage der molekularen Diagnose mittels eines früher etablierten und validierten Testverfahrens erfolgte („*nearest neighbor analysis*“<sup>66</sup>, vgl. Kapitel 2.6.6), entwickelten Bioinformatiker des Max-Planck-Instituts für Molekulare Genetik Berlin ein neues statistisches Verfahren für den mBL-Prädiktor („*Core-Group-Extension-Algorithmus*“). Dieser Algorithmus ermöglichte die Berechnung eines mBL-Index, der für die Einordnung der Proben in mBL- (mBL-Index>95) bzw. Non-mBL-Fälle (mBL-Index<5) sowie Intermediate (mBL-Index: 5-95) diene. Trotz der vollkommen unterschiedlichen Ansätze in der Entwicklung der molekularen Prädiktoren, zeigte sich bezüglich der vorhergesagten molekularen Diagnose eine Übereinstimmung von 92,5% der hier untersuchten Fälle (37 von 40 Proben, vgl. Kapitel 4.2, 4.3 und 4.4).

Zeitgleich zur Entwicklung der oben beschriebenen Prädiktoren arbeitete eine unabhängige Arbeitsgruppe aus den USA um Sandeep Dave ebenfalls an der molekularen Abgrenzung des BL vom DLBCL<sup>67</sup>. Die Arbeiten von Dave et al. konnten exakt zum selben Zeitpunkt wie das oben genannte Verbundprojekt zu einem Abschluss gebracht werden und resultierten in einer „Rücken-an-Rücken-Publikation“ in derselben Ausgabe des *New England Journal of Medicine*<sup>57, 67</sup>. Die Gruppe um Dave et al. verfolgte zur Entwicklung ihres molekularen Klassifikators (im Folgenden Dave-Prädiktor genannt) wiederum einen anderen biostatistischen Ansatz: Durch den paarweisen Vergleich der Genexpressionsmuster von histologisch eindeutigen BL und den vormals definierten molekularen Subgruppen des DLBCL<sup>49</sup> (Wright et al.) wurden differentiell exprimierte Gene zur Unterscheidung definiert. Um der zentralen Rolle der MYC-Deregulation im BL gerecht zu werden, wurden die oben identifizierten Gene mit so genannten MYC-Target-Genen (Gene, die durch das MYC-Gen reguliert werden) kombiniert und daraus der endgültige molekulare Klassifikator gebildet (217 Gene).

Im Folgenden sollen die Ergebnisse der Markergenprädiktion im Vergleich mit den Ergebnissen der anderen beschriebenen molekularen Prädiktionen diskutiert werden.

## **4.2 Molekular definierte Burkitt Lymphome (Markergen-BL)**

Mittels des Markergenprädiktors ließen sich in dem Gesamtkollektiv von 40 aggressiven B-NHL-Proben (*Training- und Test-Set*) 20 Gewebeproben mit einem für das BL charakteristischen Genexpressionsmuster identifizieren (Markergen-BL). Dabei wurden alle 16 Lymphome, bei denen die sechs Referenzpathologen die WHO-Kriterien für ein BL oder BLL eindeutig erfüllt sahen (MYC-positive BL, n=10; MYC-positive BLL, n=10), vom molekularen Prädiktor als Markergen-BL korrekt vorhergesagt. Dieser Befund spiegelt die hohe Sensitivität des Markergenprädiktors bei der Identifizierung von BL wider.

Insgesamt vier von 20 (20%) Markergen-BL wiesen jedoch eine anders lautende pathologische Diagnose auf. Während eine dieser vier Proben ein MYC-positives DLBCL darstellte (histologisch/immunologisch: DLBCL, zytogenetisch: IG-MYC-Translokation), entsprachen drei Proben MYC-negativen DLBCL (histologisch/immunologisch: DLBCL, zytogenetisch: keine MYC-Translokation). Des Weiteren zeigten insgesamt 5 der 20 Markergen-BL eine Expression des BCL2-Proteins. Es lässt sich somit zusammenfassen, dass 5 der identifizierten Markergen-BL Merkmale aufweisen, die gemäß der klassischen WHO-Diagnosekriterien nicht vereinbar mit der Diagnose eines BL gewesen wären<sup>3</sup>. Im Zusammenhang mit diesem Befund erscheint es

aufschlussreich, inwieweit die Klassifizierung des unabhängigen mBL-Prädiktors mit der Markergenprädiktion übereinstimmt: Hierbei zeigt sich interessanterweise eine 95%ige Übereinstimmung (19 von 20 Fällen) zwischen den beiden molekularen Prädiktoren (lediglich eine der 20 Markergen-BL klassifizierte der mBL-Prädiktor als Intermediärfall).

Die hohe Kongruenz der beiden molekularen Prädiktoren kann als gegenseitige Validierung aufgefasst werden. Beide Prädiktoren zeigen somit unabhängig voneinander, dass es sich bei mindestens 2 der 3 MYC-negativen DLBCL sowie dem MYC-positiven DLBCL um BL handelt, die mit herkömmlichen Diagnoseverfahren bislang nicht diagnostizierbar waren. Betrachtet man das Gesamtkollektiv aus dem genannten Verbundprojekt (n=220, davon n=44 mit mBL-Diagnose), so bestätigen sich diese Ergebnisse auch in einem größeren Maßstab: Insgesamt 11 von 44 (25%) identifizierten molekularen BL-Fällen wiesen eine DLBCL-Morphologie auf, vier von 44 (9%) zeigten außerdem einen negativen MYC-Translokationsstatus und neun von 44 Fälle (20%) waren BCL2 positiv<sup>57</sup>.

Bezüglich der morphologischen und immunphänotypischen Charakteristika von molekular definierten BL kamen Dave et al. in ihrer vollkommen unabhängig und zeitgleich entstandenen Arbeit zu ähnlichen Ergebnissen: Insgesamt 7 der 53 (13%) molekular definierten BL wiesen in der histomorphologischen Betrachtung das typische Bild eines DLBCL auf. Und ebenfalls 7 von 53 (13%) molekularen BL wiesen darüber hinaus die Expression von BCL2 auf. Entgegen den Ergebnissen der vorliegenden Arbeit bzw. des Verbundprojektes identifizierten Dave et al. in ihrer Arbeit jedoch keine MYC-negativen BL<sup>67</sup>. Dieser Befund wird bei Betrachtung des biostatistischen Vorgehens von Dave et al. zur Entwicklung ihres molekularen Prädiktors jedoch plausibel: Da die Wissenschaftler zuvor MYC-Target-Gene als Teil ihrer molekularen BL-Signatur selektionierten, wird der MYC-Status konsekutiv zu einem wichtigeren Kriterium für die Gruppenunterteilung.

