

**Identifizierung altersassoziierter
genetischer und epigenetischer Aberrationen
in reifzelligen Keimzentrums-B-Zell-Lymphomen**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Ulrike Paul

Kiel, 2016

Erster Gutachter: Prof. Dr. rer. nat. Manuela Dittmar

Zweiter Gutachter: Prof. Dr. med. Wolfram Klapper

Tag der mündlichen Prüfung: 10.06.2016

Zum Druck genehmigt: 10.06.2016

Gez. der Dekan Prof. Dr. rer. nat Wolfgang J. Duschl

Inhaltsverzeichnis

| | |
|----------------------------------------------------------------------------------------------------------------------------------------|-----------|
| Glossar | I |
| Abkürzungen | II |
| 1 Einleitung | 1 |
| 1.1 Immunsystem und Alternsprozess | 1 |
| 1.2 Lymphatisches System | 3 |
| 1.2.1 Keimzentrums-B-Zell-Entwicklung | 4 |
| 1.3 Keimzentrums-B-Zell-Lymphome | 8 |
| 1.3.1 Follikuläre Lymphome | 8 |
| 1.3.2 Diffus großzellige B-Zell-Lymphome | 9 |
| 1.3.3 Therapie und klinische Studien von Patienten mit einem Keimzentrums-B-Zell-Lymphom | 10 |
| 1.4 Molekulare Pathogenese und Forschungsstudien von Patienten mit einem Keimzentrums-B-Zell-Lymphom | 13 |
| 1.5 Keimzentrums-B-Zell-Lymphome im Senium | 18 |
| 1.6 Hochdurchsatz-Sequenzierung zur Identifikation von genetischen Aberrationen bei Patienten im Senium | 22 |
| 1.6.1 Target-Resequenzierung mit dem Verfahren von Illumina | 22 |
| 1.7 Methylierungsanalysen zur Identifikation von epigenetischen Aberrationen bei Patienten im Senium | 25 |
| 1.7.1 Histonmodifikation | 25 |
| 1.7.2 DNA-Methylierung..... | 26 |
| 1.7.3 Infinium HumanMethylation450 BeadChIP..... | 27 |
| 1.8 Ziele der Arbeit | 29 |
| 2 Material und Methoden | 31 |
| 2.1 Material | 31 |
| 2.1.1 Probenmaterial für die Optimierung der Hochdurchsatz-Sequenzierung | 32 |
| 2.1.1.1 Patienten mit einer T-Prolymphozyten-Leukämie..... | 32 |
| 2.1.1.2 Keimzentrums-B-Zell-Lymphom-Zelllinien | 32 |
| 2.1.1.3 FFPE-Material von Nicht-Tumor-Proben..... | 33 |
| 2.1.1.4 Extrahierte DNA aus peripherem Blut | 33 |
| 2.1.2 Patienten-Kohorten für die altersassoziierten Analysen an Keimzentrums-B-Zell-Lymphomen | 35 |
| 2.1.2.1 Vergleichskohorte von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus verschiedenen klinischen und Forschungsstudien | 36 |
| 2.1.2.2 Sehr alte Patienten mit extrahierter DNA aus FFPE-Tumormaterial (<i>Jack</i>) | 42 |
| 2.1.2.3 Übersicht aller analysierten senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom (<i>DLBCL/FL-Senil</i>) | 48 |
| 2.1.2.4 Validierungskohorte für das Kandidatengen <i>CD58</i> | 49 |
| 2.1.2.5 Kontrollen für die Analyse genetischer und epigenetischer Aberrationen | 50 |
| 2.1.3 Primer | 52 |
| 2.1.4 Reagenzien..... | 54 |
| 2.1.5 Enzyme und Kits..... | 54 |
| 2.1.6 Puffer und Lösungen..... | 55 |
| 2.1.7 Geräte..... | 56 |
| 2.1.8 Software und Datenbanken..... | 57 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 2.2 Methoden | 58 |
| 2.2.1 DNA Extraktion | 59 |
| 2.2.1.1 DNA Extraktion aus Frischmaterial | 59 |
| 2.2.1.2 DNA Extraktion aus FFPE-Tumormaterial | 59 |
| 2.2.2 Qualitätskontrolle der DNA- und PCR-Fragmente | 60 |
| 2.2.2.1 Quantifizierung der DNA und PCR-Fragmente mittels Qubit | 60 |
| 2.2.2.2 Qualitätskontrolle der extrahierten DNA mittels quantitativer Real Time PCR | 61 |
| 2.2.2.3 Qualitätskontrolle und Quantifizierung von PCR-Fragmenten mittels Bioanalyzer | 62 |
| 2.2.3 Klassisches Verfahren der Sanger-Sequenzierung | 63 |
| 2.2.3.1 Polymerasekettenreaktion und Agarose-Gelelektrophorese | 64 |
| 2.2.3.2 DNA Sequenzierung nach Sanger | 65 |
| 2.2.4 Hochdurchsatz-Sequenzierung | 67 |
| 2.2.4.1 PCR-basierte Anreicherungsverfahren | 69 |
| 2.2.4.2 Hybridisierungs-basierte Anreicherungsverfahren | 77 |
| 2.2.4.3 Normalisierung der Library und PhiX-Kontrolle | 91 |
| 2.2.4.4 Erstellen von <i>Sample Sheet</i> und <i>Manifest</i> für die Sequenzierung | 93 |
| 2.2.4.5 Sequenziervorgang mit dem MiSeq | 94 |
| 2.2.5 Bioinformatische Auswertung der Hochdurchsatz-Sequenzierungs-Daten | 94 |
| 2.2.5.1 Qualitätskontrolle des MiSeq-Laufs mittels der Software <i>Sequencing Analysis Viewer</i> von Illumina | 98 |
| 2.2.5.2 Qualitätskontrolle der mittels Hochdurchsatz-Sequenzierung identifizierten Mutationen im <i>Integrativen Genome Browser (IGV)</i> des Broad Instituts | 99 |
| 2.2.5.3 Analyse mittels der <i>Software Variant Studio</i> von Illumina | 99 |
| 2.2.5.4 Analyse mittels der Software <i>Sequence Pilot</i> von JSI medical systems | 100 |
| 2.2.5.5 Analyse mittels der <i>inhouse Pipeline</i> des Deutschen Krebsforschungszentrums | 101 |
| 2.2.6 Analyse der DNA-Methylierung mittels Methylierungsarrays | 105 |
| 2.2.6.1 Natriumbisulfit-Konvertierung von DNA-Proben | 105 |
| 2.2.6.2 Infinium HumanMethylation450 BeadChip (450K) | 105 |
| 2.2.7 Bioinformatische Auswertung der DNA-Methylierungsdaten | 106 |
| 2.2.7.1 Qualitätskontrolle und Analyse der Rohdaten | 108 |
| 2.2.7.2 Methylierungsanalyse von Keimzentrums-B-Zell-Lymphomen | 111 |
| 2.2.8 Vergleiche der Hochdurchsatz-Sequenzierungs- und DNA-Methylierungs-Daten | 112 |
| 2.2.8.1 Rekurrent mutierte Gene in Patienten mit einem hohen biologischen und epigenetischen Alter | 113 |
| 2.2.8.2 Gene die das epigenetische Alter von Keimzentrums-B-Zell-Lymphomen beeinflussen | 113 |
| 2.2.9 Statistische Auswertung | 114 |
| 2.2.9.1 Logistische Regression | 114 |
| 2.2.9.2 Mann-Whitney U Test | 115 |
| 2.2.9.3 Kruskal-Wallis Test | 116 |
| 2.2.9.4 Clustering zur Assoziationsanalyse | 116 |
| 2.2.10 Sonstige Verfahren zur Visualisierung und Auswertung | 117 |
| 2.2.10.1 Scatterplot-Boxplot-Grafiken | 117 |
| 2.2.10.2 MutationMapper zur Darstellung von Mutationen im Gen-Kontext | 117 |
| 2.2.10.3 Stringanalyse und Venn-Diagramme | 118 |

| | | |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 3 | Ergebnisse | 119 |
| 3.1 | Austestung und Optimierung verschiedener methodischer Verfahren zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung | 120 |
| 3.1.1 | Myeloid-Panel (<i>TruSight Amplicon Panel</i>) | 123 |
| 3.1.2 | <i>ID3</i> -Assay (<i>Two-Step-PCR-Assay</i>)..... | 124 |
| 3.1.3 | Cancer-Panel (<i>TruSight Rapid Capture Panel</i>) | 129 |
| 3.1.4 | Lymphom-Panel (<i>SureSelectXT Panel</i>) | 131 |
| 3.2 | Altersassoziierte genetische und epigenetische Aberrationen bei reifzelligen Keimzentrums-B-Zell-Lymphomen | 135 |
| 3.2.1 | Genetische Aberrationen bei reifzelligen Keimzentrums-B-Zell-Lymphomen..... | 137 |
| 3.2.1.1 | Charakterisierung genetischer Aberrationen bei folliculären Lymphomen..... | 138 |
| 3.2.1.2 | Charakterisierung genetischer Aberrationen bei diffus großzelligen B-Zell- Lymphomen..... | 144 |
| 3.2.1.3 | Vergleich von altersassoziierten genetischen Markern bei folliculären und diffus großzelligen B-Zell-Lymphomen | 166 |
| 3.2.2 | Epigenetische Aberrationen bei reifzelligen Keimzentrums-B-Zell-Lymphomen | 169 |
| 3.2.2.1 | Qualitätskontrolle der DNA-Methylierungsdaten des gesamten Datensatzes | 170 |
| 3.2.2.2 | Differenziell methylierte Regionen in diffus großzelligen B-Zell-Lymphomen | 172 |
| 3.2.2.3 | Epigenetisches Alter nach Horvath von Patienten mit einem diffus großzelligen B- Zell-Lymphom und Normalkontroll-Proben | 176 |
| 3.2.2.4 | Rekurrent mutierte Gene mit einem hohen biologischen und epigenetischen Alter bei Patienten mit einem diffus großzelligen B-Zell-Lymphom..... | 178 |
| 3.2.2.5 | Rekurrent mutierte Gene die das epigenetischen Alter bei Patienten mit einem diffus großzelligen B-Zell-Lymphom beeinflussen | 180 |
| 4 | Diskussion | 183 |
| 4.1 | Methodische Aspekte zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung.. | 184 |
| 4.2 | Senile Patienten mit einem reifzelligen Keimzentrums-Lymphom | 188 |
| 4.2.1 | Lymphom-Entitäten und Geschlechterverteilung | 189 |
| 4.2.2 | Genetische Aberrationen in diffus großzelligen B-Zell-Lymphomen vom ABC-Subtyp | 192 |
| 4.2.3 | Mutationen und strukturelle Veränderungen im Kandidatengen <i>CD58</i> | 197 |
| 4.2.4 | Epigenetische Aberrationen in diffus großzelligen B-Zell-Lymphomen | 198 |
| 5 | Ausblick..... | 202 |
| 6 | Zusammenfassung..... | 203 |
| 7 | Summary | 205 |
| 8 | Referenzen..... | 207 |
| 9 | Anhang | 222 |
| 10 | Danksagung..... | 237 |
| 11 | Lebenslauf | 239 |
| 12 | Publikationsliste | 240 |
| 13 | Erklärung..... | 242 |

Für eine bessere Übersicht sind im Inhaltsverzeichnis nur die Hierarchieebenen 1 bis 4 aufgelistet.

Glossar

Adapter:

Spezielle Oligonukleotide von ca. 30 bp, die als Startpunkt für eine Barcode-PCR benötigt werden und am 5' Ende der DNA angefügt sind.

Alignment:

Ein Vergleich von 2 oder mehr Sequenzabfolgen zur Identifikation von Mutationen in den Nukleotidsequenzen..

Assay:

In der vorliegenden Arbeit wird der Begriff für die Analyse von einem Gen mittels Hochdurchsatz-Sequenzierung verwendet. Die Analyse von mehr als einem Gene wird als Panel bezeichnet.

AVG-Beta-Wert:

Der AVG-Beta-Wert bezeichnet den Methylierungswert und stellt den durchschnittlich ermittelten Beta-Wert über einen spezifischen Locus in der Methylierungsanalyse einer Probe dar.

Bam-File/Datei (engl. Binary Alignment/Map format):

Eine binär kodierte Datei mit den ermittelten Konsensus-Sequenzen aus der Hochdurchsatz-Sequenzierung.

Barcode/Index:

Ein Oligonukleotid mit einer spezifischen/einzigartigen Sequenz von 6-8 Nukleotiden, das an die Nukleotidsequenz jeder zu analysierenden Probe angefügt wird. Mit Hilfe dieser spezifischen Oligonukleotidsequenzen ist es möglich zeitgleiche mehrere Proben zu sequenzieren und in der anschließenden bioinformatischen Auswertung die Sequenzinformationen dem jeweiligen Patienten zuzuordnen. Basierend auf dem verwendeten Anreicherungsverfahren werden an beiden oder nur an einem Ende der DNA-Fragmente die spezifischen Oligonukleotide angefügt.

Batch-Effek:

Die Unterschiede zwischen zwei gleichen Analysen von dem gleichen Ausgangsmaterialien. Die Differenzen entstehen z.B. durch die Verwendung von Reagenzien, die zu unterschiedlichen Zeitpunkten (Chargen) hergestellt wurden und mit verschiedenen Lotnummern gekennzeichnet sind.

Cluster PF:

Die Anzahl an klonalen Sequenzen, die den Illumina-spezifischen Qualitätsfilter passieren und in die weitere Analyse eingehen.

Cluster:

Ein einzelner Bereich von klonalen (identischen) Sequenzen auf dem Glasobjektträger.

Coverage:

Die Anzahl wie oft eine spezifische genomische Zielregion innerhalb der Hochdurchsatz-Sequenzierung abgelesen wurde (Sequenziertiefe).

Custom Panel:

Eine Zusammensetzung von Amplikons oder Oligonukleotiden zur Abdeckung von spezifischen Zielregionen, die speziell für die eigenen Anforderungen/Fragestellungen erstellt werden und nicht durch andere Nutzer vorgetestet sind.

Demultiplexen:

Der Prozess der Zuordnung der Sequenzinformationen nach der Sequenzierung zu den jeweiligen Proben, wobei die Proben-spezifischen Oligonukleotide aus den Sequenzen entfernt werden.

Detection p-value (Detpval):

Der Vergleich der Fluoreszenzintensität der analysierten Loci auf dem Array mit internen Negativkontrollen der Herstellerfirma Illumina in der GenomeStudio Software.

Dropout:

Nicht sequenzierte Target-Region in der Hochdurchsatz-Sequenzierung.

Duplicate Rate:

Der prozentuale Anteil einer Probe an PCR-Duplikaten gegenüber der Gesamtanzahl aller Sequenzen. PCR-Duplikate sind Fragmente mit einer identischer Sequenzlänge und Nukleotidabfolge, die bei der Amplifikation der DNA-Fragmente in der Probenvorbereitung erzeugt werden.

Enrichment:

Eine gezielte Anreicherung von Regionen aus der genomischen DNA der zu analysierenden Probe.

Fastq-File/Datei:

Eine textbasierte Datei in der die Rohdaten der Hochdurchsatz-Sequenzierung abgelegt werden.

Flow cell:

Ein Glasobjektträger mit spezifisch gebundenen Oligonukleotiden auf der Oberfläche, die komplementär zu den Sequenzier-Adaptoren der DNA-Fragmente einer Sequenzierbibliothek sind. Während der Hochdurchsatz-Sequenzierung finden auf diesem Träger die Brückenamplifikation und der anschließende Sequenzierprozess statt.

HotSpot Mutation:

Eine bekannte Position in der genomischen DNA, die häufig mutiert vorliegt.

Input:

Die Einsatzmenge bzw. Voraussetzung, die benötigt wird um eine Methode zu beginnen oder einen Prozess zu starten. In der vorliegenden Arbeit werden die benötigte/verwendete Menge an DNA zur Erstellung einer Sequenzierbibliothek (methodisch) und das SampleSheets bzw. Manifest für die Hochdurchsatz-Sequenzierung (technisch) als Input bezeichnet.

In-silico

Ein Vorgang bzw. Prozess der nur theoretisch am Computer unter Verwendung spezieller Software und Algorithmen durchgeführt wurde und nicht praktisch im Labor ausgetestet ist.

Library:

Ein fertiger Pool von DNA-Fragmente nach der Erstellung der Sequenzierbibliothek, der zur Hochdurchsatz-Sequenzierung auf den Glasobjektträger geladen wird.

Loci detection rate (LDR):

Der Quotient aus tatsächlich analysierten Loci und der Gesamtanzahl auf dem Infinium HumanMethylation450 Bead Chip (450K) abgedeckter Loci in Prozent.

Manifest:

Eine textbasierte Datei, die alle genomischen Zielregionen von Interesse (engl. region of interest, ROI) eines zu analysierenden Panels enthält. In der Datei werden dazu das Chromosom, sowie der genomische Start- und Endpunkt der jeweiligen Zielregion angegeben.

Mapping:

Der Vergleich der Konsensus-Sequenz gegen ein Referenzgenom (hg19).

Mapping-Rate:

Das prozentuale Verhältnis der Sequenzen einer Probe die dem Referenzgenom zugeordnet werden, gegenüber den nicht zuordenbaren Sequenzen.

Mutationsallelfrequenz (MAF):

Der prozentuale Anteil mit welcher Häufigkeit ein bestimmtes Allel in einem Gen des analysierten Patienten mutiert vorliegt. Die MAF liegt im Normalfall für heterozygote Mutationen bei ~ 50% und für homozygote Mutationen bei 100%. Bei Mutationsanalysen von Tumoren ist die MAF abhängig vom Tumorzellgehalt der zu analysierenden Probe, da häufig auch Zellen aus dem Normalgewebe in der Probe enthalten sind. Weiterhin kann die MAF bei Tumoranalysen aufgrund der Heterogenität bzw. Klonalität der zu analysierenden Probe verändert sein ($\leq 50\%$ bzw. $\leq 100\%$).

Mutationsfrequenz (MF):

Der prozentuale Anteil an Patienten mit einer Mutation in dem analysierten Gen gegenüber der Gesamtanzahl an analysierten Patienten.