Neben der oben dargestellten genexpressionsbasierten Definition des Markergen-BL, war in dieser Arbeit weiterhin der Frage nachgegangen worden, in welcher Beziehung das BL zu den durch Wright et al. identifizierten molekularen Subtypen des DLBCL (GCB, ABC, Gruppe-III<sup>49</sup>, vgl. Kapitel 1.5) steht. Wie in Kapitel 3.6.1 dargestellt erbrachte die Klassifizierung der 20 Markergen-BL durch den Wright-Prädiktor eine eindeutige Präferenz für das GCB-Merkmal (19 von 20; 95%). Dieser Befund, der sich auch in dem größeren Verbundkollektiv wieder finden lässt (GCB: 40 von 44; 91%), scheint die schon vormals beschriebene biologische Keimzentrums-Herkunft des BL widerzuspiegeln<sup>68-71</sup>. Die gemeinsame Expression von

Keimzentrums-B-Zell-Genen könnte somit auf eine engere biologische Verwandtschaft oder eine gemeinsame B-Zell-Differenzierungsstufe zwischen BL und GCB-DLBCL hinweisen. Die Tatsache, dass GCB-DLBCL - ebenso wie BL - gegenüber anderen DLBCL eine günstigere klinische Prognose aufweisen sollen<sup>48, 49, 57, 72, 73</sup>, unterstützt diese Hypothese. Der Wright-Prädiktor stellt somit ein Instrument dar, das in aggressiven B-NHL Rückschlüsse auf die biologische Herkunft der Zellen geben kann. Die Abgrenzung des BL/BLL vom DLBCL jedoch kann vom Wright-Prädiktor erwartungsgemäß nicht geleistet werden. Gleichzeitig kann somit die eingangs in dieser Arbeit gestellte Frage nach der biologischen Beziehung der molekularen Subtypen des DLBCL mit dem BL als beantwortet betrachtet werden (vgl. Kapitel 1.6.3).

Zusammenfassend lässt sich feststellen, dass die WHO-Klassifikation des BL durch den in der vorliegenden Arbeit entwickelten Markergenprädiktor auf Lymphome mit DLBCL-Morphologie, negativem MYC-Status sowie BCL2-Expression erweitert wird. Diese Ergebnisse wurden durch den Vergleich mit den oben erwähnten zeitgleich und unabhängig entstandenen anderen Arbeiten bestätigt. Des Weiteren ergibt die Einordnung der Markergen-BL mittels des vormals beschriebenen Wright-Prädiktors eine klare Präferenz für die GCB-Gruppe und weist somit auf die schon vormals beschriebene Keimzentrumsherkunft der BL hin. Eine engere biologische Verwandtschaft zwischen dem GCB-DLBCL und dem BL ist deshalb anzunehmen.

### **4.3 Molekular definierte diffuse großzellige B-Zell-Lymphome (Markergen-DLBCL)**

Insgesamt 18 der 40 untersuchten B-NHL-Proben zeigten in der Markergenanalyse das charakteristische Genexpressionsmuster eines DLBCL und wurden dementsprechend als Markergen-DLBCL klassifiziert. In allen 18 Fällen (18/18; 100%) stimmte die Markergendiagnose mit der ursprünglichen histologischen Diagnose der Referenzpathologen überein. Unter Berücksichtigung der zytogenetischen Analysen wiesen dabei 17 von 18 Proben (94%) keine MYC-Translokation auf (MYC-negative DLBCL), wohingegen lediglich eine Gewebeprobe (1/18; 6%) eine MYC-positive DLBCL darstellte. In diesem Fall handelte es sich interessanterweise jedoch nicht um eine – für das BL typische - IG-MYC-Translokation sondern um die Translokation des MYC-Gens an einen Non-Immunglobulin-Lokus (Non-IG-MYC). Keines der in dieser Arbeit identifizierten Markergen-DLBCL wies also eine BL/BLL-Morphologie bzw. eine IG-MYC-Translokation auf oder umgekehrt betrachtet: Alle B-NHL mit



BL-Morphologie wurden vom Markergenprädiktor ebenfalls als BL klassifiziert (hingegen zeigten 20% der Markergen-BL eine DLBCL-Morphologie, vgl. Kapitel 4.2).

Im Vergleich der molekularen Diagnosen von Markergenprädiktor und mBL-Prädiktor<sup>57</sup> wurden alle 18 Markergen-DLBCL übereinstimmend als Non-mBL vorhergesagt (Übereinstimmung: 18/18; 100%). Diese hohe Kongruenz zwischen den unabhängig voneinander entwickelten Prädiktoren unterstreicht die hohe Präzision der molekularen Diagnosevorhersagen, die in dem vorliegenden Kollektiv – wie oben dargestellt – zu 100% mit der histologischen Diagnose übereinstimmt. Der in dieser Arbeit beobachtete Trend, dass die BL-Morphologie sehr stark, die DLBCL hingegen wesentlich schwächer mit der jeweils gleichnamigen molekularen Vorhersage korreliert, ließ sich auch im Gesamtverbundkollektiv wieder finden (lediglich 3 von 128 (2%) Non-mBL-Fällen: BL/BLL-Morphologie vs. 11 von 44 (25%) mBL-Fällen: DLBCL-Morphologie). Die IG-MYC-Translokation wurde im Verbundkollektiv bei Fällen mit Non-mBL-Signatur allerdings deutlich häufiger gefunden als in der vorliegenden Arbeit (5 von 128 (4%) Non-mBL-Fällen: IG-MYC-positiv).