Natasha (engl. Nucleotide Assessment of Tumor Alternations for Somatic point mutations with High Accuracy):

Ein *inhouse* erstelltes Programm des deutschen Krebsforschungszentrums in Heidelberg zur Identifizierung von somatischen Punktmutationen bei Mutationsanalysen von Hochdurchsatz-Sequenzierungs-Daten.

Nicht-Duplikat-Reads:

Die Anzahl an Sequenzen einer Probe, die keine PCR-Duplikate sind und somit nicht die gleichen Sequenzlängen und Nukleotidabfolgen aufweisen.

On-Target-Reads:

Die Anzahl an Sequenzen einer Probe, die eindeutig einer Zielregion zugeordnet werden.

Ø-On-Target-Coverage/ Ø-OTC:

Die durchschnittliche Sequenziertiefe aller tatsächlich sequenzierten Zielregionen einer Probe.

Ø-On-Target-Rate/ Ø-OTR:

Das Verhältnis der zu sequenzierenden Zielregion zu den tatsächlich sequenzierten Zielregion in Prozent. Hierbei werden nur vollständig sequenzierte Zielregion, in denen alle Basenpaare sequenziert wurden, einberechnet.

Ø-Readlänge:

Die durchschnittliche Sequenzlänge aller tatsächlich sequenzierten Zielregionen einer Probe.

Output:

Die Ausgabe bzw. der Ergebnis-Ordner einer Hochdurchsatz-Sequenzierung, der alle generierten Dateien enthält.

Paired-End:

Ein Sequenzierungsverfahren, wobei die DNA-Fragmente der Sequenzierbibliothek von beiden Leserichtungen analysiert werden (3'- und 5'-Richtung)

Panel:

Als Panel wird die Gesamtheit von vielen Oligonukleotiden für die gezielte Anreicherung von spezifischen genomischen Regionen bezeichnet. Mit Hilfe dieses Panels ist eine gleichzeitige Analyse von zahlreichen genomischen Regionen möglich.

Passing Filter (PF):

Der Prozess während der ersten 25 Sequenzierzyklen der Hochdurchsatz-Sequenzierung indem alle Cluster aus der weiteren Analyse ausgeschlossen werden bei denen mehr als 1 Base der Sequenzabfolge mit einem *chastity* value unter 0,6 detektiert wurden. Die *chastity* ist das Verhältnis der Fluoreszenzintensität aus der Base mit der am höchsten gemessenen Intensität und der Summe aus den Basen der höchsten und zweit höchsten Intensität.

PhiX:

Eine fertige Sequenzierbibliothek des PhiX Virus der Firma Illumina, die kommerziell erhältlich ist. PhiX wird als Kontrolle in geringer Konzentration bei den Hochdurchsatz-Sequenzierungen mitgeführt, um technische Artefakte der besser beurteilen zu können.

Pipeline

Eine automatisierte Abfolge von Arbeitsschritten zur bioinformatischen Auswertung der generierten Daten. In der vorliegenden Arbeit werden die Analysen der Hochdurchsatz-Sequenzierungs-Daten und DNA-Methylierungsdaten als Pipeline bezeichnet.

Platypus

Ein Programm des Wellcome Trust Centre for Human Genetics der Universität von Oxford identifiziert zur Identifizierung von somatischen Insertionen oder Deletionen (Indels) bei Mutationsanalysen von Hochdurchsatz-Sequenzierungs-Daten.

Q30-Wert:

Der Q30-Wert ist ein Qualitätsparameter der Firma Illumina, der den prozentualen Anteil an Basen angibt, deren Fehlerrate geringer ist als 1 Fehler in 1000 Basen.

Read:

Eine generierte Nukleotidabfolge (Sequenz) nach der Hochdurchsatz-Sequenzierung.

Reads Identified:

Der prozentuale Anteil der gefilterten Reads, die von dem Software-Programm MiSeq-Reporter nach dem Demultiplexen eindeutig einer Probe zugeordnet werden können.

Reproduzierbarkeit:

In der vorliegenden Arbeit wird die Reproduzierbarkeit aus dem Vergleich der mittels Hochdurchsatz-Sequenzierung analysierten Duplikate berechnet. Hierzu wurden einzelne Proben zweifach (Duplikat) angereichert und sequenziert. Die identifizierten Mutationsallelfrequenzen (MAF) wurden anschließend miteinander verglichen und der Korrelationskoeffizient zu einander bestimmt. Alle Mutationen mit einem Korrelationskoeffizienten größer als 0,9 wurden als reproduzierbar bezeichnet.

rs-Loci:

Der Locus der einen bekannten Single Nukleotid Polymorphismus (SNP) abdeckt.

SampleSheet:

Datei in der die verwendeten Proben-spezifischen Oligonukleotide den jeweiligen Proben zugeordnet werden. Zusätzlich werden die Sequenzier- und Auswertungsparameter mit angegeben wie z.B. Anreicherungsverfahren, Sequenzierlänge, Referenzgenom, Manifest).

Sensitivität (richtig Positiv-Rate):

In der vorliegenden Arbeit wird die Sensitivität aus dem Vergleich der mittels Hochdurchsatz- und Sanger-Sequenzierung detektierten Mutationen ermittelt. Alle Mutationen, die in der Hochdurchsatz- und Sanger-Sequenzierung identifiziert wurden, sind richtig positiv.

Sequenzier-Adapter:

Spezielle Oligonukleotide von ca. 30 bp die zusätzlich am 5'-Ende der DNA-Fragmente mit den Proben-spezifischen Oligonukleotiden angefügt sind. Mit Hilfe dieser Oligonukleotide bindet die Sequenzierbibliothek an die komplementären Oligonukleotiden auf der Oberfläche des Glasobjektträgers.

Single Nukleotid Polymorphismus (SNP):

In der vorliegenden Arbeit werden alle Sequenzveränderungen, die auch in der Normalbevölkerung mit einer Häufigkeit von über 1% auftreten, als SNPs bezeichnet.

Skript:

Der Quellcode eines Programms, wobei es sich um eine Folge von Programmieranweisungen handelt, die einen Algorithmus darstellen. Algorithmen werden zur Automatisierung von Prozessen erstellt, die immer nach dem gleichen Prinzip ablaufen (z.B. Statistische Analysen, Mutationsanalysen).

Spezifität (richtig Negativ-Rate):

In der vorliegenden Arbeit wird die Spezifität aus dem Vergleich der mittels Hochdurchsatz- und Sanger-Sequenzierung detektierten Mutationen ermittelt. Alle Proben mit bereits bekannten Mutationen aus der Sanger-Sequenzierung bei denen keine weiteren Mutationen in der Hochdurchsatz-Sequenzierung identifiziert wurden, sind richtig negativ.

Target-Region:

Eine spezifische zu analysierende genomische Region (Zielsequenz).

Total-Mapped-Reads:

Die Anzahl an Sequenzen einer Probe, die beim Vergleich der Konsensussequenz gegen das Referenzgenom eindeutig zugeordnet werden.

Total-Reads:

Die Gesamtanzahl an Sequenzen einer Probe die mittels Hochdurchsatz-Sequenzierung generiert wurden.

VCF-File/Datei (engl. Variant Call File):

Eine textbasierte Datei, in der alle Abweichungen von dem Referenzgenom durch die Angabe von dem Chromosom und der genomischer Position aufgelistet sind.

Abkürzungen

| | |
|------------------------|-------------------------------------------------------------------------------------------------------|
| °C | Grad Celsius |
| 450K | Infinium HumanMethylation450 BeadChip |
| 5hmC | 5-Hydroxymethylcytosin |
| 5mC | 5-Methylcytosin |
| A | Adenin |
| AG | Arbeitsgruppe |
| ATP | Adenosintriphosphat |
| AVG-Beta | engl. average Beta, DNA-Methylierungswert eines CpG-Lokus |
| <i>B2M</i> | engl. Beta-2-Microglobulin |
| BA | Bioanalyzer |
| <i>BCL2</i> | engl. B-Cell Lymphoma 2 |
| <i>BCL6</i> | engl. B-Cell Lymphoma 6 |
| BCR | B-Zell-Rezeptor |
| BL | Burkitt-Lymphom |
| bp | Basenpaar |
| bzw. | Beziehungsweise |
| C | Cytosin |
| ca. | Circa |
| CD | engl. Cluster of differentiation, Immunphänotypische Oberflächenmarker von Zellen |
| CD10 | engl. Cluster of differentiation 10 (Antigen) |
| <i>CD58</i> | engl. Lymphocyte Function-Associated Antigen 3 |
| <i>CD79B</i> | engl. Immunoglobulin-Associated Beat |
| cg | Infinium HumanMethylation450 BeadChip Bezeichnung einer Sonde die einen CpG-Lokus repräsentiert |
| ch | Infinium HumanMethylation450 BeadChip Bezeichnung einer Sonde die einen nicht-CpG-Lokus repräsentiert |
| CHOP | Chemotherapie mit den Wirkstoffen Cyclophosphamid, Doxorubicin, Vincristin und Prednison |
| cm | Zentimeter |
| CpG | Cytosin-Guanin-Dinukleotid |
| <i>CREBBP</i> | <i>CREB Binding Protein</i> |
| CSV-Datei | Dateiformat, Comma-separated values |
| ddDNA | Doppelsträngige DNA |
| Dest. H ₂ O | Destilliertes Wasser |
| DetPval | engl. detection P-value, detektierter P-Wert |
| DKFZ | Deutsches Krebsforschungszentrum |
| DLBCL | Diffus großzelliges B-Zell-Lymphom |
| DLBCL-ABC | Diffus großzelliges B-Zell-Lymphom vom aktivierten B-Zelltyp |
| DLBCL-GCB | Diffus großzelliges B-Zell-Lymphom vom Keimzentrumsähnlichen B-Zelltyp |
| DMR | Differenziell methylierte Region |
| DNA | Desoxyribonukleinsäure |
| DNMT | DNA-Methyltransferase |
| dNTP | Desoxynukleotidtriphosphat |

Abkürzungen

| | |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DSMZ | Deutsche Sammlung für Mikroorganismen und Zellkulturen |
| EDTA | Ethylendiamintetraessigsäure |
| FACS | engl. Fluorescence-activated cell sorting, Sortieren der Zellen in einer Durchflusszytometrie mittels Fluoreszenzfarbstoffen |
| FDR | engl. false discovery rate, Falsch-Positiv-Rate |
| FFPE | engl. Formalin-Fixed Paraffin-Embedded, Formalin-fixiertes Material |
| FISH | Fluoreszenz-in-situ-Hybridisierung |
| FL | Follikuläres Lymphom |
| g | Gramm |
| G | Guanin |
| Gb | Gigabasen |
| GC | engl. Germinal center, Keimzentrum |
| gDNA | genomische DNA |
| GEP | Genexpression |
| GEP-Klassifikator | Klassifikation der DLBCL-Subtypen in GCB und ABC nach Genexpressionsdaten von Mikroarrays oder RNA-Seq |
| ggf | Gegebenfalls |
| GS | GenomeStudio |
| GYN | Gynäkologie |
| h | Stunden |
| Hans-Klassifikator | Klassifikation der DLBCL-Subtypen in GCB und nonGCB nach Hans et al. |
| HG | Humangenetik |
| HGVS | engl. Human Genome Variation Society, Gesellschaft für die Nomenklatur des humanen Genomes |
| HS | engl. High sensitivity, Hoch sensitive |
| ICGC | engl. International Cancer Genome Consortium, Internationales Krebsgenomkonsortium |
| <i>ID3</i> | engl. Inhibitor of DNA Binding 3, Dominant Negative Helix-Loop-Helix Protein |
| IDAT | engl. intensity data, Intensitätsdaten |
| <i>IGH</i> | Immunoglobulin Heavy chain Locus |
| IHC | Immunhistochemie |
| Indel | Insertion oder Deletion |
| IPI | Internationaler prognostischer Index |
| <i>IRF4</i> | Interferon Regulatory Factor 4 |
| J. | Jahre |
| Jack | „Jackstädt-Studie“- Projekt zu : “Mechanismen der B-Zell-Lymphomentstehung im Senium als Grundlage für die Entwicklung altersangepasster Therapiestrategien“ gefördert durch die Dr. Werner Jackstädt Stiftung |
| K | Kluster |
| k/mm ² | Kluster pro Quadratmillimeter |
| Kb | Kilobasen |
| kg | Kilogramm |
| KR | Kontrollregion im Intron des <i>ID3</i> -Gens |
| LDR | engl. Loci detection rate, Loci-Detektionsrate |
| LK | Lymphknoten |
| m | Männlich |
| max. | Maximal |

Abkürzungen

| | |
|--------------|--------------------------------------------------------------------------------------------------------------------------|
| Mb | Megabasen |
| mBL | Molekulares Burkitt-Lymphom |
| MF | Mutationsfrequenz |
| mg | Milligramm |
| min | Minute |
| min. | Minimal |
| ml | Milliliter |
| mM | Millimolar |
| mmHG | Torr – Einheit der physikalischen Größe „Druck“ |
| MMML | Verbundprojekt “Molekulare Mechanismen bei malignen Lymphome” |
| <i>MTOR</i> | engl. Mechanistic Target of Rapamycin (Serine/Threonine Kinase) |
| MUM1 | Melanoma Associated Antigen (Mutated) 1(Antigen) |
| mut | Mutiertes Individuum |
| <i>MYC</i> | engl. V-Myc Avian Myelocytomatosis Viral Oncogene Homolog |
| <i>MYD88</i> | engl. Myeloid Differentiation Primary Response 88 |
| N | Anzahl |
| NA | Nicht verfügbar |
| NaCl | Natriumchlorid |
| NaOH | Natriumhydroxid |
| ng | Nanogramm |
| NGS | engl. Next Generation Sequencing, Hochdurchsatz-Sequenzierung |
| NHL-B | Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphomen bis zu einem Ersterkrankungsalter von 75 Jahren |
| NK | Normalkontrolle |
| nMZL | Nodales Marginalzonen-Lymphom |
| non-mBL | Nicht-molekulares Burkitt-Lymphom |
| NOS | engl. not other specified, Nicht weiter spezifiziert |
| Ø | Durchschnittlich |
| OTC | engl. On-Target-Coverage |
| OTR | engl. On-Target-Rate |
| P | P-Wert, Signifikanzwert |
| PCA | engl. principal component analysis, Hauptkomponentenanalyse |
| PCR | Polymerasekettenreaktion |
| pmol | Picomolar |
| QC | Qualitätskontrolle |
| qPCR | quantitative Realtime PCR |
| R | Rituximab |
| RICOVER | Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom mit einem Ersterkrankungsalter von 61 bis 80 Jahren |
| Rkt | Reaktion |
| RNA | Ribonukleinsäure |
| <i>RRAS</i> | engl. Related RAS Viral (R-Ras) Oncogene Homolog |
| RT | Raumtemperatur |
| sec | Sekunde |
| Seq. | Sequenzierung |
| SNP | engl. single nucleotide polymorphism, Einzelnukleotidpolymorphismus |
| SNV | engl. single nucleotide variation, Punktmutation |
| sog. | sogenannt |

Abkürzungen

| | |
|-------------|---------------------------------------------------------------------------------------|
| SV | Strukturelle Varianten |
| SYBR-Green | Cyanin-Farbstoff |
| T | Thymin |
| t | Translokation |
| T-PLL | T-Prolymphozyten-Leukämie |
| TBE | Tris-Borat-EDTA |
| <i>TET1</i> | engl. Ten-Eleven Translocation-1 bzw. Tet Methylcytosine Dioxygenase 1 |
| <i>TET2</i> | engl. Ten-Eleven Translocation-2 bzw. Tet Methylcytosine Dioxygenase 2 |
| TMA | engl. Tissue Micro Array, Kleine Gewebestanden von Tumor-Proben auf Glasobjektträgern |
| TRIS | Tris(hydroxymethyl-)aminomethan |
| TSS | Transkriptionsstartstelle |
| U | Uracil |
| u.a. | Unter anderem |
| u.U. | Unter Umständen |
| UKSH | Universitätsklinikum Schleswig-Holstein |
| V | Volt |
| vs. | (engl. versus) gegen |
| w | weiblich |
| WGS | engl. Whole Genome Sequencing, Ganz-Genom-Sequenzierung |
| WHO | engl. World Health Organization, Weltgesundheitsorganisation |
| wt | Wildtyp (unmutiert) |
| x | -faches |
| xg | -faches der Erdbeschleunigung |
| z.B. | Zum Beispiel |
| z.T. | Zum Teil |
| µl | Mikroliter |

Anmerkung: Die Gennamen wurden nicht in die deutsche Sprache übersetzt sondern sind nur in Englischer Sprache angegeben und stammen von der Internetseite GeneCards (<http://www.genecards.org/>, Stand November 2015).

1 Einleitung

1.1 Immunsystem und Alternsprozess

Das Immunsystem ist der Abwehrmechanismus des Körpers und schützt den Organismus vor körpereigenen und körperfremden, potentiell gefährlichen Pathogenen. Menschen mit einer Immunschwäche besitzen eine hohe Anfälligkeit gegenüber opportunistischen Keimen, Tumorgenese sowie Autoimmunkrankheiten^{1,2}. Ein gesunder Organismus schützt den Körper täglich vor derartigen Krankheiten, indem einzelne bösartige Zellen durch Effektorzellen des Immunsystems eliminiert werden². Das Immunsystem ist ein komplexes Netzwerk im Körper, das sich aus dem nativen und dem adaptiven Immunsystem zusammensetzt und aus frei beweglichen Immunzellen, löslichen Serumbestandteilen und lymphatischen Organen besteht^{2,3}. Die zellulären Bestandteile finden ihren Ursprung in der Hämatopoese im Knochenmark. Aus pluripotenten Stammzellen werden myeloide oder lymphoide Vorläuferzellen gebildet. Die myeloiden Vorläuferzellen differenzieren u.a. zu Thrombozyten, Erythrozyten, Granulozyten sowie Monozyten. Aus den lymphoiden Vorläuferzellen gehen natürliche Killerzellen (NK), B-Zellen, T-Zellen sowie dendritische Zellen hervor⁴⁻⁶.