Wie in Kapitel 3.6.2 dargestellt brachte der Vergleich der 18 Markergen-DLBCL mit dem Wright-Prädiktor-Merkmal<sup>49</sup> erwartungsgemäß keine präferentielle Zuordnung zu einem der vormals beschriebenen Subgruppen des DLBCL (ABC-Signatur: n=10 (56%); GCB-Signatur: n=6 (33%); Typ-III-Signatur: n=2 (11%) vgl. Kapitel 3.6.2). Dieser Befund zeigt, dass sich Gewebeproben mit Markergen-DLBCL-Merkmal noch weiter in verschiedene molekulare Subgruppen unterteilen lassen. Wie retrospektive Untersuchungen im Rahmen des erwähnten Verbundprojektes zeigen, könnte diese weitere Unterteilung von klinisch-prognostischer Bedeutung sein: Gewebeproben mit Non-mBL-Signatur und GCB-Signatur wiesen gegenüber Non-mBL-Fällen mit ABC-Signatur eine signifikant längeres Überleben nach Chemotherapie auf<sup>48, 72</sup>.

Die Ergebnisse der Arbeit von Dave et al.<sup>67</sup> zeigen bezüglich Morphologie und Zytogenetik ähnliche Ergebnisse wie die Arbeit des erwähnten Verbundprojektes: So fand sich in dem großen Kollektiv von 242 molekular diagnostizierten DLBCL lediglich eine Gewebeprobe mit der Morphologie eines BL/BLL und 7% der molekularen DLBCL zeigten eine MYC-Translokation.

Zusammenfassend lässt sich feststellen, dass Lymphome mit dem Markergenprofil eines DLBCL in der vorliegenden Arbeit - in allen Fällen - auch eine DLBCL-Morphologie aufwiesen.

Umgekehrt zeigten Gewebeproben mit DLBCL-Morphologie in 17% der Fälle eine Markergen-BL-Signatur. Wie in Kapitel 4.2 dargestellt wiesen hingegen Gewebeproben mit BL-Morphologie in 100% der Fälle auch in der Markergenprädiktion das BL-Merkmal auf. Diese Beobachtungen ließen sich im Trend auch in der Arbeit des Verbundprojektes und in der Arbeit von Dave et al. wieder finden (B-NHL mit DLBCL-Morphologie und molekularem BL-Profil: vorliegende Arbeit: 4/20 (20%); Verbundprojekt: 11/44 (25%), Dave et al.: 13/53(13%) vs.: Gewebeproben mit BL-Morphologie und molekularem DLBCL-Profil; Markergenprädiktor: 0/20 (0%); mBL-Prädiktor: 3/128 (2%); Dave-Prädiktor: 1/45 (2%)).

Für die Praxis ließe sich daraus folgende Hypothese ableiten: Eine typische BL-Morphologie korreliert sehr gut mit der molekularen Einordnung als BL (mindestens 98% der Fälle) und die Diagnose eines BL scheint auch ohne weitere molekulare Einordnung angemessen. Liegt hingegen die Morphologie eines DLBCL vor, verbirgt sich dahinter in etwa 20% (Median) der Fälle ein molekulares BL und eine molekulare Evaluation des Lymphoms scheint zur adäquaten Diagnosestellung unabdingbar.

Weiter lässt sich feststellen, dass sich Gewebeproben mit der molekularen Signatur eines Markergen-DLBCL bzw. eines Non-mBL durch Anwendung des Wright-Prädiktors weiter in molekulare Subtypen unterteilen lassen. Angesichts der vormals beschriebenen klinischen Relevanz dieser Einteilung<sup>48, 49, 72</sup> erscheint die Möglichkeit einer Kombination des Markergenprädiktors bzw. des mBL-Prädiktors mit dem Wright-Prädiktor in weiterführenden Analysen prüfenswert.

#### **4.4 Molekular nicht-klassifizierbare aggressive B-Zell-Non-Hodgkin Lymphome (nicht klassifizierbare B-NHL)**

Nicht alle 40 Gewebeproben der vorliegenden Arbeit konnten in der Markergenprädiktion eindeutig einer der beiden oben genannten diagnostischen Gruppen zugeordnet werden: In dem Kollektiv von 40 untersuchten aggressiven B-NHL fanden sich zwei Gewebeproben, die hinsichtlich der Expression der Markergene Eigenschaften beider diagnostischer Gruppen aufwiesen und demzufolge vom Markergenprädiktor als „nicht-klassifizierbar“ eingestuft wurden. Verglichen mit der pathologischen Diagnose entsprach eine der Proben einem MYC-positivem DLBCL (IG-MYC-Translokation), wohingegen die zweite der beiden Proben auf Grund der uncharakteristischen Morphologie von den Pathologen schon zuvor als nicht näher klassifizierbares aggressives B-NHL eingestuft worden war. Auch diese Probe wies in der

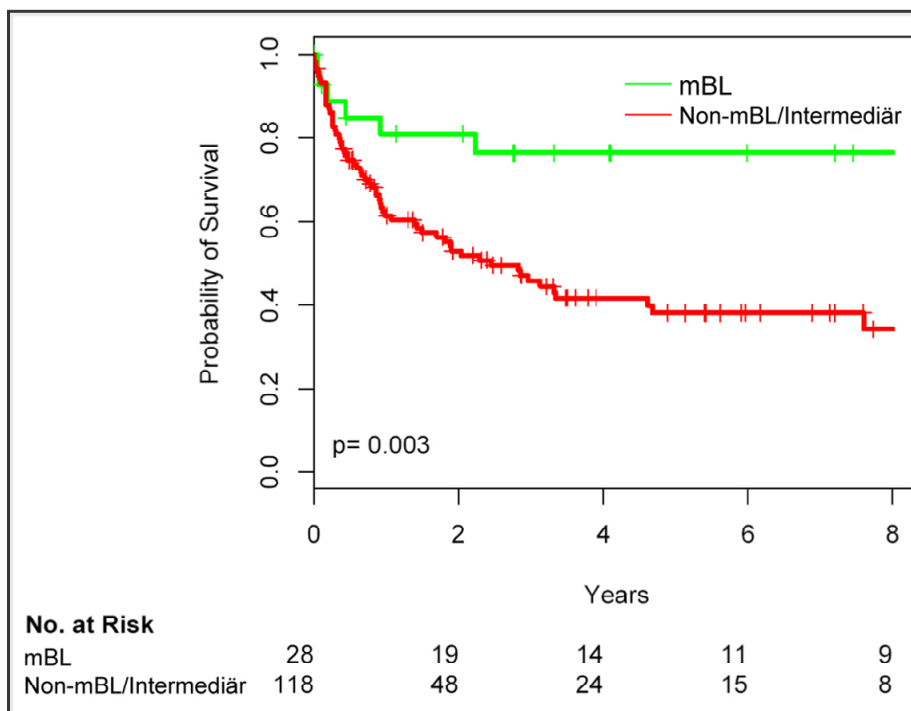
zytogenetischen Analyse eine IG-MYC-Translokation auf. Es ist bemerkenswert, dass beide nicht klassifizierbare Proben das zytogenetische Charakteristikum von BL/BLL tragen (IG-MYC-Translokation), jedoch bezüglich ihrer Morphologie eindeutig nicht in diese Kategorie fallen. Betrachtet man die Expressionsstärke der 100 Markergene, so fällt auf, dass die beiden Gewebeproben sowohl die BL-typischen Gene, als auch die DLBCL-typischen Gene exprimieren (vgl. Abb. 6). Auf Grund dieser Befunde könnte man spekulieren, dass es sich bei den beiden Proben quasi um biologische Hybride handelt, welche – vielleicht ursprünglich als DLBCL – das genetische Markenzeichen der BL (die IG-MYC-Translokation) akquiriert haben und somit biologische Eigenschaften beider Lymphomentitäten aufweisen.

Interessanterweise ließen sich die beiden nicht klassifizierbaren Gewebeproben auch in der Wright-Prädiktion weder der GCB- noch der ABC-Gruppe eindeutig zuordnen und erhielten dementsprechend das Merkmal „Klasse-III-DLBCL“. Auch dieser Befund könnte auf die biologische Ambivalenz dieser beiden Gewebeproben hinweisen.

Der im Rahmen des erwähnten Verbundprojektes entwickelte mBL-Prädiktor klassifizierte die beiden Gewebeproben als mBL-Fälle, was möglicherweise durch die Expression der BL-spezifischen Gene (s.o.) bzw. der Expression der MYC-Target-Gene dieser beiden Proben zu erklären ist. Trotz der enorm hohen Übereinstimmung zwischen Markergen- und mBL-Prädiktor bezüglich der Diagnosevorhersage der untersuchten 40 Proben ist grundsätzlich anzumerken, dass die Verwendung geringerer Fallzahlen naturgemäß mit einer geringeren Signifikanz einhergeht. Vor diesem Hintergrund scheint es nicht ausgeschlossen, dass die Diskrepanz in der Diagnosevorhersage dieser zwei Fälle auf eine geringere Signifikanz des Markergenprädiktors gegenüber dem mBL-Prädiktor zurückzuführen ist. Betrachtet man jedoch das Gesamtkollektiv des Verbundprojektes (n=220), so zeigen auch hier immerhin 48 (22%) Gewebeproben ein ambivalentes Genexpressionsmuster und wurden dementsprechend als Intermediate-Fälle klassifiziert. In diesem Zusammenhang ist auch die im Rahmen des Verbundprojektes durchgeführte Analyse des Langzeitüberlebens der zugehörigen Patienten nach Chemotherapie unter Berücksichtigung ihrer Gruppenzugehörigkeit interessant: Während sich die Non-mBL und Intermediate-Fälle klinisch ähnlich verhalten, so zeigen diese beiden Gruppen kombiniert, gegenüber den mBL-Fällen eine signifikant schlechtere Prognose (5-Jahres-Überleben: mBL: 75% vs. Non-mBL/Intermediate: 39%; p=0,003; vgl. Abbildung-7). Die molekulare Einordnung der beiden in der Markergenanalyse nicht identifizierbaren Gewebeproben muss auf Grund der

kontroversen Einordnung der Proben durch die unterschiedlichen Prädiktoren jedoch als nicht geklärt betrachtet werden.

**Abbildung-7:** Kaplan-Meier-Schätzung des Überlebens entsprechend der molekularen Signatur



Gesamtüberleben von Patienten mit mBL-Signatur war signifikant höher als von Patienten mit Non-mBL/Intermediär-Lymphomsignatur ( $p=0,003$  mittels log-rank test). Die Strichmarkierungen kennzeichnen lebende Patienten zum Zeitpunkt der letzten Nachsorgeuntersuchung (nach Hummel et al.<sup>57</sup>)

Zusammenfassend lässt sich sagen, dass sowohl die vorliegende als auch die im Rahmen des genannten Verbundprojektes entstandene Arbeit eine Gruppe von aggressiven B-NHL identifizierten, welche biologische Eigenschaften von BL und DLBCL aufwiesen. Während retrospektive Überlebensanalysen der zugehörigen Patienten ein signifikant besseres Überleben von Patienten mit mBL-Profil gegenüber mBL/Intermediate-Profil aufwiesen und somit die hohe klinische Relevanz der molekularen Klassifikatoren untermauerte, wies die Intermediate-Gruppe gegenüber der Non-mBL-Gruppe keinen signifikanten prognostischen Unterschied auf. Ob Patienten mit einem nicht eindeutig klassifizierbaren Genexpressionsprofil von einer anderen Chemotherapie profitieren, kann jedoch nur anhand von prospektiven kontrollierten klinischen Studien überprüft werden.