Das Immunsystem unterliegt ähnlich wie alle Organe dem natürlichen Prozess des Alterns, was häufig zu einer geringeren funktionellen Leistungsfähigkeit und Stressresistenz des Organismus führt. Zusätzlich ist der Alternsprozess mit einem erhöhten Risiko für Erkrankungen wie z.B. Krebs assoziiert^{7,8}. Jedoch korreliert das chronologische Alter nicht immer mit dem biologischen Alter des Organismus. So gibt es beispielweise junge Menschen, deren Immunsystem bereits Symptome des Alterns aufweist und andererseits gibt es Senioren, die bis ins hohe Alter ein sehr abwehrstarkes (jungendliches) Immunsystem aufweisen^{9,10}. Dabei spielen viele Mechanismen eine zentrale Rolle u.a. die Telomerverkürzung, die Akkumulation von genetischen Aberrationen, oxidativer Stress oder auch der molekulare Stoffwechsel. Aus verschiedenen Studien ist bereits bekannt, dass Umweltfaktoren und der Lebensstil (z.B. Rauchen und körperliche Aktivitäten) den Prozess des Alterns und entsprechend das Immunsystem des Organismus beeinflussen können¹¹⁻¹³. Zusätzlich können genetische und epigenetische Modifikationen während dem Altern die zelluläre Funktion beeinflussen und eine Auswirkung auf die Entwicklung altersassoziierter Phänotypen und Krankheiten wie z.B. neurologische Störungen, Diabetes oder Tumore haben¹⁴⁻¹⁸.

Es besteht die Hypothese, dass epigenetische Mechanismen während des Alternsprozesses möglicherweise eine essenzielle Rolle in der kontrollierten Veränderung der Genexpression und der genomischen Instabilität spielen¹². Im Normalfall nimmt mit dem Alter die globale DNA-Methylierung ab, während die zell- und gewebespezifische lokale Promotor-Methylierung zunimmt. In verschiedenen Arbeitsgruppen wurden in den letzten Jahren altersassoziierte CpG-Loci identifiziert, die mit zunehmendem Alter eine höhere oder geringere Methylierung aufwiesen. Diese Beobachtungen bestätigten die Hypothese, dass es einen Zusammenhang zwischen dem chronologischen Alter und dem DNA-Methylierungsmuster des menschlichen Organismus gibt^{13,17,19-24}. Für einzelne altersassoziierte CpG-Loci wurde eine nahezu lineare Änderung im Methylierungsgrad während des Alternsprozesses identifiziert. Diese CpG-Loci wurden anschließend als biologische Marker in verschiedenen Modellen zur Altersvorhersage verwendet^{12,22,24-27}. Das derzeit wohl bekannteste Modell zur Altersvorhersage, das auch in der vorliegenden Arbeit zur Bestimmung des epigenetischen Alters von Tumor- und Normalkontrollproben verwendet wurde, ist das Modell nach Horvath²². Es wurde von Steve Horvath von der Universität in Kalifornien (Los Angeles) entwickelt und nutzt einen Algorithmus basierend auf 353 CpG-Loci, die auf dem Infinium HumanMethylation450 BeadChip (Illumina) lokalisiert sind, zur Bestimmung des epigenetischen Alters der analysierten Proben. Mit diesem Modell konnte das epigenetische Alter für fast alle Gewebe des Körpers bestimmt werden, wobei das ermittelte epigenetische Alter mit dem biologischen Alter der analysierten Proben korrelierte. Die Auswertungen des biologischen und epigenetischen Alters nach Horvath zeigten keine zellspezifischen Veränderungen, sondern ausschließlich altersassoziierte Veränderungen im DNA-Methylom der analysierten Proben²². Weiterhin konnten verschiedene Arbeitsgruppen zeigen, dass in Tumoren (z.B. Lymphomen) vorzugsweise Gene aberrant methyliert sind, die auch mit dem Altern des Individuums eine zunehmende Methylierung zeigen^{16,28,29}. Es besteht daher die Vermutung, dass es einen Zusammenhang zwischen dem DNA-Methylierungsprofil von Älteren und dem DNA-Methylierungsmuster von Tumorpatienten gibt. Auch bei gesunden Personen konnte kürzlich gezeigt werden, dass die Anzahl genomischer Aberrationen im Mosaik im peripheren Blut mit dem Alter zunimmt. Dabei sind typischerweise solche Regionen betroffen, die gleichermaßen bei hämatologischen Neoplasien alteriert sind und mit Leukämien und Lymphomen assoziiert werden^{30,31}. Die Immunabwehr ist von zentraler Bedeutung in der Pathogenese von B- und T-Zellen des lymphatischen Systems und wird immer wichtiger in der modernen Krebstherapie zur effektiven Behandlung von bisher unheilbaren Erkrankungen. So werden u.a. Mutationen und Rekombinationen in den Genen für die B- und T-Zell-Rezeptoren eingefügt, die zu einer effektiven Immunantwort des Organismus führen und nachhaltig die Tumorzellen zerstören bzw. die Entartung von Zellen des lymphatischen Systems zu Krebszellen unterbinden³²⁻³⁵. Detaillierte Informationen zum Aufbau und der Funktion des lymphatischen Systems im menschlichen Organismus sind in dem nachfolgenden Abschnitt zusammengefasst.

1.2 Lymphatisches System

Das lymphatische System setzt sich aus zirkulierenden Lymphozyten und dem lymphatischen Gewebe zusammen und hat die Aufgabe, den Organismus vor schädigenden Fremdkörpern zu schützen. Es erstreckt sich über den gesamten Körper und kann in ein primär lymphatisches (Knochenmark, Thymus) und sekundär lymphatisches System (Milz, Lymphknoten und mukosa-assoziierte lymphatische Gewebe) eingeteilt werden^{2,3,36}. Die lymphatischen Organe sind über Lymphgefäße miteinander verbunden wobei die Lymphe aus den peripheren in die sekundären Organe geleitet wird. Die Lymphe ist ein Ultrafiltrat des Blutplasmas mit Plasmaproteinen und dient der Zell- und Gewebeversorgung sowie dem Transport der Lymphozyten aus dem Gewebe zurück in den Blutkreislauf³. Ein Teil der zellulären Stoffwechselprodukte sammelt sich als Lymphe in den Lymphkapillaren der lymphatischen Organe und wird entlang der Lymphbahnen durch den Lymphknoten geleitet, und schließlich wieder dem Blutkreislauf zugeführt. Die zirkulierenden Lymphozyten sind weiße Blutkörperchen und von zentraler Bedeutung bei der Immunantwort. Sie entstehen aus pluripotenten Stammzellen im Knochenmark und lassen sich in T-, B- und natürliche Killer-Zellen unterteilen. Die B-Zellen reifen im Knochenmark weiter zu antikörperproduzierenden Plasmazellen und sind für die humorale Immunantwort des erworbenen Immunsystems verantwortlich. Die T-Zellen hingegen sind bei der zellulären Immunantwort von zentraler Bedeutung und werden im Thymus zu Gedächtnis-, zytotoxischen-, und T-Helfer-Zellen differenziert. Neben den Lymphozyten werden auch Pyrogene oder Pathogene mit Hilfe dieser Flüssigkeit in den Lymphknoten transportiert, der als Startpunkt für die anschließende Effektorreaktion und den Abwehrmechanismus fungiert³⁷. Die wichtigste Funktion der B-Zellen ist die Sezernierung von Antikörpern. Dazu wandern aktivierte B-Zellen aus dem lymphatischen System und dem Blutkreislauf in den nächstgelegenen Lymphknoten, worin die B-Zellen in der Keimzentrums-Reaktion bis hin zu langlebigen Antikörper sezernierenden Plasmazellen differenzieren^{2,3}. Eine detaillierte Beschreibung der B-Zell-Differenzierung und Keimzentrums-Reaktion ist in dem nachfolgenden Abschnitt nachzulesen.

1.2.1 Keimzentrums-B-Zell-Entwicklung

Der Bildungsort für die reifen B-Zellen ist das Knochenmark (engl. bone marrow). Im Vorläuferstadium der B-Zell-Entwicklung entsteht aus einer hämatopoetischen Stammzelle eine Pro-B-Zelle². Im frühen Pro-B-Zell-Stadium sind bereits einige B-Zell-spezifische Oberflächenproteine vorhanden und der Prozess der Immunglobulin (Ig)-Genumlagerung der schweren Ketten (H, heavy chain) beginnt. Bei der IgH-Genumlagerung werden aus den zahlreichen Gensegmenten der Keimbahnkonfiguration sog. V- (engl. Variability), D- (engl. Diversity) und J- (engl. Joining) Segmente zufällig ausgewählt und zu einem neuen Komplex zusammengefügt (V-D-J). Dieser Umlagerungsprozess wird auch als somatische (VDJ-)Rekombination bezeichnet. In der nachfolgenden Prä-B-Zelle werden die Ig-Leichtketten (λ und κ) umgelagert. Prinzipiell ist dies dem Prozess der IgH-Genumlagerung identisch. Der einzige Unterschied ist, dass aufgrund des Fehlens von D-Segmenten nur V- und J-Segmente verknüpft werden. Die Leichtketten-Genumlagerung ist somit nach nur einem Schritt beendet und es entsteht eine unreife B-Zelle. Die unreife B-Zelle exprimiert komplette IgM/D Antikörper, die auf der Zelloberfläche verankert sind und mit den Proteinen Ig α und Ig β zum B-Zell-Rezeptor (BCR) assoziieren³.

Die unreife B-Zelle wandert nun aus dem Knochenmark über die Blutbahn in das periphere lymphatische Gewebe und liegt dort als zirkulierende reife naive B-Zelle in Form von Primärfollikeln vor. In der sich anschließenden Keimzentrums-Reaktion transformieren die eingewanderten B-Zellen in Zentroblasten und breiten sich durch Proliferation stark aus. Die Folge ist, dass die ruhenden B-Zellen der Primärfollikel an den Rand gedrängt werden und sich der Follikelmantel ausbildet. In der weiteren Entwicklung nimmt die Proliferationsrate der Zentroblasten ab und sie differenzieren zu Zentrozyten, die sich unterhalb des Follikelmantels sammeln. Dadurch bilden sich zwei verschiedene Zonen im Keimzentrum (GC) aus: eine zentrozytenreiche (helle Zone) und eine zentroblastenreiche (dunkle Zone) Keimzentrums-Zone. Sowohl die Zentroblasten als auch die Zentrozyten durchlaufen nun den Prozess der somatischen Hypermutation (SHM), wobei es zu genetischen Aberrationen (Mutationen) innerhalb der variablen Bereiche der Ig-Gene kommt. B-Zellen, die durch eine veränderte Ig-Struktur von keiner höheren Antikörperaffinität profitieren, gelangen anschließend in den Prozess der Apoptose. Zentrozyten mit einer höheren Antigenaffinität differenzieren zu langlebigen Plasmazellen oder Gedächtnis-B-Zellen^{3,38-41}. Eine schematische Darstellung der B-Zell-Entwicklung ist in der nachfolgenden Abbildung 1A dargestellt.

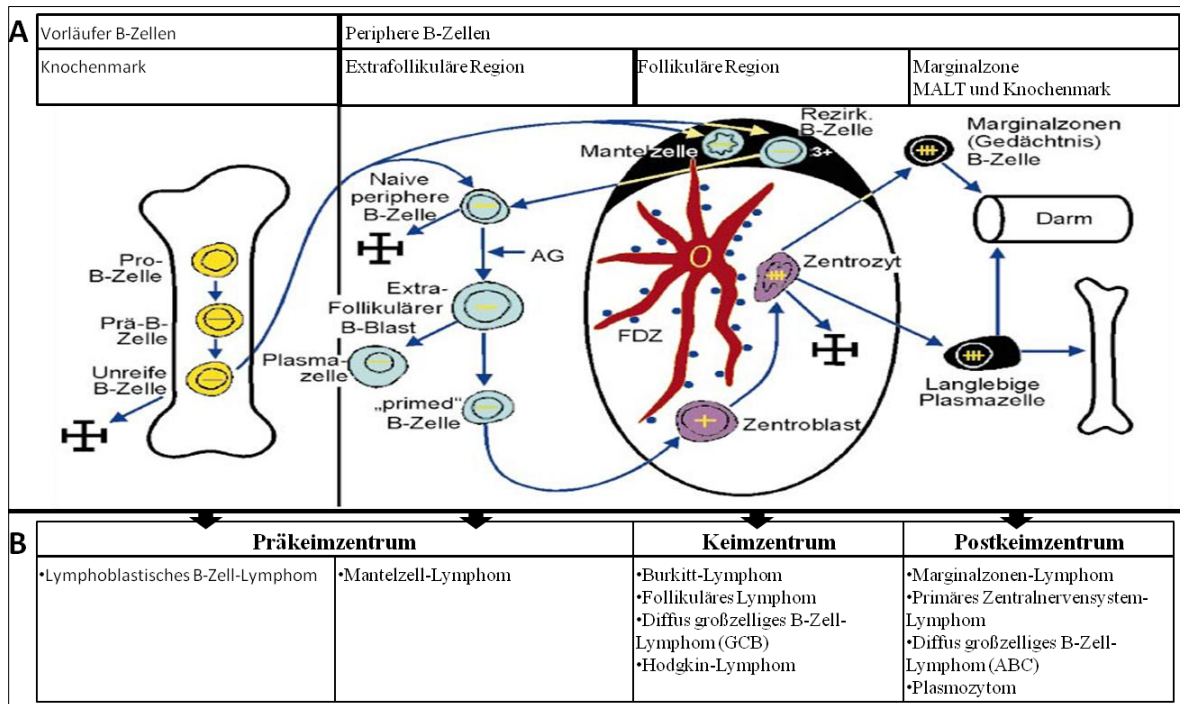


Abbildung 1: Schematische Darstellung der B-Zell-(Lymphom)-Entwicklung und Differenzierung

A: Die Grafik stellt eine Übersicht der B-Zell-Entwicklung im menschlichen Organismus von der Pro-B-Zelle im Knochenmark bis zur differenzierten Gedächtnis- oder Plasma-Zelle im peripheren lymphatischen Gewebe dar. Die Umlagerung der Immunglobulin-Gene ist durch Querbalken in der Abbildung dargestellt. Somatische Mutationen sind durch senkrechte Striche in den Querbalken symbolisiert. AG: Fremdantigen, FDC: Follikuläre dendritische Zelle, MALT: Mukosaassoziiertes lymphatisches Gewebe, GCB: Keimzentrumsähnliche B-Zellen, ABC: Aktivierte B-Zellen (modifiziert nach Stein et al., 2007⁴¹). B: In der Grafik sind einzelne Zeitpunkte und Differenzierungsstadien gekennzeichnet an denen normale B-Zellen innerhalb der B-Zell-Entwicklung eine maligne Transformation durchlaufen können und infolgedessen Lymphome entstehen. Beispielsweise gehen folliculäre Lymphome und diffus großzellige B-Zell-Lymphome vom GCB-Subtyp aus B-Zellen im Keimzentrum hervor, wohingegen der Ursprung der Tumorzellen der diffus großzelligen B-Zell-Lymphome vom ABC-Subtyp vermutlich im Postkeimzentrum liegt.

Die Entwicklung der B-Zellen unterliegt einem komplexen Netzwerk, worin verschiedene Transkriptionsfaktoren und Regulatoren miteinander interagieren und zusammen die Prozesse von der Initiation (Naive B-Zelle) bis zur terminalen Differenzierung (Gedächtnis- oder Plasma-Zelle) steuern⁴². Die nachfolgende Abbildung 2 zeigt eine schematische Übersicht des transkriptionellen Netzwerks der Keimzentrums-B-Zell-Reaktion inkl. einzelner Transkriptionsfaktoren, von denen bereits bekannt ist, dass diese eine essentielle Rolle in der Lymphogenese von Keimzentrums-B-Zell-Lymphomen besitzen.

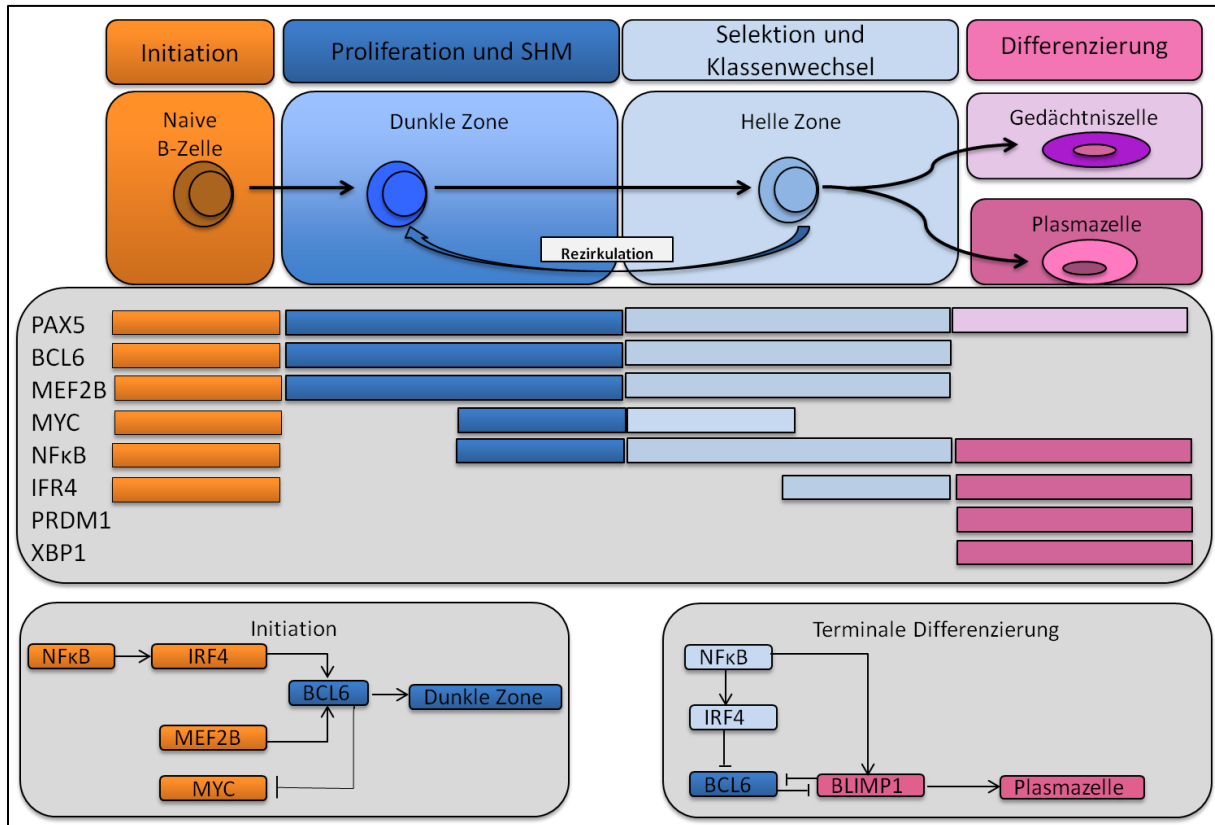


Abbildung 2: Schematische Darstellung des transkriptionellen Netzwerkes in der Keimzentrums-B-Zell-Reaktion

Verschiedene Transkriptionsfaktoren und Regulatoren sind an der Regulation der Keimzentrums-Reaktion beteiligt. Das Mastergen der B-Zell-Entwicklung ist *PAX5*, wobei das entsprechende Protein in fast jedem Differenzierungsstadium exprimiert wird und erst zur Plasmazell-Differenzierung herunterreguliert ist. Hingegen werden *BCL6* und *MEF2B* in der terminalen Differenzierung nicht exprimiert. *MYC* wird während der Initiation und von einem Subset an B-Zellen in der hellen Zone exprimiert. *NF-κB* und *IRF4* liegen in der Initiation der Keimzentrums-Reaktion transkriptionell aktiv vor, werden in der dunklen Zone nicht exprimiert, wohingegen es zur Re-Expression der Gene in der hellen Zone kommt. *PRDM1* und *XBP1* sind ausschließlich während der terminalen Differenzierung zur Plasmazelle transkriptionell aktiv. Die verschiedenen Farben in der Grafik kennzeichnen die unterschiedlichen Signalwege und Auswirkungen (modifiziert nach Basso et al., 2015⁴²).