**Tabelle-10:** Diagnostische Einordnung der 40 Gewebeproben mit verschiedenen Klassifizierungsalgorithmen

Proben-ID und WHO-Diagnose	Markergen-Prädiktor	mBL-Prädiktor	Wright-Prädiktor
ID-017-BL (MYC+)	Markergen-BL	mBL	GCB
ID-032 BL (MYC+)	Markergen-BL	mBL	GCB
ID -033 BL (MYC+)	Markergen-BL	mBL	GCB
ID -077 BL (MYC+)	Markergen-BL	mBL	GCB
ID -079 BL (MYC+)	Markergen-BL	mBL	Typ-III
ID -080 BL (MYC+)	Markergen-BL	mBL	GCB
ID -084 BL (MYC+)	Markergen-BL	mBL	GCB
ID -085 BL (MYC+)	Markergen-BL	mBL	GCB
ID -089 BL (MYC+)	Markergen-BL	mBL	GCB
ID -114 BL (MYC+)	Markergen-BL	mBL	GCB
ID -001 BLL (MYC+)	Markergen-BL	mBL	GCB
ID -002 BLL (MYC+)	Markergen-BL	mBL	GCB
ID -027 BLL (MYC+)	Markergen-BL	mBL	GCB
ID I-043 BLL (MYC+)	Markergen-BL	mBL	GCB
ID -044 BLL (MYC+)	Markergen-BL	mBL	GCB
ID -082 BLL (MYC+)	Markergen-BL	mBL	GCB
ID -006 DLBCL (MYC-)	Markergen-BL	mBL	GCB
ID -018 DLBCL (MYC +)	Markergen-BL	mBL	GCB
ID -065 DLBCL (MYC-)	Markergen-BL	mBL	GCB
ID -008 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -009 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -010 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -013 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -039 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -040 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -063 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	Typ-III
ID -073 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	Typ-III
ID -074 DLBCL (MYC +)*	Markergen-DLBCL	Non-mBL	ABC
ID -113 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -118 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -119 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -130 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -132 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -133 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -134 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -135 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	ABC
ID -136 DLBCL (MYC-)	Markergen-DLBCL	Non-mBL	GCB
ID -034 DLBCL (MYC+)	Nicht-klassifizierbar	mBL	Typ-III
ID -087 B-NHL**	Nicht-klassifizierbar	mBL	Typ-III

BL = Burkitt-Lymphom; DLBCL = diffuses großzelliges B-Zell-Lymphom; mBL = molekulares mBL; Non-mBL = Non-molekulares BL; GCB = *germinal center B-cell-type*; ABC = *activated B-cell type*

\* Non-IG-MYC-Translokation; \*\*nicht weiter klassifizierbares MYC-positives aggressives B-NHL

## 5 Zusammenfassung

Hintergrund: Die differentialdiagnostische Abgrenzung des Burkitt-Lymphoms (BL) vom diffusen großzelligen B-Zell-Lymphom (DLBCL) ist bislang unscharf definiert. Die derzeit gültigen WHO-Diagnosekriterien ermöglichen eine reproduzierbare Diagnosestellung basierend auf Morphologie, Immunphänotyp und Zytogenetik nur in einem Teil der Fälle. Da die beiden Lymphomentitäten unterschiedlich therapiert werden, ist diagnostische Präzision jedoch von entscheidender Bedeutung. In der vorliegenden Arbeit wurde untersucht, ob sich durch transkriptomweite Genexpressionsanalysen von Lymphomgewebeproben eine verlässliche Unterscheidung der beiden Entitäten etablieren lässt.

Methoden: Von 294 reifen aggressiven B-Zell-Lymphomgewebeproben wurden 40 Proben mit einem Tumorzellgehalt > 70% selektioniert und immunhistochemisch charakterisiert. Affymetrix U133A *GeneChips* wurden verwendet, um von den ausgewählten 40 Lymphomgewebeproben genomomweite Genexpressionsanalysen zu erstellen.

Ein *Training-Set* mit Gewebeproben, die mittels WHO-Diagnosekriterien eindeutig in BL und DLBCL zu unterteilen waren, wurde verwendet, um zwischen den beiden diagnostischen Gruppen differentiell exprimierte Markergene zu identifizieren. Der Expressionsstatus der besten 100 Markergene wurde anschließend für die Vorhersage einer molekularen Diagnose aller Gewebeproben verwendet. Die - als Teil eines Verbundprojektes der Deutschen Krebshilfe - in dieser Arbeit generierten Ergebnisse aus Untersuchungen an 40 Lymphomproben wurden mit weiterführenden Analysen dieses Gesamtverbundprojektes an zusätzlichen 180 Lymphomproben verglichen.

Ergebnisse: Anhand der Markergenexpression wurden 20 Gewebeproben des Kollektivs (n=40) als BL identifiziert: Darin eingeschlossen waren alle 16 Proben, die entsprechend der WHO-Kriterien zuvor eindeutig als BL oder BLL diagnostiziert worden waren. Drei der molekular definierten BL zeigten das morphologische Bild eines DLBCL und ein Fall das Bild eines unklassifizierbaren aggressiven B-Zell-Lymphoms. Zytogenetisch wiesen 17 der 20 molekular definierten BL eine IG-MYC-Translokation auf.

18 Gewebeproben klassifizierte der im Rahmen dieser Arbeit entwickelte 100 Gene umfassende Markergenprädiktor als DLBCL. Alle 18 Fälle zeigten histologisch das Bild eines DLBCL, wobei ein Fall gleichzeitig eine IG-MYC-Translokation aufwies. Insgesamt zwei der 40 untersuchten Gewebeproben ordnete der Prädiktor als unklassifizierbar ein, da sie ein

Genexpressionsprofil mit Eigenschaften beider Lymphomentitäten zeigten. Histomorphologisch entsprach eine der beiden Proben einem unklassifizierbarem aggressiven B-Zell-Lymphom, die zweite Probe zeigte die Morphologie eines DLBCL. Der zytogenetische MYC-Translokationsnachweis war bei diesen beiden Fällen positiv.

Die in dieser Arbeit identifizierten molekularen Gruppierungen wurden außerdem mit den im Rahmen des oben genannten Verbundprojektes unabhängig ermittelten molekularen Diagnosen verglichen. Hierbei ergab sich eine Übereinstimmung in 37 von 40 Fällen (93%).

Konklusion: Der Markergenprädiktor repräsentiert eine neue Signatur, um das BL vom DLBCL auf Basis ihrer Genexpression differentialdiagnostisch reproduzierbar abzugrenzen. Dabei erweitert der Markergenprädiktor die WHO-Definition des BL auf Gewebeproben mit der Morphologie eines DLBCL, auf Proben ohne IG-MYC-Translokation sowie auf Proben mit Expression des BCL2-Proteins. Eine direkte Übertragbarkeit dieser Befunde in die klinische Praxis sollte - unter Berücksichtigung der hochsignifikanten Ergebnisse des Verbundprojektes - in prospektiven kontrollierten Studien überprüft werden.

## 6 Abkürzungsverzeichnis

ABC-DLBCL	<i>activated B-cell-like DLBCL</i> , DLBCL vom aktivierten B-Zell-Typ
ACTB	<i>beta-actin</i> , beta-Aktin
ALL	Akute lymphatische Leukämie
ANOVA	<i>analysis of variance</i> , Varianzanalyse
APAAP-Methode	Alkalische Phosphatase Anti-Alkalische Phosphatase Methode
BL	<i>Burkitt lymphoma</i> , Burkitt-Lymphom
BLL	<i>Burkitt-like lymphoma</i> , atypisches Burkitt-Lymphom
cDNA	<i>complementary DNA</i> , komplementäre DNA
cRNA	<i>complementary RNA</i> , komplementäre RNA
DEPC	Diethylpyrocarbonat
DLBCL	<i>diffuse large B-cell lymphoma</i> , diffuses großzelliges B-Zell-Lymphom
DNA	<i>deoxyribonucleic acid</i> , Desoxyribonukleinsäure
dNTP	Desoxy-Nukleotidtriphosphat
DTT	Dithiothreitol
EBV	Epstein-Barr-Virus
EST	<i>expressed sequence tag</i>
FISH	Fluoreszenz-in-situ-Hybridisierung
GAPDH	Glyceraldehyd-3-Phosphat-Dehydrogenase
GCB	<i>germinal center B-cell-type DLBCL</i> , DLBCL vom Keimzentrumstyp
Ig	Immunglobulin
IgA,-D,-E,-G,-M	Ig mit verschiedenen konstanten Regionen
IPI	<i>international prognostic index</i>
ISGF-3	<i>interferon-stimulated-transcription-factor-3</i>
IVT	in-vitro-Transkription
mBL	<i>molecular BL</i> , molekulares BL
MES	2-Morpholinoethansulfonsäure
MM	<i>mismatch</i>
mRNA	<i>messenger RNA</i> , Boten RNA
NFκB	<i>nuclear factor kappa B</i>
NHL	Non-Hodgkin-Lymphom
NTP	Nukleosidtriphosphat



OD	optische Dichte
PM	<i>perfect match</i>
RNA	<i>ribonucleic acid</i> , Ribonukleinsäure
RNase	<i>ribonuclease</i> , Ribonuklease
RPMI	<i>Roswell Park Memorial Institute</i> (Zellkulturmedium)
SSPE	<i>Sodium chloride-sodium hydrogen phosphate-EDTA</i> (Puffer)
TBE	Tris-Borat-EDTA (Puffer)
WHO	<i>World Health Organization</i> , Weltgesundheitsorganisation
ZNS	Zentralnervensystem

## 7 Literaturverzeichnis

- (1) BURKITT DP. Primary hydrocele and its treatment. Review of two hundred cases. *Lancet* 1951 June 23;1(25):1341-3.
- (2) O'CONNOR GT. Malignant lymphoma in African children. II. A pathological entity. *Cancer* 1961 March;14:270-83.:270-83.
- (3) Diebold J, Jaffe ES, Raphael M, Warnke RA. Burkitt Lymphoma. In: Jaffe ES, Harris NL, Stein H, Vardiman J, editors. *World Health Organization Classification of Tumours, Pathology and Genetics, Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press: Lyon; 2001. p. 181-4.
- (4) A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood* 1997 June 1;89(11):3909-18.
- (5) Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. *Blood* 2004 November 15;104(10):3009-20.
- (6) Harris NL, Jaffe ES, Stein H et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood* 1994 September 1;84(5):1361-92.
- (7) Zech L, Haglund U, Nilsson K, Klein G. Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas. *Int J Cancer* 1976 January 15;17(1):47-56.
- (8) Magrath I. The pathogenesis of Burkitt's lymphoma. *Adv Cancer Res* 1990;55:133-270.:133-270.
- (9) Kelly K, Siebenlist U. Mitogenic activation of normal T cells leads to increased initiation of transcription in the c-myc locus. *J Biol Chem* 1988 April 5;263(10):4828-31.
- (10) Kelly K, Underwood B. Cell growth associated regulation of c-myc and c-fos in normal human T cells. *Adv Exp Med Biol* 1987;213:241-7.
- (11) Kelly K, Siebenlist U. The regulation and expression of c-myc in normal and malignant cells. *Annu Rev Immunol* 1986;4:317-38.
- (12) Kelly K, Siebenlist U. The role of c-myc in the proliferation of normal and neoplastic cells. *J Clin Immunol* 1985 March;5(2):65-77.
- (13) Armelin HA, Armelin MC, Kelly K et al. Functional role for c-myc in mitogenic response to platelet-derived growth factor. *Nature* 1984 August 23;310(5979):655-60.
- (14) Adams JM, Harris AW, Pinkert CA et al. The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* 1985 December 12;318(6046):533-8.