MYC ist von zentraler Bedeutung in der Induktion und Amplifikation von transkriptionellen Mechanismen die mit der Stimulation des Metabolismus, der Telomerasefunktion und der DNA-Reparatur assoziiert sind⁴³. Zudem ist *MYC* ein essentieller Faktor in der frühen Keimzentrums-Reaktion und charakterisiert durch ein bimodales Expressionsmuster, wobei die Expression zum späteren Zeitpunkt der Keimzentrums-Reaktion durch den transkriptionellen Repressor *BCL6* inhibiert wird⁴⁴⁻⁴⁹.

BCL6 ist ein transkriptioneller Repressor, der u.a. ein Faktor für die Regulation der Mechanismen der Entstehung und Aufrechterhaltung der Keimzentrums-Reaktion innerhalb der B-Zell-Entwicklung darstellt⁵⁰. Insgesamt ist festzuhalten, dass *BCL6* als Masterregulator der Keimzentrums-Reaktion fungiert und verschiedene Target-Gene kontrolliert, die mit dem transkriptionellen Netzwerk der Keimzentrums-Reaktion assoziiert sind. *BCL6* ist dabei von essentieller Bedeutung u.a. in der Inhibierung des Zellzyklus, der Regulierung der Apoptose und der Anfälligkeit von Keimzentrums-B-Zellen für DNA-Schäden⁵¹⁻⁵⁸.

In der interfollikulären Zone der Keimzentrums-Reaktion exprimieren bereits einzelne naive B-Zellen, die durch den Kontakt mit Antigenen oder T-Zellen aktiviert wurden, das Protein BCL6^{59,60}. Weiterhin wird die Expression von BCL6 durch die Transkriptionsfaktoren MEF2B⁶¹ und IRF4⁵⁷ beeinflusst.

Der Transkriptionsfaktor MEF2B wird zum frühen Zeitpunkt der Keimzentrums-Reaktion (Initiation) von B-Zellen exprimiert⁶¹. Vergangene Studien zeigen, dass MEF2B das *BCL6*-Gen direkt transaktivieren kann und genetische Aberrationen in MEF2B zu einem veränderten Expressionsmuster von BCL6 führen^{61,62}. IRF4 ist ein Transkriptionsfaktor, der zur Familie der Interferon Regulator Faktoren gehört und von entscheidender Bedeutung bei der Regulation des Immunsystems und der Immunantwort im menschlichen Organismus ist. IRF4 wird zum frühen Zeitpunkt der Keimzentrums-Reaktion induziert und bewirkt anschließend die Expression von BCL6. In der dunklen Zone ist IRF4 transkriptionell inaktiv, wohingegen es in der nachfolgenden hellen Zone bei einzelnen B-Zellen, die weiter zu Plasmazellen differenzieren, zur Re-Expression von IRF4 kommt^{57,63,64}.

NF-κB ist ein Transkriptionsfaktor, der aus einem Homo- oder Heterodimer von Proteinen der NF-κB-Familie besteht und für die Entwicklung und Funktion von Immunzellen u.a. B-Zellen von entscheidender Bedeutung ist⁶⁵. Im inaktiven Zustand ist NF-κB durch einen Inhibitor (IκB) gebunden und liegt im Zytosol vor. In Folge der Degradierung des IκB wird eine Signalkaskade gestartet, wodurch es zur Aktivierung von NF-κB und dem zugehörigen Signalweg kommt^{65,66}. Die Aktivierung von NF-κB wird posttranskriptionell reguliert, d.h. das Protein liegt bereits in jeder normalen Zelle vor und muss zur Aktivierung der NF-κB-Signalkaskade nicht neu synthetisiert werden. NF-κB wird zum frühen Zeitpunkt der Keimzentrums-Reaktion (Initiation) exprimiert, wohingegen in der dunklen Zone der Keimzentrums-Reaktion NF-κB transkriptionell inaktiv ist^{47,67,68}.

PAX5 gilt als Masterregulator der B-Zell-Entwicklung und Leukämogenese im menschlichen Organismus⁵⁵ und spielt eine essentielle Rolle in der B-Zell-Differenzierung. In allen normalen hämatopoetischen Zellen ist PAX5 vorhanden und wird zudem in fast allen B-Zell-Differenzierungsstadien exprimiert. Besonders ist jedoch, dass die Expression von PAX5 im Rahmen der terminalen Differenzierung zur Plasmazelle herunterreguliert wird⁶⁹⁻⁷².

PRDM1 (BLIMP1) wird als Mastergen der terminalen B-Zell-Entwicklung bezeichnet, da die Expression des Proteins von zentraler Bedeutung bei der Differenzierung von reifen B-Zellen zu Ig-sezierenden Plasmazellen ist. PRDM1 steuert u.a. die Mobilität und das Wachstum der Plasmazellen in das Knochenmark und sorgt für die Hochregulierung der Antikörperproduktion. Insgesamt ist festzuhalten, dass der menschliche Organismus ohne PRDM1 keine Antikörper-vermittelte Immunantwort ausüben kann^{73,74}.

XBP1 ist ein Transkriptionsfaktor, der als Co-Faktor zusammen mit PRDM1 und IRF4 eine wichtige Bedeutung in der B-Zell-Differenzierung von reifen B-Zellen zu Plasmazellen besitzt⁷⁵. Dabei ist das Protein XBP1 insbesondere für die Sekretion der Antikörper von Plasmazellen verantwortlich^{76,77}.

1.3 Keimzentrums-B-Zell-Lymphome

Genetische oder epigenetische Aberrationen in den Genen des transkriptionellen Netzwerkes können zur Entartung der B-Zellen innerhalb der Keimzentrums-Reaktion führen⁷⁸⁻⁸³. B-Zellen können zu jedem Zeitpunkt während der Entwicklung entarten, woraufhin die Regulation und/oder Aktivierung der normalen B-Zell-Differenzierung gestört wird. Die Folge ist das Wachstum von Tumoren durch eine unkontrollierte Proliferation der entarteten B-Zellen (Tumorzellen). Tumore des lymphatischen Systems werden als Lymphome bezeichnet, wobei je nach Reifungs- und Differenzierungsstadium der entarteten B-Zelle verschiedene B-Zell-Lymphome entstehen. Alle Lymphome bei denen die maligne Transformation der B-Zellen innerhalb der Keimzentrums-Reaktion stattfindet, werden als Keimzentrums-B-Zell-Lymphome bezeichnet⁸⁴⁻⁸⁸. In der Abbildung 1 ist die B-Zell-Entwicklung (A) und die Zuordnung einzelner B-Zell-Lymphome zu bestimmten Differenzierungsstadien (B) schematisch dargestellt. Eine Gruppe von malignen Keimzentrums-Lymphomen, die aus B-, T- oder natürlichen Killerzellen hervorgehen, werden als Non-Hogkin-Lymphome (NHL) bezeichnet. Die zwei häufigsten Formen der NHL sind die folliculären Lymphome (FL) und diffus großzelligen Lymphome (DLBCL), die zusammen jedes Jahre ca. 60% der neu diagnostizierten NHL der B-Zellreihe ausmachen⁶².

1.3.1 Follikuläre Lymphome

Follikuläre Lymphome (FL) sind die häufigsten indolenten (niedermalignen) Lymphome und machen etwa 20-30% aller B-Zell-Lymphome in unserer Bevölkerung aus^{89,90}. Die Inzidenz dieser Lymphom-Entität liegt in den westlichen Ländern bei 5-7/100.000 Einwohner, wobei die Anzahl stetig zunimmt. Das FL ist eine Erkrankung des fortgeschrittenen Alters mit einem medianen Erkrankungsalter zwischen 55 und 60 Jahren⁹¹. Der klinische Verlauf der Erkrankung ist gekennzeichnet durch ein langsames Tumorwachstum mit häufigen Rezidiven, wobei die Überlebensrate für Patienten bei 15-20 Jahren nach der Diagnose liegt. Aufgrund des langsamen Wachstums eines indolenten Lymphoms wird die Diagnose eines FL häufig erst im fortgeschrittenen Stadium von III gestellt. Die Graduierung der Lymphome erfolgt nach der Anzahl der Zentroblasten im High-Power-Field (hpf). Die WHO-Klassifikation unterscheidet dabei wie folgt: Grad 1-2 (0-15 Zentroblasten pro hpf) und Grad 3 (>15 Zentroblasten pro hpf). FL vom Grad 3 werden zusätzlich in A und B unterteilt, wobei FL vom Grad 3B nicht mehr als indolente sondern hochmaligne Lymphome mit einem aggressiven klinischen Verlauf angesehen werden. Patienten mit einem FL Grad 3B werden mit dem gleichen Protokoll wie Patienten mit einem DLBCL behandelt. Etwa ein Drittel aller FL transformiert nach einigen Jahren zu einem aggressiven Lymphom (z.B. DLBCL)⁹².

FL sind kleinzellige B-Zell-Lymphome, die innerhalb der Keimzentrums-Reaktion entstehen. Während der normalen Keimzentrums-Reaktion der B-Zell-Entwicklung werden die Follikelmäntel gebildet (Abbildung 1A: Follikuläre Region), was im Verlauf der Entstehung von FL ausbleibt^{77,93}. Beim FL bilden die entarteten Keimzentrums-B-Zellen innerhalb der B-Zell-Entwicklung sog. Pseudofollikel aus, die nicht wie gewohnt an den Rand wandern und den Follikelmantel darstellen, sondern im GC bestehen bleiben. Diese Umstrukturierung der Follikel führt zur Aufhebung der physiologischen Lymphknotenstruktur^{89,94}. Nach der WHO-Klassifikation von 2008 werden die FL in vier unterschiedliche Wachstumsmuster eingeteilt: erstens.) rein follikulär, zweitens.) follikulär und diffus, drittens.) fokal-follikulär bei weniger als 25% follikulärem Muster und viertens.) rein diffus^{89,95}. Immunhistologisch sind die Tumorzellen der FL häufig durch einen charakteristischen Phänotyp gekennzeichnet und weisen spezifische Keimzentrums-assoziierten B-Zell-Markern auf: CD10⁺, CD19⁺, CD20⁺, CD22⁺, CD45RA⁺, CD79A⁺, BCL2⁺, BCL6⁺, CD5⁻, CD23⁻, CD43⁻^{89,93}. Die Tumorzellen des FL sind B-Zellen, wobei es innerhalb der Keimzentrums-Reaktion im Prozess der somatischen Hypermutation zu Aberrationen in den Ig-V-Genen kam. Diese aktive somatische Hypermutation resultiert daraufhin in eine intraklonale V-Gen-Diversität in den Tumorzellen der FL⁹⁶.

1.3.2 Diffus großzellige B-Zell-Lymphome

Diffus großzellige B-Zell-Lymphome (DLBCL) sind eine sehr heterogene Entität hinsichtlich ihrer Morphologie, Biologie und molekularen Mechanismen ihrer Pathogenese^{97,98}. DLBCL sind im hohen Alter die häufigsten malignen Lymphome in unserer Bevölkerung und umfassen etwa 30 bis 40% aller Non-Hodgkin-Lymphome (NHL)⁹⁷. In den letzten Jahren nahm die Anzahl an DLBCL-Patienten aus bisher ungeklärten Gründen stetig zu, wobei das mediane Ersterkrankungsalter für Patienten mit einem DLBCL bei 64 Jahren lag^{91,99}. DLBCL sind durch einen hoch aggressiven klinischen Verlauf gekennzeichnet und führen unbehandelt schnell zum Tod^{89,90}. Bei einer Diagnose im frühen Erkrankungsstadium und einer anschließenden Therapie liegt die Heilungsrate der DLBCL-Patienten bei 60 bis 70%. Bei den restlichen 30 bis 40% der Fälle liegt eine refraktäre Erkrankung vor oder es kommt im Laufe der Zeit zu einem Rezidiv, was mit einer hohen Morbidität und Mortalität assoziiert ist¹⁰⁰⁻¹⁰³. Trotz der Heterogenität weisen DLBCL Gemeinsamkeiten auf. So ist beispielsweise der Zellkern aller DLBCL etwa doppelt so groß oder größer wie bei normalen B-Zellen. Die Mehrzahl der DLBCL exprimieren verschiedene Oberflächenmarker wie. z.B. CD19, CD20, CD22 und CD79A, wobei nicht immer in jedem Tumor alle Marker exprimiert werden. 75% aller DLBCL besitzen zellmembranständige oder zytoplasmatische Immunglobuline (IgM, IgG oder IgA). Zusätzlich können bei einigen DLBCL die Oberflächenmarker CD5 (10%) oder die Keimzentrumsmarker CD10 und BCL6 (25-50%) nachgewiesen werden. Die Expression von BCL2 liegt bei etwa 30 bis 50% in DLBCL, wohingegen Plasmazell-assoziierte B-Zell-Marker (z.B. CD138) in DLBCL eher selten exprimiert werden. Im Normalfall sind mind. 40% aller DLBCL positiv für den Wachstumsfaktor Ki-67⁸⁹.

1.3.3 Therapie und klinische Studien von Patienten mit einem Keimzentrums-B-Zell-Lymphom

Aus klinischer Sicht gibt es verschiedene Faktoren die zur Beurteilung der Prognose der Patienten herangezogen werden. Für jeden Patienten mit einem hochmalignen Non-Hodgkin-Lymphom (NHL) wird der Internationale prognostische Index (IPI¹⁰⁴) ermittelt, der sich aus dem Alter, dem Krankheitsstadium nach Ann-Arbor-Klassifikation (Stadien I-IV), der Lokalisation und Ausbreitung der Tumore (nodal und/oder extranodal), dem Laktatdehydrogenase (LDH)-Spiegel und dem Allgemeinzustand nach Einteilung der *Eastern Cooperative Oncology Group* (ECOG) zusammensetzt. Anhand dieser Faktoren wird der Patient in eine der vier Risikogruppen eingeteilt (niedrig, niedrig-intermediär, hoch-intermediär, hoch) und es werden statistische Werte für die Remissionsdauer und das Gesamtüberleben des Patienten berechnet. Bei der Berechnung des IPI wird bislang ausschließlich das 60. Lebensjahr als Altersgrenze gewählt, wobei Patienten über 60 Jahren eine ungünstigere Prognose als jüngeren Patienten prophezeit wird. Eine schematische Übersicht der R-CHOP-Therapie bei DLBCL-Patienten ist in der nachfolgenden Abbildung 3 dargestellt.

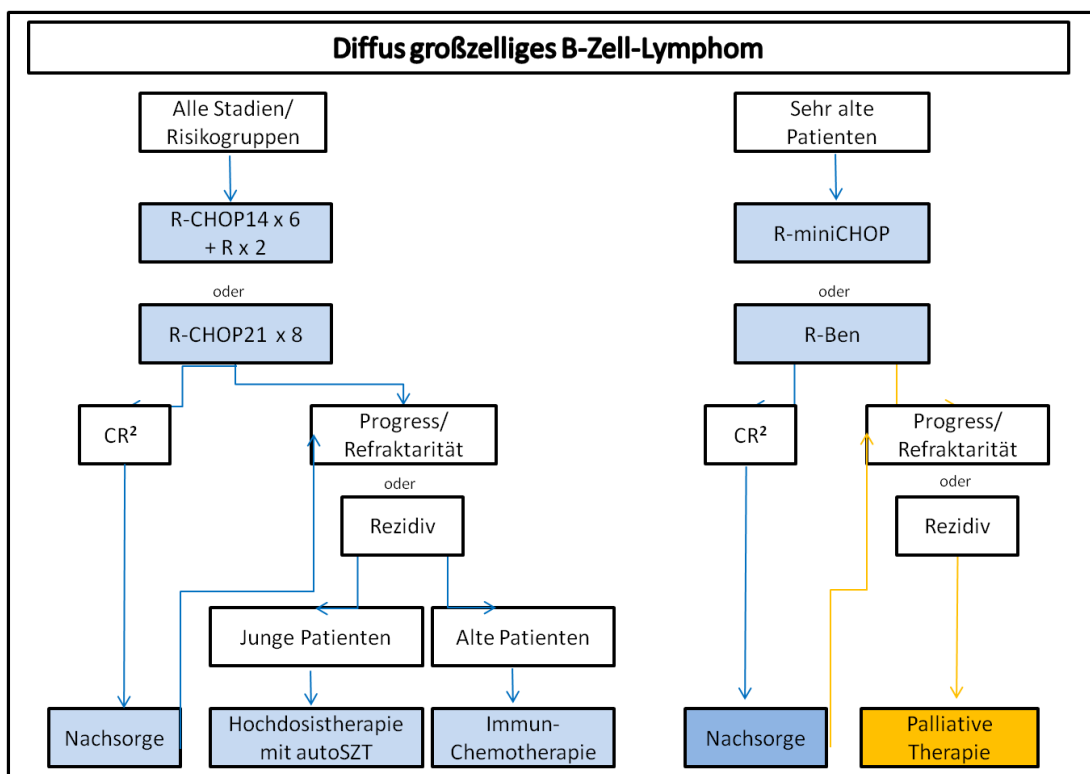


Abbildung 3: Standardmäßige Therapie bei Patienten mit einem DLBCL

Standardmäßig werden heutzutage alle Patienten mit DLBCL mit einer Kombination aus Chemotherapie und dem monoklonalen CD20-Antikörper Rituximab behandelt. Als standardmäßige Behandlung für Rezidive bei jungen Patienten (<60 Jahre) wird eine autologe Stammzelltransplantation und alten Patienten (>60 Jahre) eine erneute Immunchemotherapie gewählt. Patienten im Senium (~80 Jahre) werden mit einer dosisreduzierten Chemotherapie oder einem palliativen Therapieansatz behandelt. Abkürzungen und Verweise: Ben: Bendamustin, CHOP: Cyclophosphamid, Hydroxydaunorubicin, Oncovin, Prednisolon, Cyclophosphamid. CHOP14: Wiederholung der Chemotherapie alle 14 Tage. CHOP21: Wiederholung der Chemotherapie alle 21 Tage. miniCHOP: dosisreduziertes CHOP. R: Rituximab. CR: komplette Remission. autoSzt: autologe Stammzelltransplantation. Junge Patienten: Patienten unter 60 Jahre. Alte Patienten: Patienten über 60 Jahre. Blau markiert: kurative Therapie. Gelb markiert: palliative Therapie (modifiziert nach <https://www.onkopedia.com>).