- (15) Murphy SB, Hustu HO. A randomized trial of combined modality therapy of childhood non-Hodgkin's lymphoma. *Cancer* 1980 February 15;45(4):630-7.
- (16) BURKITT D. A sarcoma involving the jaws in African children. *Br J Surg* 1958 November;46(197):218-23.
- (17) Wright D, McKeever P, Carter R. Childhood non-Hodgkin lymphomas in the United Kingdom: findings from the UK Children's Cancer Study Group. *J Clin Pathol* 1997 February;50(2):128-34.
- (18) Wright DH. Burkitt's lymphoma: a review of the pathology, immunology, and possible etiologic factors. *Pathol Annu* 1971;6:337-63.:337-63.
- (19) Bittencourt AL, Araujo I, Carvalho EG, Vieira MG, Hummel M. Burkitt's lymphoma of the scapula. *Pediatr Pathol Mol Med* 2003 July;22(4):271-6.
- (20) Soussain C, Patte C, Ostronoff M et al. Small noncleaved cell lymphoma and leukemia in adults. A retrospective study of 65 adults treated with the LMB pediatric protocols. *Blood* 1995 February 1;85(3):664-74.
- (21) Reiter A, Schrappe M, Ludwig WD et al. Favorable outcome of B-cell acute lymphoblastic leukemia in childhood: a report of three consecutive studies of the BFM group. *Blood* 1992 November 15;80(10):2471-8.
- (22) Thomas DA, Faderl S, O'Brien S et al. Chemoimmunotherapy with hyper-CVAD plus rituximab for the treatment of adult Burkitt and Burkitt-type lymphoma or acute lymphoblastic leukemia. *Cancer* 2006 April 1;106(7):1569-80.
- (23) Cohen LF, Balow JE, Magrath IT, Poplack DG, Ziegler JL. Acute tumor lysis syndrome. A review of 37 patients with Burkitt's lymphoma. *Am J Med* 1980 April;68(4):486-91.
- (24) Gatter KC, Warnke RA. Diffuse large B-cell lymphoma. In: Jaffe ES, Harris NL, Stein H, Vardiman J, editors. *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon: IARC Press: Lyon; 2001. p. 171-4.
- (25) Weiss LM, Warnke RA, Sklar J, Cleary ML. Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas. *N Engl J Med* 1987 November 5;317(19):1185-9.
- (26) Kramer MH, Hermans J, Wijburg E et al. Clinical relevance of BCL2, BCL6, and MYC rearrangements in diffuse large B-cell lymphoma. *Blood* 1998 November 1;92(9):3152-62.
- (27) Coiffier B, Lepage E, Briere J et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* 2002 January 24;346(4):235-42.
- (28) A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* 1993 September 30;329(14):987-94.

- (29) Milpied N, Deconinck E, Gaillard F et al. Initial treatment of aggressive lymphoma with high-dose chemotherapy and autologous stem-cell support. *N Engl J Med* 2004 March 25;350(13):1287-95.
- (30) Braziel RM, Arber DA, Slovak ML et al. The Burkitt-like lymphomas: a Southwest Oncology Group study delineating phenotypic, genotypic, and clinical features. *Blood* 2001 June 15;97(12):3713-20.
- (31) Frost M, Newell J, Lones MA, Tripp SR, Cairo MS, Perkins SL. Comparative immunohistochemical analysis of pediatric Burkitt lymphoma and diffuse large B-cell lymphoma. *Am J Clin Pathol* 2004 March;121(3):384-92.
- (32) Akasaka T, Akasaka H, Ueda C et al. Molecular and clinical features of non-Burkitt's, diffuse large-cell lymphoma of B-cell type associated with the c-MYC/immunoglobulin heavy-chain fusion gene. *J Clin Oncol* 2000 February;18(3):510-8.
- (33) Vitolo U, Gaidano G, Botto B et al. Rearrangements of bcl-6, bcl-2, c-myc and 6q deletion in B-diffuse large-cell lymphoma: clinical relevance in 71 patients. *Ann Oncol* 1998 January;9(1):55-61.
- (34) Haralambieva E, Boerma EJ, van Imhoff GW et al. Clinical, immunophenotypic, and genetic analysis of adult lymphomas with morphologic features of Burkitt lymphoma. *Am J Surg Pathol* 2005 August;29(8):1086-94.
- (35) Cogliatti SB, Novak U, Henz S, Schmid U, Moller P, Barth TF. Diagnosis of Burkitt lymphoma in due time: a practical approach. *Br J Haematol* 2006 August;134(3):294-301.
- (36) DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 1997 October 24;278(5338):680-6.
- (37) Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000 January 7;100(1):57-70.
- (38) Yeoh EJ, Ross ME, Shurtleff SA et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002 March;1(2):133-43.
- (39) Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med* 2001 December 17;194(12):1861-74.
- (40) 't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002 January 31;415(6871):530-6.
- (41) Wiestner A, Rosenwald A, Barry TS et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood* 2003 June 15;101(12):4944-51.
- (42) Bullinger L, Dohner K, Bair E et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004 April 15;350(16):1605-16.

- (43) Rosenwald A, Wright G, Wiestner A et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell* 2003 February;3(2):185-97.
- (44) Chen HY, Yu SL, Chen CH et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 2007 January 4;356(1):11-20.
- (45) Alizadeh AA, Eisen MB, Davis RE et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000 February 3;403(6769):503-11.
- (46) Savage KJ, Monti S, Kutok JL et al. The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. *Blood* 2003 December 1;102(12):3871-9.
- (47) Shipp MA, Ross KN, Tamayo P et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002 January;8(1):68-74.
- (48) Rosenwald A, Wright G, Chan WC et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002 June 20;346(25):1937-47.
- (49) Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* 2003 August 19;100(17):9991-6.
- (50) Rosenwald A, Wright G, Leroy K et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J Exp Med* 2003 September 15;198(6):851-62.
- (51) Monti S, Savage KJ, Kutok JL et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood* 2004 November 18;.
- (52) Staudt LM, Dave S. The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv Immunol* 2005;87:163-208.
- (53) Davis RE, Staudt LM. Molecular diagnosis of lymphoid malignancies by gene expression profiling. *Curr Opin Hematol* 2002 July;9(4):333-8.
- (54) Küppers R, Klein U, Schwering I et al. Identification of Hodgkin and Reed-Sternberg cell-specific genes by gene expression profiling. *J Clin Invest* 2003 February;111(4):529-37.
- (55) Freeman WM, Robertson DJ, Vrana KE. Fundamentals of DNA hybridization arrays for gene expression analysis. *Biotechniques* 2000 November;29(5):1042-55.
- (56) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B* 1995;57:289-300.