Derzeit werden DLBCL-Patienten standardmäßig mit einer Kombination aus Chemotherapie (CHOP: Cyclophosphamid, Hydroxydaunorubicin, Oncovin, Prednisolon) und dem monoklonalen CD20-Antikörper Rituximab behandelt. Diese Form von Immunchemotherapie wird oftmals als R-CHOP bezeichnet und erfolgt entweder mit einem 14- oder 21-tägigen Intervallprotokoll¹⁰⁵⁻¹⁰⁹. Als standardmäßige Behandlung für Rezidive bei Patienten mit einem Ersterkrankungsalter unter 60 Jahren wird die autologe Stammzelltransplantation gewählt¹¹⁰. Voraussetzung hierfür ist neben dem Alter auch, dass keine Therapie-limitierenden Komorbiditäten bei dem Patienten vorliegen. Bei älteren Patienten mit einem Rezidiv wird aufgrund ihres Alters und der Komorbiditäten von einer Stammzelltransplantation abgesehen. Als erste Alternative zur Standardtherapie besteht für diese Patienten des Seniums die Möglichkeit ein Dosis-reduziertes CHOP-Protokoll (R-miniCHOP) zur Behandlung einzusetzen¹¹¹. Patienten des Seniums mit einer sehr schlechten Anamnese (hoher IPI) oder Rezidiven werden häufig nur palliativ behandelt, da das Risiko höher ist, dass diese sehr alten Patienten an den Folgen der Nebenwirkungen der Immunchemotherapie als an ihrem Lymphom sterben. Bei kurzen Intervallen zwischen der primären Diagnose und dem Rezidiv oder bei chemorefrakteren Erkrankungen wird größtenteils von einem kurativen Therapieansatz Abstand genommen und eine palliative Therapie gewählt¹¹².

Die Arbeit von Peyrade et al.¹¹¹ beschreibt die bisher einzige klinische Studie, die gezielt Patienten über 80 Jahren betrachtet hat, die zuvor mit R-miniCHOP behandelt wurden. Insgesamt wurden Daten von 149 DLBCL-Patienten (80-95 Jahre, Median: 83 Jahre) zwischen 2006 und 2009 aus 38 Zentren in Frankreich und Belgien rekrutiert und mit den Daten jüngerer DLBCL-Patienten (<80 Jahren) verglichen. Die Ergebnisse der Studie zeigten, dass R-miniCHOP für sehr alte Patienten mit einer guten Anamnese (geringer IPI) eine vielversprechende und wirksame Therapie darstellt. Jedoch besteht weiterhin der Bedarf an neuen Behandlungsstrategien für sehr alte Patienten mit einer schlechteren Anamnese (hoher IPI)¹¹¹. Als Alternative für Patienten mit Herzinsuffizienz, die das Chemotherapieprotokoll CHOP aufgrund der Anthracycline nicht erhalten können, bietet sich z.B. die Therapie mit Gemcitabin¹¹³ oder eine Kombination aus Rituximab, Gemcitabin und Oxaliplatin (R-GemOx¹¹⁴). Ansonsten werden bislang alle Patienten mit einer sehr schlechten Anamnese (hoher IPI) ausschließlich palliativ behandelt. Aktuell wird im Rahmen der deutschen Studiengruppe für hochmaligne Non-Hodgkin Lymphome (DSHNHL) die sog. BRENDA-Studie durchgeführt, worin Patienten über 80 Jahren oder ältere Patienten mit Komorbidität mit einer Kombination aus Rituximab und Bedamustin behandelt werden. Ergebnisse zu Brenda liegen zum aktuellen Zeitpunkt noch keine vor, da diese Studie noch aktiv ist¹¹⁵. Demgegenüber wurden in den vergangenen Jahren bereits verschiedene klinische Studien in Deutschland durchgeführt (z.B. NHL-B und RICOVER-60), die Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem Ersterkrankungsalter von max. 80 Jahren analysiert haben.

NHL-B ist eine bereits abgeschlossene klinische Studie der deutschen Studiengruppe Hochmaligne Non-Hodgkin-Lymphome (DSHNHL) zur Analyse von Patienten mit einem aggressiven Non-Hodgkin-Lymphom (NHL) der B-Zellreihe, die sich aus den Teilstudien NHL-B1 und –B2 zusammensetzt. NHL-B1 umfasst junge Patienten (<61 Jahre) mit einem aggressiven NHL und einem normalen Laktat-Dehydrogenasewert (LDH)¹¹⁶. Hingegen wurden in NHL-B2 ältere Patienten aller Stadien und Risikogruppen im Alter von 61 bis 75 Jahren rekrutiert¹¹⁷. Ziel der Studie war eine Verbesserung der Immunchemotherapie mittels CHOP für Patienten mit einem aggressiven Keimzentrums-B-Zell-Lymphom. Hierzu wurde in einer randomisierten Analyse die zusätzliche Gabe von Etoposid (CHOEP) und/ oder die Reduzierung der Therapieintervalle von 21 (CHOPE21) auf 14 Tagen (CHOPE14) bewertet. Im Rahmen von NHL-B wurden von 09/1993 bis 06/2000 die Therapiedaten von insgesamt 1399 Patienten mit einem Keimzentrums-B-Zell-Lymphom (größtenteils DLBCL) im Alter bis zu 75 Jahren rekrutiert. Die Ergebnisse der Studie zeigten, dass die Behandlung von älteren Patienten mit einem aggressiven Keimzentrums-B-Zell-Lymphom mittels CHOEP nicht effektiver ist als eine CHOP-Behandlung. Zusätzlich waren keine signifikanten Unterschiede in den Überlebensraten zwischen einem 14- und 21-tägigen Therapieintervall zu erkennen. Als Ergebnis der NHL-B-Studie wurde daraufhin als Behandlung für ältere Patienten mit einem Keimzentrums-B-Zell-Lymphom die Immunchemotherapie mittel CHOP14 empfohlen¹¹⁷.

RICOVER-60 ist eine bereits abgeschlossene klinische Studie von Patienten mit einem aggressiven NHL. Ziel der Studie war eine Therapieoptimierung der Immunchemotherapie CHOP im 14-tägigen Intervall (CHOP14). Hierzu wurde in einer randomisierten Analyse die zusätzlich Gabe von dem monoklonalen Antikörper Rituximab und/ oder die Erhöhung der Zyklusanzahl von 6 auf 8 Chemotherapien bewertet. Im Rahmen von RICOVER-60 wurden von 07/2000 bis 06/2005 die Therapiedaten von insgesamt 1222 Patienten mit einem Keimzentrums-B-Zell-Lymphom (CD20⁺, größtenteils DLBCL) im Alter von 61 bis 80 Jahren rekrutiert. Für jeden Patienten wurde eine dreijährige Nachsorge durchgeführt. Das Hauptkriterium zur Bewertung der Therapie war die Zeit bis zum Therapieversagen. Zusätzlich wurden die langfristigen Nebenwirkungen, mögliche Rezidive und die Überlebensrate begutachtet. Die Ergebnisse der Studie zeigten, dass die Behandlung von älteren Patienten mit einem aggressiven Keimzentrums-B-Zell-Lymphom mit Rituximab und 6 Zyklen von CHOP14 am effektivsten war und als Therapieansatz empfohlen wurde¹¹⁸.

1.4 Molekulare Pathogenese und Forschungsstudien von Patienten mit einem Keimzentrums-B-Zell-Lymphom

In den vergangenen Jahren wurden in verschiedenen internationalen Verbundprojekten und Kompetenz-Netzwerken molekulargenetische Analysen zur Charakterisierung von Keimzentrums-B-Zell-Lymphomen durchgeführt, wobei insbesondere DLBCL und FL hinsichtlich genetischer und epigenetischer Aberrationen analysiert wurden. In Deutschland wurden derartige Analysen von Patienten mit einem Keimzentrums-B-Zell-Lymphom in den Forschungsstudien MMML und ICGC MMML-Seq durchgeführt.

Molekulare Mechanismen der malignen Lymphome (MMML) ist ein Verbundprojekt von deutschen Forschungsgruppen und Versorgungseinrichtungen, die im Bereich der Lymphomerkrankungen tätig sind¹¹⁹. Die wissenschaftlichen Arbeiten des MMML-Verbundprojektes wurden von 10/2003 bis 03/2011 von der Deutschen Krebshilfe gefördert, wobei in diesem Zeitraum ein Gesamtkollektiv von 895 aggressiven B-Zell-Lymphomen zusammengestellt wurde (größtenteils DLBCL, FL und BL). Ziel der Forschungsstudien war die Verbesserung der Prognose und Therapie von Patienten mit einem malignen Lymphom durch die Nutzung klassischer histopathologischer und molekulargenetischer Methoden (z.B. IHC, FISH) sowie neuer molekulargenetischer Verfahren (z.B. Array-CGH, SNP-Arrays). Es erfolgte eine zentrale Speicherung der klinischen und molekulargenetisch erhobenen Daten, sodass diese Daten auch für weiterführende Analysen und Fragestellungen nachhaltig zur Verfügung standen.

Die Analyse Molekularer Mechanismen in Malignen Lymphomen durch Sequenzierung (ICGC MMML-Seq) ist ein Teilprojekt des internationalen Krebsgenomprojektes und wurde vom Institut für Humangenetik in Kiel durch Herrn Prof. Dr. Siebert koordiniert¹²⁰. ICGC MMML-Seq ist ein Verbundprojekt, das an das zuvor beschriebene MMML Projekt anknüpfte. Zielführend in diesem Forschungsprojekt war die Identifizierung von tumorspezifischen genetischen und epigenetischen Aberrationen bei Keimzentrums-B-Zell-Lymphomen unter Verwendung von histopathologischen und molekulargenetischen Methoden (z.B. IHC, FISH) und modernen Hochdurchsatz-Verfahren (z.B. NGS, SNP-Arrays). Dazu wurden alle rekrutierten Patienten mittels Fluoreszenz-in-situ-Hybridisierung (FISH), Immunhistochemie (IHC) und HumanMethylation450 Bead Chip (450K) analysiert. Alle Patienten von denen die korrespondierenden nicht-Tumor-Probe vorhanden war, wurden mittels *Whole Genome Sequencing* (WGS) untersucht. Zusätzlich wurde für eine Auswahl von etwa 60 Patienten ein *Whole Genome Bisulfite Sequencing* (WGBS) durchgeführt. Ein weiteres Ziel der Studie war die Erstellung von Genexpressionsprofilen maligner Lymphome, weshalb zusätzlich von allen Patienten (wenn möglich) eine RNA- und ggf. miRNA-Sequenzierung durchgeführt wurde.

Das ICGC MMML-Seq Projekt wurde seit 2010 vom Bundesministerium für Bildung und Forschung gefördert und endete offiziell im Januar 2016. Insgesamt wurden über die komplette Förderperiode des ICGC MMML-Seq über 250 aggressive Keimzentrums-B-Zell-Lymphome mittels WGS analysiert.

Im Rahmen der klinischen und Forschungsstudien in Deutschland sowie den Analysen anderer internationaler Arbeitsgruppen konnte eine Charakterisierung von Keimzentrums-B-Zell-Lymphomen erfolgen, worin genetische und epigenetische Aberrationen identifiziert wurden, die spezifisch in verschiedenen Lymphom-Entitäten nachweisbar sind.

Beispielsweise ist inzwischen bekannt, dass 85 bis 90% aller FL die Ig-assoziierte Translokation $t(14;18)^{121-124}$. Bei dieser chromosomalen Imbalanz gelangt das *BCL2*-Gen (Chr. 18) in die unmittelbare Nähe der regulatorischen Elemente des *IgH*-Gens (Chr. 14) und wird daraufhin dauerhaft überexprimiert (Abbildung 4A)¹²⁵. Das anti-apoptotische Protein BCL2 ist in der Mitochondrienmembran lokalisiert und verhindert die Freisetzung von Cytochrom C, einem kleinen Protein von etwa 100 Aminosäuren, das von essentieller Bedeutung in der Initiierung der Apoptose-Signalkaskade ist^{123,126}. Ursächlich für die Translokation $t(14;18)$ sind Fehlern in der frühen B-Zell-Entwicklung während der (V)DJ-Gensegmentumlagerung oder in selteneren Fällen innerhalb der Keimzentrums-Reaktion im Verlauf des Ig-Klassenwechsels⁹². Zusätzlich wurden in über 80% aller FL mono- oder biallelische Aberrationen in *KMT2D/MLL2* identifiziert, die zur Inaktivierung des Gens führten^{62,127}. Im Regelfall wird durch die Aktivierung von *KMT2D/MLL2* ein Protein exprimiert, das funktionell als H3K4 Methyltransferase agiert und von entscheidender Bedeutung für die Regulation von Histonmodifikationen ist (Abbildung 4B)^{128,129}. Weiterhin wurden genetische Aberrationen in den Histon-modifizierenden Genen *EZH2* (~20%)^{130,131}, *CREBBP* (~30%)^{62,132} und *EP300* (~40%)^{62,132} identifiziert.

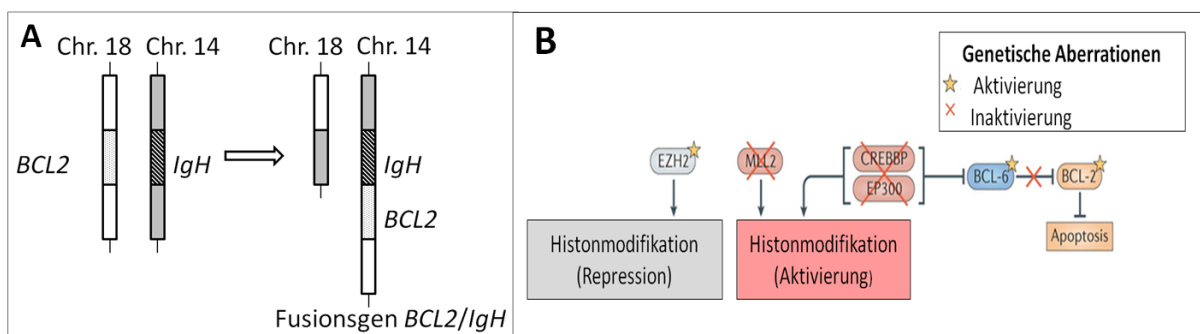


Abbildung 4: Schematische Darstellung der Translokation $t(14;18)$ und des veränderten Signalweges bei der Lymphogenese in FL. A: Die grafische Darstellung zeigt die für folliculäre Lymphome charakteristische reziproke Translokation $t(14;18)$. Hierbei wird die Region des *BCL2*-Gens (gepunktet) von dem Chromosom 18 (weiß) an die funktionelle Region des *IgH*-Gens (gestreift) auf dem Chromosom 14 (grau) umgelagert, wodurch das Fusionsgen *BCL2/IgH* entsteht. B: Die Überexpression von BCL2 und epigenetische Aberrationen in Histonmodifizierenden Genen sind die häufigsten onkologischen Veränderungen in FL. Aufgrund der verstärkten Expression des antiapoptotischen Proteins BCL2 wird der Prozess der Apoptose unterbunden und die malignen B-Zellen überleben. Weiterhin führen die Mutationen in den Methyltransferase *KMT2D/MLL2*, *CREBBP*, *EP300* und *EZH2* zu Veränderungen im Epigenom der malignen B-Zellen. Die verschiedenen Farben in der Grafik kennzeichnen die unterschiedlichen Signalwege und Auswirkungen (modifiziert nach Basso et al., 2015⁴²). Abkürzungen: Die Gennamen wurden nicht in die deutsche Sprache übersetzt sondern sind nur in Englischer Sprache angegeben und stammen von der Internetseite GeneCards (<http://www.genecards.org/>, Stand Januar 2016).

Ebenso konnte in klinischen und Forschungsstudien gezeigt werden, dass DLBCL durch eine Vielzahl von genetischen Aberrationen gekennzeichnet sind, wovon die für FL typische Translokation t(14;18) auch bei DLBCL eine der häufigsten Veränderungen darstellt (20-30%)¹³³. Möglicherweise handelt es sich bei einigen dieser t(14;18)-positiven DLBCL um sekundär transformierte FL. Weiterhin wurde in bis zu 30% aller DLBCL eine genetische Aberration in der 3q27-Region nachgewiesen, was normalerweise zu Veränderungen im *BCL6*-Gen führt^{76,102}.

In den letzten Jahren konnten verschiedene Arbeitsgruppen zeigen, dass genetische Aberrationen in DLBCL wie z.B. somatische Hypermutation (SHM), chromosomale Deletionen oder balanzierte Translokationen die Genexpression der Proteine von verschiedenen Protoonkogenen (z.B. *BCL6*, *REL*, *BCL2*, *MYC*) beeinflusst und eine Deregulation der Protoonkogene oftmals mit der Apoptose oder Defekten in den DNA-Reparaturmechanismen assoziiert ist^{102,103,127,134-137}. Aus diversen *Whole Genome Sequencing* (WGS)- und *Whole Exome Sequencing* (WES)-Studien sind bereits über 300 Gene bekannt, die als rekurrent mutiert in primären DLBCL detektiert wurden^{62,127,132,134-140}. Einige dieser rekurrent mutierten Gene wurden bereits als funktionell relevant in DLBCL identifiziert, wohingegen von vielen Genen die funktionelle Relevanz in DLBCL noch ungeklärt ist^{127,135,137,140}. So ist aus diesen vergangenen WGS- und WES-Studien bekannt, dass chromosomale Imbalancen in den Onkogenen *BCL6*, *BCL2*, *MYC* oder *REL* ein primäres bzw. frühes Event in der Lymphomentstehung sind und zusätzliche sekundäre bzw. spätere Events die Klonalität beeinflussen. Diese sekundären Aberrationen weisen häufig rekurrente Mutationen bzw. genetische Aberrationen u.a. in den Genen *PRDM1*, *MYD88*, *CREBBP*, *TP53*, *EZH2*, *KMT2D/MLL2*, *PIM1*, *B2M*, *CD79B* und *CD58* auf^{62,127,132,134-140}.

Vor den molekulargenetischen Analysen im Rahmen des MMML Projektes wurden DLBCL ausschließlich hinsichtlich der histopathologischen Ergebnisse der Expression der Proteine CD10, *BCL6* und *MUM1/IRF4* nach Hans et al.¹⁴¹ in die Subtypen nonGCB und GCB eingeteilt. Diese Klassifikation wird in der vorliegenden Arbeit als *Hans-Klassifikator* bezeichnet. Innerhalb der Forschungsstudien konnten neue Erkenntnisse über die molekulare Pathogenese von Keimzentrum-B-Zell-Lymphomen erzielt und publiziert werden^{142,143}. Basierend auf den Erkenntnissen der Studien lassen sich DLBCL trotz der genetischen und klinischen Heterogenität nach ihrem Genexpressionsprofil (GEP) in die molekularen Subtypen hinsichtlich ihrer Ursprungs-B-Zelle (sog. „cell of origin“) in ABC und GCB einteilen^{142,143}. Insgesamt machen Lymphome des ABC- und GCB-Subtypen etwa 80% aller DLBCL aus^{134,144}. Zusätzlich gibt es einen Subtypen *unclassified*, worin alle Lymphome zusammengefasst werden, deren GEP nicht eindeutig dem ABC- oder GCB-Typ zugeordnet werden kann. Diese Klassifikation wird in der vorliegenden Arbeit als *GEP-Klassifikator* bezeichnet. Detaillierte Informationen zu der molekularen Pathogenese des GCB- und ABC-Subtypen sind in den nachfolgenden Abschnitten beschrieben.

GCB-Subtyp

Der zelluläre Ursprung der Tumorzelle vom GCB-Subtyp ist vermutlich die Keimzentrums-B-Zelle, da das GEP-Muster dem normaler Keimzentrums-B-Zellen ähnelt¹⁰⁵. Es besteht die Hypothese, dass die Entstehung von DLBCL-GCB in der Keimzentrums-Reaktion der B-Zell-Differenzierung vor dem Zeitpunkt der Differenzierung zur aktivierten B-Zelle einzuordnen ist. Charakteristisch für DLBCL-GCB sind genetische Aberrationen in Keimzentrums-spezifischen Oberflächenproteinen (z.B. CD10, CD38)^{142,145}, Transkriptionsfaktoren (z.B. BCL6)¹⁴⁶ und den Acetyltransferasen CREBBP und EP300. In DLBCL-GCB weisen beide Acetyltransferasen verstärkt inaktivierende Mutationen auf, was zu Acetylierungsdefekten und einer erhöhten BCL6-Aktivität führt^{132,143,146-148}. Bei DLBCL-GCB können verschiedene molekulare Mechanismen die Regulation der Apoptose inhibieren, wobei chromosomale Imbalancen der Gene *MYC* und *BCL2* sowie die Translokation t(14;18), die charakteristisch für FL sind, zu den wichtigsten Mechanismen zählen^{51,126,146,149,150}. Dabei ist das gleichzeitige Auftreten von Aberrationen in *MYC* und *BCL2* mit einem aggressiveren klinischen Krankheitsverlauf assoziiert^{151,152}.

ABC-Subtyp

Der zelluläre Ursprung der Tumorzellen vom ABC-Subtyp konnte bisher nicht eindeutig geklärt werden, jedoch ähnelt das GEP-Muster dem aktivierter B-Zellen. Daher besteht die Hypothese, dass die Entstehung des DLBCL-ABC im fortgeschrittenen Prozess der B-Zell-Differenzierung (aktivierte B-Zellen, Prä-Keimzentrum) vor dem Zeitpunkt der Differenzierung zur Plasmazelle einzuordnen ist^{142,145}. DLBCL-ABC weisen zahlreiche Mutationen in den Ig-Genen auf, was dafür spricht, dass die Tumorzellen den Prozess der SHM durchlaufen haben¹⁵³. Weiterhin wurden beim DLBCL-ABC-Subtypen u.a. Aberrationen im Transkriptionsfaktor BLIMP1 (PRDM1) detektiert, was eine essentielle Rolle in der terminalen Differenzierung der aktivierten B-Zellen zu Plasmazellen hat^{134,147,154-157}. Zudem konnte in über 50% der DLBCL-ABC eine Mutation in Genen nachgewiesen werden, die den NF-κB-Signalweg regulieren wie z.B. *MYD88* und *CD79B*¹³⁸. Die Pathogenese von DLBCL-ABC wird demnach von zwei Ereignissen dominiert: zum Einen der Blockierung der terminalen Differenzierung zur Plasmazelle und zum Anderen der Aktivierung des NF-κB-Signalweges. Eine schematische Übersicht der Lymphogenese in DLBCL vom ABC-Subtyp ist in der nachfolgenden Abbildung 5 dargestellt.

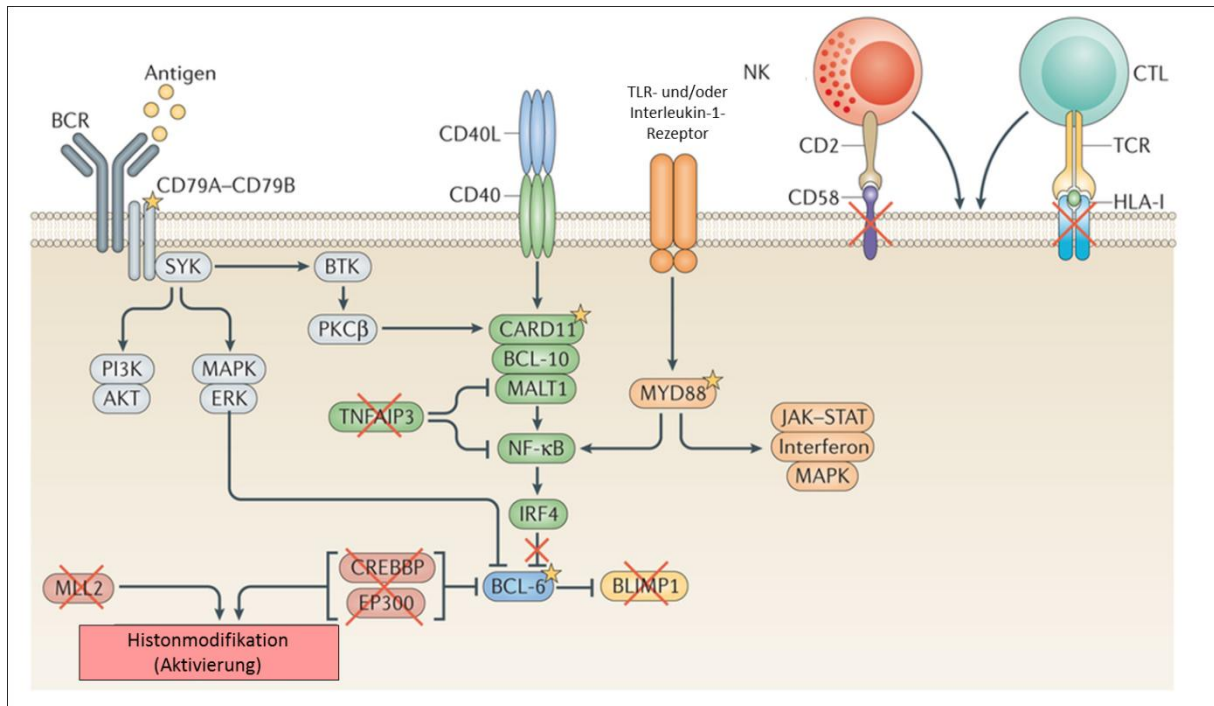


Abbildung 5: Schematische Darstellung der veränderten Signalwege bei der Lymphogenese in DLBCL-ABC

DLBCL-ABC sind assoziiert mit der Aktivierung des NFκB-Signalweges, wobei die transkriptionelle Aktivität von NFκB zum einen durch aktivierende Mutationen in den Transkriptionsfaktoren *CD79A*, *CD79B* oder *CARD11* und zum Anderen durch inaktivierenden Mutationen in *TNFAIP3* reguliert werden können. Infolgedessen kommt es zur Blockade der terminalen Differenzierung der Keimzentrums-B-Zellen zu Plasmazellen, wodurch *BLIMP1* transkriptionell inaktiviert wird. Weiterhin weisen die Keimzentrums-B-Zellen verschiedene Mutationen auf, die möglicherweise zur Hemmung des Immunsystems führen können. Diese genetischen Aberrationen wurden u.a. in den Genen *CD58*, *KMT2D/MLL2*, *CREBBP* und *EP300* detektiert (modifiziert nach Basso et al., 2015⁴²). Abkürzungen: BCR: B-Zell-Rezeptor. CTL: Zytotoxische T-Zelle. NK: Natürliche Killerzelle. TCR: T-Zell-Rezeptor.

Vergleich GCB- und ABC-Subtyp

In den vergangenen Jahren wurden zahlreiche genetische Aberrationen identifiziert, die charakteristisch für die jeweiligen DLBCL-Subtypen sind. Ein Teil der genetischen Aberrationen für DLBCL-ABC und –GCB sind in der nachfolgenden Tabelle 1 kurz zusammengefasst.

Tabelle 1: Genetische Aberrationen bei Diffus großzelligen B-Zell-Lymphomen

| Gen | Genetische Aberration | DLBCL-ABC [%] | DLBCL-GCB [%] | Funktion | Referenz |
|---------------------|------------------------|---------------|---------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| <i>BCL2</i> | Translokation t(14;18) | 0 | 45 | Inaktivierung der Apoptose | Lenz et al., 2008 ¹³⁴ |
| | Amplifikation 18q21 | 47 | 16 | | |
| <i>BCL6</i> | Translokation t(3q27) | 24 | 10 | | Iqbal et al., 2007 ¹⁴⁷ |
| | Mutation | 44 | 74 | | |
| <i>TP53</i> | Mutation | 14- 24 | 10- 34 | Zainuddin et al., 2009 ¹⁵⁸ Young et al., 2008 ¹⁵⁹ | |
| <i>KMT2D</i> | Mutation | 20 | 27 | Epigenetische Modifikationen | Pasqualucci et al., 2011 ¹²⁷ |
| <i>EZH2</i> | Mutation | 0 | 21 | | Morin et al., 2010 ¹³⁰ |
| <i>CREBBP</i> | Mutation | 17 | 41 | | Pasqualucci et al., 2011 ¹³² |
| <i>EP300</i> | Mutation | 1 | 6 | | Pasqualucci et al., 2011 ¹²⁷ |
| <i>TNFAIP3</i> | Mutation | 24 | 2 | | Aktivierung des NFκB-Signalweg |
| | Deletion | 50 | 22 | | |
| <i>MYD88</i> | Mutation | 29 | 1 | Ngo et al., 2011 ¹⁶⁰ | |
| <i>CARD11</i> | Mutation | 10 | 4 | Lenz et al., 2008 ¹⁶¹ | |
| <i>CD79B</i> | Mutation | 21 | 3 | Davis et al., 2010 ¹⁶² | |
| <i>SPIB</i> | Amplifikation 19q13.3 | 26 | 3 | Störung in der B-Zell-Differenzierung | Lenz et al., 2008 ¹³⁴ |
| <i>BLIMP1/PRDM1</i> | Mutation/ Deletion | 24 | 0 | | Tam et al., 2006 ¹⁵⁴ Pasqualucci et al., 2006 ¹⁵⁵ |
| <i>CDKN2A</i> | Deletion (9p21) | 30 | 4 | Verlust der Zellzyklusregulation | Lenz et al., 2008 ¹³⁴ |

Abkürzungen: DLBCL: Diffus großzelliges B-Zell-Lymphom. ABC: Aktivierte B-Zellen. GCB: Keimzentrumsähnliche B-Zellen.

1.5 Keimzentrums-B-Zell-Lymphome im Senium

In den vergangenen Jahren haben zahlreiche Studien molekulargenetische Analysen von Keimzentrums-B-Zell-Lymphomen im Kindes- und Erwachsenenalter durchgeführt, jedoch sind nur wenige Daten von Patienten mit einem sehr hohen Ersterkrankungsalter vorhanden¹⁶³⁻¹⁶⁵. In Zentraleuropa sind FL und DLBCL die weitaus häufigsten Lymphome des Erwachsenenalters und Seniums und umfassen etwa 70-80% aller B-Zell-Lymphome im hohen Alter⁸⁹. In der Literatur wird die Altersgrenze des Seniums variabel verwendet und überwiegend zwischen 60 und 80 Jahren gesetzt^{166,167}.

Aus biologischer Sicht steigt die Lebenserwartung der Menschen in den westlichen Ländern kontinuierlich auch im Sinne des "Healthy Ageing". Darüber hinaus ist zu erwarten, dass unter der Hypothese, dass es altersassoziierte Aberrationen in den Mechanismen der B-Zell-Lymphomentstehung gibt, eine „Extremgruppe“ letztlich biologisch homogener ist als eine weniger stringent definierte Population. Vergangene Projekte haben diese Strategie der „Extremgruppenanalyse“ bereits erfolgreich bei FL und DLBCL im jungen Alter belegen können. Durch die Analyse von FL und DLBCL im Kindesalter (bis 18 Jahre) konnte gezeigt werden, dass diese, wenn auch selten, so doch biologisch distinkt sind und sich von denselben Lymphomen des Erwachsenenalters unterscheiden. Nach der Definition der molekularen Mechanismen der B-Zell-Lymphom-Entstehung in der "Extrempopulation" des Kindesalters (z.B. IRF4-Translokation positives FL3/DLBCL, t(14;18)-negatives pädiatrisches FL) konnten in der Folge gezeigt werden, dass diese Lymphome auch im höheren Alter mit abnehmender Inzidenz auftreten, in der Gesamtfrequenz dort aber viel seltener sind¹⁶³⁻¹⁶⁵.

Aus klinischer Sicht stellt derzeit das Ersterkrankungsalter von 60 Jahren eine Altersgrenze zwischen niedrigem und hohem Risiko dar (Risikofaktoren: IPI¹⁰⁴, FLIPI¹⁶⁸, MIPI¹⁶⁹), weshalb in einer Reihe von klinischen Studien die Patienten mit einem Ersterkrankungsalter ab 60 Jahren analysiert wurden (z.B. RICOVER-60, Teile der NHL-B Studie). In diesen Studien wurden im wesentlichen Patienten mit einem Ersterkrankungsalter von 60 bis 80 Jahren rekrutiert¹¹⁸.

Die epidemiologischen Daten (siehe Abbildung 6) zeigen, dass die Inzidenz von DLBCL und FL (im Gegensatz z.B. zu BL oder B-Linien ALLs) mit dem Alter kontinuierlich zunimmt¹⁷⁰. Jedoch zeigt die genaue Betrachtung der SEER-Daten, dass die Inzidenz bei FL (und auch z.B. bei Mantelzell-Lymphomen) ab dem 65.Lebensjahr quasi konstant bleibt. Hingegen steigt die Inzidenz von DLBCL (und auch z.B. bei Marginalzonen-Lymphomen) auch nach dem 65. Lebensjahr weiter an.

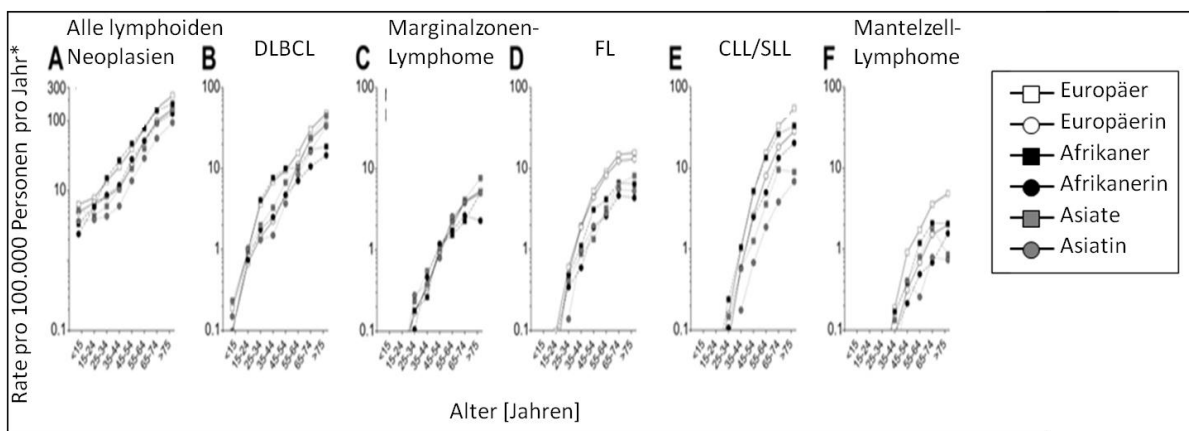


Abbildung 6: Inzidenz von lymphoiden Neoplasien entsprechend Subtyp, Ethnie, Geschlecht und Alter

12 SEER Register, 1992-2001 (Morton et al., Blood, 2006¹⁷⁰; Figure 2). *Alle Inzidenzraten sind innerhalb der Altersgruppen zu der "2000 United States Population" alters adjustiert. Insgesamt ist zu erkennen, dass die Inzidenz bei FL-Patienten ab dem 65.Lebensjahr konstant bleibt (D) und die Inzidenz von DLBCL weiter zunimmt (B).