- (57) Hummel M, Bentink S, Berger H et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* 2006 June 8;354(23):2419-30.
- (58) Coiffier B. State-of-the-art therapeutics: diffuse large B-cell lymphoma. *J Clin Oncol* 2005 September 10;23(26):6387-93.
- (59) Bishop PC, Rao VK, Wilson WH. Burkitt's lymphoma: molecular pathogenesis and treatment. *Cancer Invest* 2000;18(6):574-83.
- (60) Butler RD, Hainsworth JD. Optimal therapy for small noncleaved cell lymphoma. *Cancer Treat Res* 1993;66:65-79.
- (61) Jaffe ES, Harris NL, Stein H, Vardiman J. *World Health Organization Classification of Tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues*. Lyon, France: IARC Press: Lyon 2001; 2001.
- (62) Magrath I, Jaffe E.S., Bhatia K. Burkitt's Lymphoma. In: Knowles DM, editor. *Neoplastic Hematopathology*. Philadelphia: Lippinkott Williams & Wilkins; 2001. p. 953-86.
- (63) A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood* 1997 June 1;89(11):3909-18.
- (64) Stein H, Hummel M. [Burkitt's and Burkitt-like lymphoma : Molecular definition and value of the World Health Organisation's diagnostic criteria.]. *Pathologie* 2007 January 11.
- (65) Stein H, Hummel M. Histopathology in the light of molecular profiling. *Ann Oncol* 2006 May;17 Suppl 4:iv5-iv7.
- (66) Belacel N, Wang Q, Cuperlovic-Culf M. Clustering methods for microarray gene expression data. *OMICS* 2006;10(4):507-31.
- (67) Dave SS, Fu K, Wright GW et al. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med* 2006 June 8;354(23):2431-42.
- (68) Chapman CJ, Wright D, Stevenson FK. Insight into Burkitt's lymphoma from immunoglobulin variable region gene analysis. *Leuk Lymphoma* 1998 July;30(3-4):257-67.
- (69) Tamaru J, Hummel M, Marafioti T et al. Burkitt's lymphomas express VH genes with a moderate number of antigen-selected somatic mutations. *Am J Pathol* 1995 November;147(5):1398-407.
- (70) Chapman CJ, Wright D, Feizi HP, Davis Z, Stevenson FK. V(H) gene analysis of Burkitt's lymphoma in children from north-western Iran. *Br J Haematol* 1998 December;103(4):1116-23.
- (71) Bellan C, Lazzi S, Hummel M et al. Immunoglobulin gene analysis reveals 2 distinct cells of origin for EBV-positive and EBV-negative Burkitt lymphomas. *Blood* 2005 August 1;106(3):1031-6.

- (72) Rosenwald A, Staudt LM. Clinical translation of gene expression profiling in lymphomas and leukemias. *Semin Oncol* 2002 June;29(3):258-63.
- (73) Bea S, Zettl A, Wright G et al. Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood* 2005 November 1;106(9):3183-90.

## 8 Danksagung

Danken möchte ich zunächst Herrn Prof. Dr. Harald Stein für die freundliche Überlassung des Themas sowie die Unterstützung und Gesprächsbereitschaft, die er mir stets entgegengebracht hat.

Mein besonderer Dank gilt Herrn PD Dr. Michael Hummel. Er hat mich bei der Erstellung dieser Arbeit betreut, in jeder Hinsicht unterstützt und mich die Grundlagen und Feinheiten des wissenschaftlichen Arbeitens und Schreibens gelehrt. Durch seine außergewöhnliche wissenschaftliche Kompetenz, aber auch durch seine Freundlichkeit und Führung ist er mir zum Vorbild geworden.

Herrn Prof. Dr. Christoph Loddenkemper danke ich für die Anleitung und Lehre bei der histologischen Begutachtung der Lymphomgewebeproben und die vielen persönlichen Gespräche, die mir immer sehr gut getan haben.

Besondere Dankesworte möchte ich an Hedwig Lammert für die Einführung in die RNA- und *Microarray*-Techniken richten und alle weiteren Hilfestellungen, die sie so zahlreich waren, dass ich sie nicht einzeln benennen kann. Ein ganz besonderes Dankeschön auch an Erika Berg für die große Hilfe bei den immunhistochemischen Färbungen und Henning Müller für die Einführung in die PCR-Techniken. Für die täglich kompetente fachliche Beratung und das freundschaftliche Verhältnis, das die Arbeit sehr angenehm gemacht hat, sei den Dres. rer. nat. Dido Lenze, Anke Ehlers und Ulrike Treichel herzlich gedankt. Im Zusammenhang mit der guten Arbeitsatmosphäre sei auch allen anderen Mitarbeitern des Instituts für Pathologie gedankt.

Danken möchte ich auch meinem Cousin Benedikt Weber sowie meinen Freunden Jan Dehne-Niemann, Robert Glumm und Dirk Schnapauff für ermutigende Gespräche und Ablenkung zum rechten Zeitpunkt.

Meiner Frau Selina danke ich für Unterstützung und Beratung, die mich getragen hat und für ihre liebevolle Geduld, wenn es im Labor mal wieder später wurde.

Ferner danke ich meinen Eltern Ingrid und Karlwilhelm sowie meinen Geschwistern Angela und Alexander für die moralische und fachliche Hilfe.



## **9 Lebenslauf**

Mein Lebenslauf wird aus Datenschutzgründen in der elektronischen Version meiner Arbeit nicht mit veröffentlicht.

## 10 Publikation

Hummel M, Bentik S, Berger H, Klapper W, Wessendorf S, Barth TFE, Bernd HW, Cogliatti SB, Dierlamm J, Feller AC, Hansmann ML, Haralambieva E, Harder L, **Kühn M**, Lenze D, Lichter P, Ignacio MS, Möller P, Müller-Hermelink HK, Ott G, Parwaresch RM, Pott C, Rosenwald A, Rosolowski M, Schwaenen C, Stürzenhofecker B, Szcapanowski M, Trautmann H, Wacker HH, Spang R, Loeffler M, Stein H, Siebert R  
A Biologic Definition of Burkitt's Lymphoma from Transcriptional and Genomic Profiling.  
*New England Journal Medicine* 2006; 354:2419-30

## 11 Erklärung

„Ich, Michael Kühn, erkläre, dass ich die vorgelegte Dissertationsschrift mit dem Thema: „Molekulare Abgrenzung des Burkitt-Lymphoms vom diffusen großzelligen B-Zell-Lymphom durch Microarray-basierte Genexpressionsanalysen“ selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, ohne die (unzulässige) Hilfe Dritter verfasst und auch in Teilen keine Kopien anderer Arbeiten dargestellt habe.“

Datum

Unterschrift