Diese Beobachtungen stehen im Einklang mit den molekulargenetischen Daten verschiedener Forschungsgruppen, die zeigen, dass mit dem Alter eine Abnahme von Lymphomen mit der für FL charakteristischen Translokation t(14;18) zu verzeichnen ist¹⁴³. Mit zunehmendem Alter bzw. im sehr hohen Alter wurden verstärkt DLBCL vom ABC-Subtyp ohne t(14;18) diagnostiziert^{144,171,172}. Aufgrund der Epidemiologie und den molekulargenetischen Daten besteht die Hypothese, dass eine Aktivierung des NF-κB –Signalweges eine Rolle in der Lymphomentstehung bei sehr alten Patienten mit einem Keimzentrums-Lymphom spielt und dies möglicherweise eine Therapieoption darstellt.

Zusätzlich zeigen die epidemiologischen Daten, dass es Unterschiede in den Charakteristika von alten Patienten mit Keimzentrums-B-Zell-Lymphomen gibt, die abhängig von dem Lebensraum der betroffenen Patienten sind (z.B. EBV-Status). Das EBV-positive DLBCL des "älteren Patienten" ("of the elderly") ist eine provisorische Kategorie der WHO-Klassifikation und setzt den Nachweis von EBV-Virus in den Lymphomzellen voraus. Die EBV-Aktivierung in den Lymphomzellen wird als Hinweis auf einen Verlust der Immunkontrolle über das EBV-Virus im höheren Lebensalter interpretiert. DLBCL, die im Krebszentrum Nord innerhalb der letzten 10 Jahre behandelt wurden, wurden im Archiv des Lymphknotenregisters identifiziert und systematisch auf EBV-Virus-kodierte RNA (EBER) mittels in situ Hybridisierung untersucht. In Tabelle 2 sind die bisher erzielten Ergebnisse im Vergleich zu den publizierten Studien aus Japan (dem Ort der Erstbeschreibung) und West-Europa zusammengefasst.

Tabelle 2: EBV-positive DLBCL wurden in Patienten >50 Jahre über die EBER in situ Hybridisierung identifiziert

| Land | Charakteristika der identifizierten EBV-positive DLBCL | | | Publikation |
|-------------------------|--------------------------------------------------------|-------------|------------|---------------------------------------------|
| | >50 Jahre | 60-80 Jahre | >80 Jahre | |
| Japan | 22/403 (5%) | 16/22 (73%) | 6/22 (27%) | Oyama et al. 2003 ¹⁷³ |
| Mexiko | 9/136 (7%) | 7/9 (78%) | 0/9 (0%) | Hofscheier et al. 2011 ¹⁷⁴ |
| Deutschland | 4/169 (2%) | 3/4 (75%) | 1/4 (25%) | Hofscheier et al. 2011 ¹⁷⁴ |
| Schweiz, Österreich | 8/243 (3%) | 5/9 (56%) | 2/8 (26%) | Hoeller et al. 2011 ¹⁷⁵ |
| CCC Nord Deutschland | 7/224 (3%) | 6/7 (85%) | 0/7 (0%) | Stuhlmann-Laeisz et al. 2015 ¹⁷⁶ |

So treten beispielsweise EBV-positive DLBCL in West-Europa im Senium (>80 Jahre) eher selten auf. Zählt man alle hier aufgeführten Studien zusammen sind nur 16% der EBV-positiven DLBCL-Patienten über 80 Jahre alt. Insgesamt stellen die EBV-positiven DLBCL des "älteren Patienten" nach diesen Daten nur eine sehr kleine Gruppe von Lymphomen im höheren Alter dar. Der überwiegende Teil der DLBCL in älteren Patienten muss eine EBV-unabhängige Pathogenese aufweisen. Der besonders geringe Anteil von EBV-assoziierten Lymphomen im Senium kann als weiterer Hinweis auf eine für dieses Patientenalter besondere Pathogenese interpretiert werden.

Insgesamt legen die genetischen und epidemiologischen Daten nahe, dass die Lymphome im Senium auf der Akkumulation genetischer Schäden beruhen (im Gegensatz z.B. zu Burkitt-Lymphomen, wo wenige Ereignisse die Tumorigenese determinieren). In diesem Kontext ist interessant, dass nicht nur genetische Schäden sondern auch epigenetische Alterationen mit dem Alter zunehmen. So konnten verschiedene Arbeitsgruppen eine Zunahme der DNA-Methylierung mit dem Alter zeigen^{118,170,177}. Interessanterweise konnten Teschendorff et al.¹⁶ zeigen, dass in Tumoren gerade solche Gene aberrant methyliert sind, die auch mit dem Altern eine zunehmende Methylierung zeigen¹⁶.

Zusammenfassend ist festzuhalten, dass die epidemiologischen Daten zeigen, dass das Alter der Bevölkerung in den nächsten Jahren weiter kontinuierlich ansteigt und auch die Anzahl an Patienten mit einem Keimzentrums-B-Zell-Lymphom stetig zunehmen wird. Demnach wird es immer wichtiger Keimzentrums-B-Zell-Lymphom bei Patienten im Senium zu therapieren und alternative Therapieoptionen zur R-CHOP-Therapie zu entwickeln. In den vorangegangenen Studien wurden bisher größtenteils Kinder und Erwachsene molekular charakterisiert und spezifische Biomarker für betroffene Patienten dieser Altersgruppen identifiziert. Jedoch zeigten die Ergebnisse der molekularen Analysen, dass es möglicherweise genetische und epigenetische Aberrationen in Keimzentrums-B-Zell-Lymphomen gibt, die vorwiegend bei Patienten mit einem hohen Ersterkrankungsalter auftreten und die als mögliche molekulargenetische Marker für eine altersangepasste Therapie für Patienten im Senium dienen können. Das primäre Ziel der vorliegenden Arbeit bestand daher in der Identifizierung von altersassoziierten Aberrationen bei Keimzentrums-B-Zell-Lymphomen, wozu folgende Fragestellungen bearbeitet werden sollten:

a. Genetische Aberrationen bei sehr alten Patienten

Gibt es Unterschiede zwischen jüngeren und sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom:

- 1.) In der Frequenz von chromosomalen Imbalancen der Gene *BCL2*, *BCL6*, *MYC* und *IGH* ?
- 2.) In der Expression von *BCL2*, *BCL6*, *CD10*, *Ki67* und *MUM1* ?
- 3.) In dem Mutationsprofil von Genen, die häufig in Lymphomen mutiert auftreten ?

b. Epigenetische Aberrationen bei sehr alten Patienten

Gibt es Unterschiede in dem DNA-Methylierungsmuster zwischen:

- 1.) Verschiedenen Keimzentrums-B-Zell-Lymphom-Entitäten ?
- 2.) Jüngeren und sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom?

1.6 Hochdurchsatz-Sequenzierung zur Identifikation von genetischen Aberrationen bei Patienten im Senium

Zur Identifizierung der genetischen Aberrationen in den Keimzentrums-B-Zell wurde die moderne Technologie der Hochdurchsatz-Sequenzierung genutzt, die es ermöglicht zeitgleich tausende Nukleinsäure-Fragmente parallel zu sequenzieren, um Sequenzveränderungen (z.B. Mutationen, kleine Insertionen oder Deletionen, Inversionen) zu identifizieren, was mit der herkömmlichen Sanger-Sequenzierung nicht bzw. nur unter hohem Zeit- und Kostenaufwand möglich ist. Mit Hilfe dieser modernen Technologieplattform können mittlerweile ganze Genome, Epigenome, Transkriptome, Exome oder auch nur spezifische genetische Regionen bzw. Gene analysiert werden. Letzteres wird als Panel-Analyse bezeichnet und findet durch die gezielte Analyse genetischer Regionen sowie Zeit- und Kostenersparnis immer mehr Anwendung in der Klinik und Diagnostik.

1.6.1 Target-Resequenzierung mit dem Verfahren von Illumina

Das eingesetzte Verfahren der Firma Illumina beruht auf dem Prinzip der Sequenzierung bei Synthese (engl. sequencing by synthesis, SBS). Die Sequenzierung und Analyse der DNA-Fragmente erfolgt zumeist von beiden Leserichtungen (3'- und 5'-Richtung), was als sog. „Paired-End-Run“ bezeichnet wird. Zu Beginn der Probenvorbereitung wird die genomische DNA fragmentiert und anschließend werden universelle Sequenzier-Adapter und spezifische Barcodes (Index) für jede zu analysierende Probe an beide Enden der zu analysierende genomischen DNA-Fragmente ligiert (Abbildung 7A). Barcodes sind einzigartige Sequenzen, die spezifisch an die zu analysierende Probe angefügt werden und es ermöglichen zeitgleich mehrere Proben zu analysieren. Die Sequenz-Adapter werden für die Bindung der zu analysierenden DNA-Fragmente an die Flow cell benötigt, wobei letztere einen Glasobjektträger mit immobilisierten Oligonukleotiden darstellt, worauf die Amplifikation und der Sequenzier-Prozess stattfindet. Anschließend werden die verschiedenen Proben-Fragmente inkl. Sequenzier-Adapter-Barcode-Komplexen zu einem Gesamtpool (sog. Library) zusammengeführt und auf die Flow Cell geladen, wo die DNA-Fragmente komplementär an die immobilisierten Oligonukleotide der Flow Cell binden (Abbildung 7B). Durch die Bindung der Sequenzier-Adapter an die Flow Cell-Oligonukleotide entsteht eine DNA-Brücke, welche als Startpunkt für die DNA-Synthese (sog. Brückenamplifikation) dient. Anschließend folgen mehrere Zyklen Denaturierung, Renaturierung und Neusynthese, wodurch klonale Cluster mit einer hohen Dichte von identischen DNA-Fragmenten erzeugt werden. Nun beginnt der eigentliche Sequenzier-Prozess durch Sequenzierung bei Synthese (Abbildung 7C). Dazu werden 4 verschiedene mit Fluoreszenzfarbstoff und mit einer Blockierungsgruppe markierte dNTPs hinzugegeben, wobei diese entsprechend der Basenabfolge komplementär an die DNA-Fragmente durch die Polymerase eingebaut werden. Jeder Nukleotideinbau entspricht einem Zyklus und nach jedem Zyklus wird die Fluoreszenzintensität gemessen und dokumentiert.

Die Anzahl der Zyklen entspricht der Leselänge (sog. Readlänge) der Sequenzierung, die von Library zu Library variiert und von der jeweiligen DNA-Fragmentlänge abhängig ist. Die Readlänge wird vor jeder Sequenzierung manuell in einem sog. *SampleSheet* angegeben. Nach dem Sequenzier-Prozess erfolgt die bioinformatische Auswertung der detektierten Sequenzen (sog. Reads) indem zuerst ein Alignment und Mappen der Konsensus-Sequenz gegen das Referenzgenom wie z.B. hg19 vorgenommen wird, um Veränderungen in den sequenzierten DNA-Fragmenten zu identifizieren (Abbildung 7D).

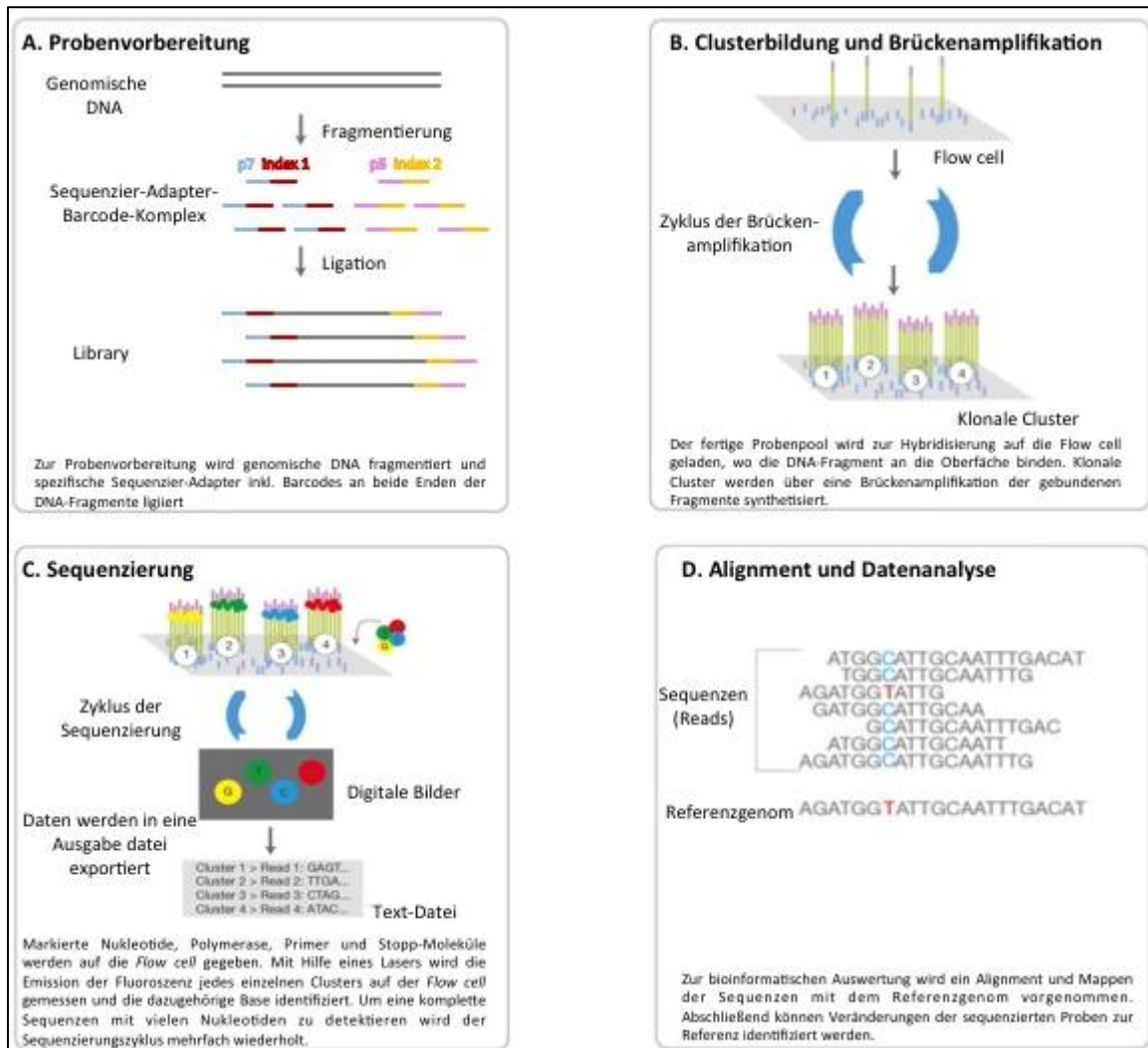


Abbildung 7: Schematische Darstellung zum Arbeitsablauf der Hochdurchsatz-Sequenzierung nach dem Verfahren der Firma Illumina. A: Zur Probenvorbereitung werden Sequenzier-Adapter-Barcode-Komplexe an die fragmentierte DNA ligiert. B: Die Adapter-DNA-Fragmente binden an die komplementären Oligonukleotide auf der Oberfläche der Flow cell wonach mittels Brückenamplifikation klonalen Cluster identischer DNA-Fragmente synthetisiert werden. C: Die Detektion der Sequenz erfolgt über die Messung der Emission von Fluoreszenz nach dem erfolgreichen Einbau des Nukleotides. D: Die detektierten Sequenzen werden mit dem Referenzgenom abgeglichen und genetische Aberrationen identifiziert (modifiziert nach Illumina¹⁷⁸).

Mit Hilfe der Hochdurchsatz-Sequenzierung ist es möglich eine gezielte Re-Sequenzierung genomischer Regionen durchzuführen. Hierbei gibt es verschiedene Möglichkeiten eine Anreicherung von spezifischen Regionen vorzunehmen¹⁷⁹, wobei vornehmlich in zwei Anreicherungsprozesse unterschieden wird: erstens.) PCR-basiert und zweitens.) Hybridisierungs-basiert Verfahren (Abbildung 8).

Bei der Probenvorbereitung mit einem PCR-basierten Verfahren handelt es sich um eine Weiterentwicklung der herkömmlichen PCR und Sanger-Sequenzierung. Mit dem Einsatz dieser Anreicherung wird zielführend eine parallele Tiefensequenzierung von einzelnen Regionen bei mehreren Patienten angestrebt. Hierzu werden spezifische genomische Regionen mittels Multiplex-PCR amplifiziert und jeder Patient mit einem eigenen Identifier (Barcode) versehen. Anschließend wird ein Gesamtpool aller Patienten erstellt und eine Tiefensequenzierung vorgenommen (Abbildung 8A). Hybridisierungs-basierte Verfahren ermöglichen eine hocheffiziente Anreicherung großer Zielregionen von mehreren Kilo- oder Megabasen aus gDNA. Bei der Hybridisierungs-basierte Methodik werden kurze Oligonukleotide von ca. 60 bp verwendet, die spezifisch an die gewünschte Target-Region binden. Es folgt eine Selektion über die spezifische Bindung der DNA an die Oligonukleotide, wobei nicht hybridisierte DNA-Fragmente durch Waschschrte entfernt werden (Abbildung 8B). Abschließend wird eine Amplifikation der Target-spezifischen Regionen durchgeführt, die Library aller Patienten gepoolt und sequenziert.

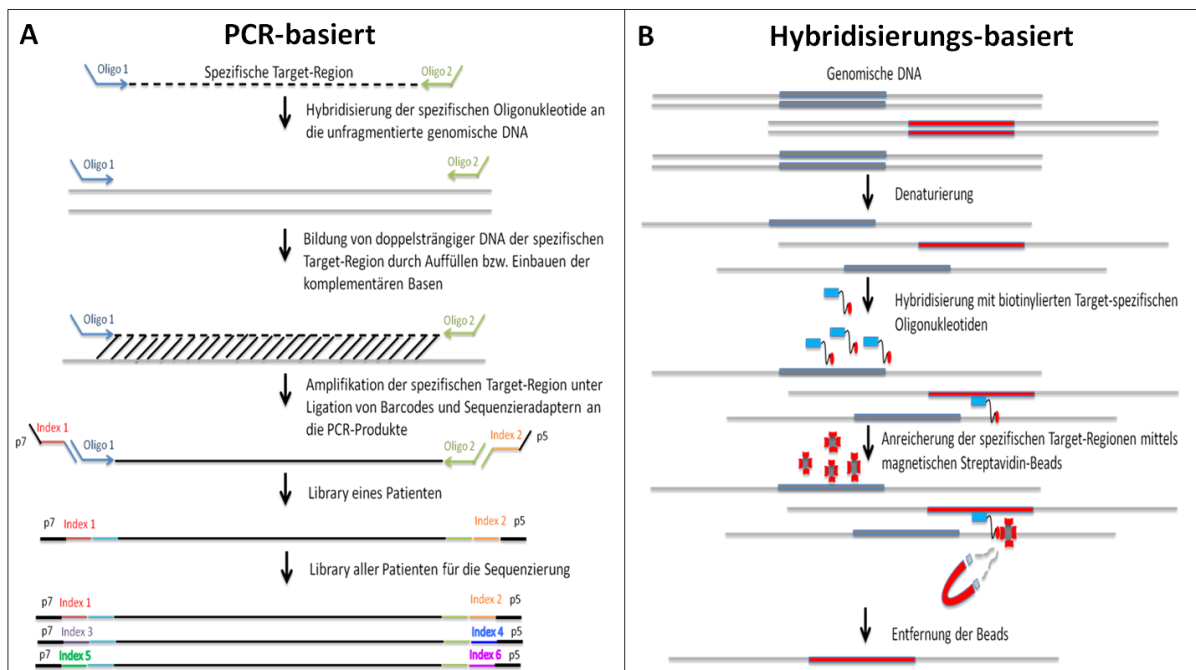


Abbildung 8: Schematische Darstellung der Verfahren zur Anreicherung von Target-spezifischen Regionen

A: PCR-basierte Anreicherung: Spezifische Oligonukleotide binden flankierend an die Target-Region und ein DNA-Doppelstrang wird synthetisiert. Daraufhin werden für jeden einzelnen Patienten Barcodes und Sequenzier-Adapter an die PCR-Produkte ligiert. Abschließend wird ein Gesamtpool aller Patienten erstellt, der als Input zur Hochdurchsatz-Sequenzierung mit dem MiSeq Gerät dient (modifiziert nach Illumina¹⁸⁰). B: Hybridisierungs-basierte Anreicherung: In der Library Präparation wird eine Hybridisierung der DNA-Fragmente mit Target-spezifischen biotinylierten Oligonukleotiden vorgenommen. Mit Hilfe von magnetischen Streptavidin-Beads werden die Target-Regionen aus dem DNA-Pool isoliert, amplifiziert und abschließend mittels NGS auf dem MiSeq sequenziert (modifiziert nach Illumina¹⁷⁸).

Nachfolgend werden die Hochdurchsatz-Sequenzierungs-Daten unabhängig vom Anreicherungsverfahren ausgewertet. Hierbei beginnt jede bioinformatische Auswertung mit dem Demultiplex-Schritt, währenddessen jede Sequenz dem zugehörigen Patienten zugeordnet wird. Durch den Abgleich mit dem Referenzgenom wird abschließend eine Mutationsanalyse durchgeführt (Abbildung 7D). Mittlerweile gibt es kommerziell erhältliche Panel für zahlreiche Krankheiten und Fragestellungen von verschiedensten Firmen. Zusätzlich bieten die meisten Firmen auch das Design von eigenen/ nicht-kommerziellen Panel an (Custom Panel), die es ermöglichen das Mutationsprofil von Proben in einem Set von eigens ausgewählten Genen/Target-Regionen zu analysieren.

1.7 Methylierungsanalysen zur Identifikation von epigenetischen Aberrationen bei Patienten im Senium

Neben genetischen Aberrationen können auch epigenetische Modifikationen zur Entstehung und Progression von Tumoren führen. Unter epigenetischen Aberrationen werden die mitotisch und meiotisch vererbaren Veränderungen in der Genexpression beschrieben, wobei die DNA-Sequenz unverändert vorliegt. Epigenetische Mechanismen sind maßgeblich an der Genregulation sowie der Struktur und Konformation der DNA beteiligt. Zu den wichtigsten epigenetischen Mechanismen der Genregulation zählen die posttranslationalen Histonmodifikationen, RNA-basierte Mechanismen und die postreplikative Methylierung der DNA. Die zwei grundlegenden Mechanismen der Epigenetik, die sowohl die Genexpression als auch die Chromatinstruktur beeinflussen, sind die DNA-Methylierung und die Histon-Modifikationen, die im Folgenden kurz beschrieben werden.

1.7.1 Histonmodifikation

Die DNA liegt innerhalb der Zelle in einer kompakten Form vor. Dabei bilden jeweils 2 Histon-Proteine der Klassen H2A, H2B, H3 und H4 ein Oktamer, dass von der DNA umwickelt wird. Diese Grundeinheit wird auch als Nukleosom bezeichnet und die einzelnen Nukleosomen sind durch Linker-DNA und dem Histon-Protein H1 miteinander verbunden. Der Komplex aus DNA und Histonen stellt das Chromatin dar. Die Enden der Histone (engl. Histone Tails) können durch verschiedene chemische Gruppen (Methyl, Acetyl- und Phosphatgruppen) und verschiedene Grade der DNA-Methylierung (Mono-, Di-, Trimethylierung) modifiziert werden. Mehr als 100 post-translationale Histon-Modifikationen sind bisher bekannt und werden hauptsächlich durch Acetylierung, Methylierung, Phosphorylierung oder Ubiquitinylierung hervorgerufen. Diese Modulationen führen zu Veränderungen in der Chromatinstruktur und beeinflussen die Bindungszugänglichkeit der Transkriptionsfaktoren an die Promotorregionen sowie die damit verbundene Genexpression.

1.7.2 DNA-Methylierung

Histonmodifikationen und DNA-Methylierung sind epigenetische Mechanismen, die stark miteinander interagieren¹⁸¹. Die DNA-Methylierung ist der bisher am besten untersuchte Mechanismus der Epigenetik und kann einzelne Gene oder auch nicht kodierende chromosomale Regionen stilllegen, indem die Bindung der Transkriptionsfaktoren durch die Modifikation der DNA verhindert wird¹⁸². DNA-Methylierung ist eine postreplikativ programmierte Modifikation und tritt beim menschlichen Organismus primär an einem Cytosin in Cytosin-Guanin-Dinukleotiden (CpG) auf, wodurch eine Modifikation am Kohlenstoffatom 5 des Pyrimidinrings der Base Cytosin erfolgt¹⁸³⁻¹⁸⁵. Ein geringer Anteil von unter 5% aller CpGs des Genoms kommen in Clustern vor, wobei diese CpG-Cluster in Satelliten-Regionen, zentromerischen Repeats oder in CG-reichen DNA-Abschnitten (CpG-Inseln) liegen. Im Normalfall liegen die CpG-Inseln in den Housekeeping-Genen in der 5' Regionen von Promotoren, der untranslatierten Regionen (UTR) oder Exon 1 und sind unmethyliert. CpG-Inseln in repetitiven Regionen der Gene sind häufig methyliert^{182,184,186}. Insgesamt sind ca. 75% der Transkriptionsstartpunkte (TSS) und ca. 90% der Promotoren mit CpG-reichen Nukleotidsequenzen assoziiert und unterliegen einer gewebs- und entwicklungsabhängigen Regulation der DNA-Methylierung¹⁸⁷⁻¹⁸⁹.

Der Methylierungsprozess wird von einer Gruppe spezifischer Enzyme, den DNA-Methyltransferasen (DNMT), katalysiert^{182,184}. Die DNMTs nutzen das Substrat S-Adenosyl-L-Methionin (SAM) als Methylgruppendonor, wodurch eine Methylgruppe auf das Kohlenstoffatom 5 des Pyrimidinrings der Base Cytosin in der DNA übertragen wird^{190,191}. Dieses modifizierte Cytosin wird anschließend als 5-Methylcytosin (5mC) bezeichnet. DNMT3A und DNMT3B sind de novo Methyltransferasen und größtenteils für die Modifizierung unmethylierter Cytosine verantwortlich, ein epigenetischer Mechanismus der z.B. in der frühen embryonalen Stammzelle oder bei Tumorzellen eine entscheidende Rolle spielt. Mit Hilfe von DNMT1 wird das Methylierungsmuster nach der Zellteilung wiederhergestellt. Während der Replikation erkennt DNMT1 die hemimethylierte Regionen in der DNA und überträgt das Methylierungsprofil des parentalen Strangs auf den neusynthetisierten Tochterstrang^{184,188,192}. Die Methylierung kann genomweit oder auch genspezifisch sein und wird durch methylbindende Proteine komplettiert.

In den letzten Jahren wurde zudem nachgewiesen, dass neben 5mC auch 5-Hydroxymethylcytosin (5hmC) im menschlichen Organismus auftritt^{193 194}. Die Arbeiten von Tahiliani et al.¹⁹³ und Ito et al.¹⁹⁵ konnte zeigen, dass 5hmC durch die Oxidation von 5mC durch Mitglieder der TET-Familie (TET1-3) katalysiert wird. Vermutlich stellt 5hmC ein Zwischenprodukt der aktiven Demethylierung von 5mC zu Cytosin dar¹⁹⁶.

5hmC kommt in verschiedenen Geweben und Entwicklungsstadien unterschiedlich stark vor, wobei große Mengen an 5hmC im Geweben des Zentralnervensystems nachgewiesen wurden, weshalb davon auszugehen ist, dass 5hmC in Assoziation zu neurobiologischen Erkrankungen steht^{194,197-199}. Zudem konnten genetische Aberrationen in *TET2* bereits in verschiedenen hämatologischen Erkrankungen nachgewiesen werden. Als Erstes konnten Langemeijer et al. Mutationen in der kodierenden Region von *TET2* bei Patienten mit myelodysplastischem Syndrom (MDS) nachweisen²⁰⁰. Daraufhin zeigten Delhommeau et al., dass Mutationen und Deletion in *TET2* bei Patienten mit akuter myeloischer Leukämie (AML), MDS und myeloproliferativen Neoplasien von zentraler Bedeutung sind²⁰¹. Mittlerweile wurden auch Mutationen in chronisch myelomonozytären Leukämien und malignen Lymphomen nachgewiesen²⁰²⁻²⁰⁴. Bisher konnte noch nicht geklärt werden, welche spezifische Funktion *TET2* in der Homöostase des Bluts spielt²⁰⁵. Es ist jedoch klar, dass Mutationen in *TET2* zur malignen Transformation von hämatopoetischen Zellen führen²⁰⁵.

1.7.3 Infinium HumanMethylation450 BeadChip

Am Institut für Humangenetik wurde in den letzten Jahren die Analyse der DNA-Methylierung mittels Infinium HumanMethylation450 Bead Chip (450K) etabliert. Unter Verwendung des 450K ist es möglich zeitgleich das Methylierungsmuster von 485.577 Loci in 21.231 Genen zu bestimmen, wobei parallel bis zu 12 Proben auf einem 450K analysiert werden können. Die CpG-Loci des 450K decken Regionen innerhalb von 200 bis 1500 bp vor dem Transkriptionsstartpunkt (TSS), der 5'-UTR, dem Exon 1, dem „Genkörper“ (TSS + 500 bp bis zum 3'-Ende des Transkriptes) und der 3'-UTR ab. 65 der auf dem 450K lokalisierten Loci repräsentieren Einzelnukleotidpolymorphismen (SNP) und sind mit „rs“ gekennzeichnet²⁰⁶. 3.091 Loci sind nicht-CpG-Loci (sog. ch-Loci) und liegen außerhalb von CpG-Inseln. Die ch-Loci liegen in Shores (2 kb Regionen, die die CpG-Inseln umgeben), Shelves (2 kb Regionen, die die Shores umgeben) und über das Genom verteilt (Open Sea)²⁰⁷. Alle weiteren 482.241 Loci sind CpG-Loci und werden mit „cg“ benannt.

Um die Methylierung der DNA als Sequenzunterschied deutlich zu machen und mittels 450K bestimmen zu können, musste zunächst eine Natriumbisulfit-Behandlung der genomischen DNA erfolgen (Abbildung 9). Dabei wurden unmethylierte Cytosine (C) zu Uracil umgewandelt und methylierte Cytosine gingen unverändert aus der Reaktion hervor²⁰⁸. Nach der Konvertierung folgte eine Amplifikationsreaktion in der Uracil (U) durch Thymin (T) ersetzt wurde, sodass bei der anschließenden Analyse unmethylierte Cytosine als Thymin detektiert wurden. Die Bisulfitkonvertierung mit anschließender Amplifikation hat zur Folge, dass zwei nicht mehr komplementäre DNA-Stränge am Ende der Reaktion vorliegen.

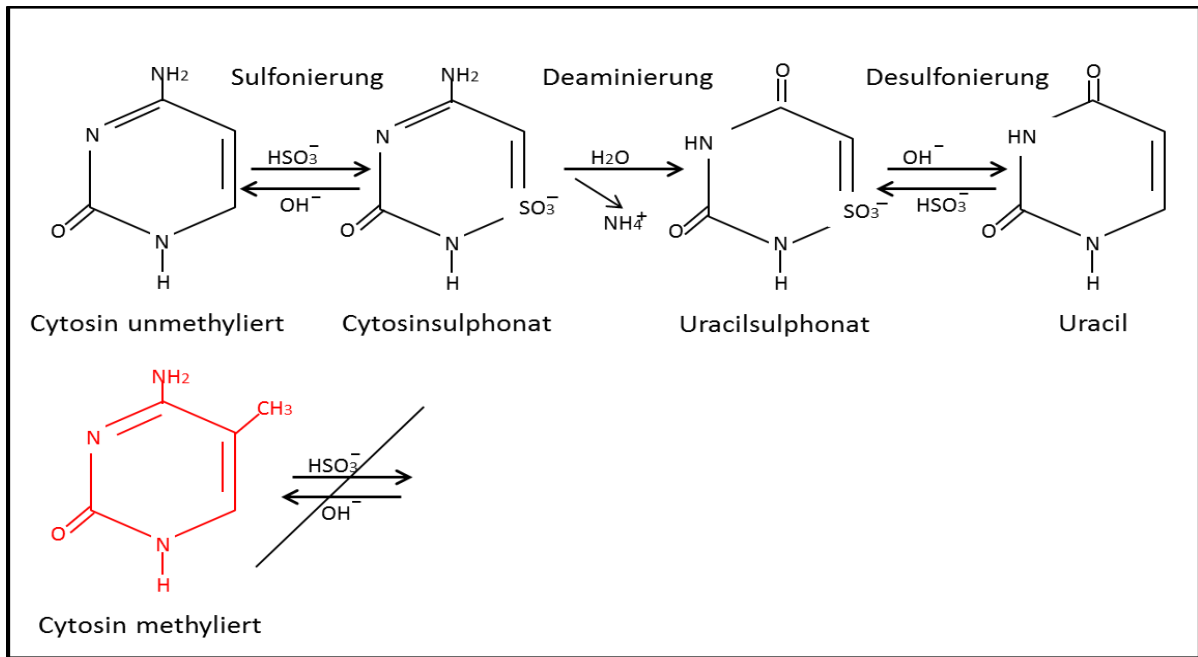


Abbildung 9: Schematische Darstellung der Natriumbisulfit-Konvertierung von genomischer DNA

In der Konvertierung der DNA durch die Behandlung mit Natriumbisulfit werden nicht methylierte Cytosin unter der Bildung verschiedener Zwischenprodukte zu Uracil umgewandelt. Cytosine bei denen eine Methylgruppe am fünften Kohlenstoffatom vorliegt, werden unter den gewählten Reaktionsbedingungen nicht zu Uracil konvertiert. (modifiziert nach Carew et al., 1985²⁰⁸)

Das Prinzip des Arrays basiert auf zwei verschiedenen Typen von Beads: Infinium I Assay und Infinium II Assay. Der Infinium I Assay besitzt für die Detektion jedes zu analysierenden CpG-Loci zwei Beadtypen: einen methylierten und einen unmethylierten Beadtyp (Abbildung 10A). Diese Beadtypen besitzen Sonden, die am 3'-Ende der Sondensequenz entweder komplementär an den methylierten oder unmethylierten CpG-Lokus binden. In der Probenvorbereitung mit dem 450K bindet die konvertierte DNA entsprechend dem Methylierungsstatus der Sequenz an den methylierten oder unmethylierten Beadtypen. Anschließend wird die Sonde komplementär zum gebundenen konvertierten DNA-Fragmente, um ein markiertes Nukleotid, aufgefüllt. Der Infinium II Assay hingegen besteht nur aus einem Beadtyp, der es ermöglicht mit einer Sonde sowohl ein methyliertes als auch ein unmethyliertes CpG Dinukleotid zu identifizieren (Abbildung 10B)^{206,209}. So wird bei einem methylierten CpG-Loci die komplementäre Base Guanin in die Nukleotidsequenz eingebaut während ein unmethylierter CpG-Lokus mit der Base Adenin aufgefüllt wird. Der Methylierungsstatus der analysierten CpG-Loci wird abschließend aus den spezifischen Fluoreszenzsignalen ermittelt. Die markierten Nukleinbasen des Infinium I und II Assay Designs werden dann von Fluoreszenz-markierten Antikörpern gebunden und die Fluoreszenzintensität für beide Beadtypen mit dem IScan Gerät der Firma Illumina bestimmt. Der Methylierungsgrad wird abschließend aus dem Verhältnis der Intensität des methylierten Allels zu der Summe der Intensitäten des methylierten und unmethylierten Allels berechnet und als AVG-Beta-Wert bezeichnet. Der berechnete AVG-Beta-Wert liegt dabei zwischen 0 (0% Methylierung) und 1 (100% Methylierung).

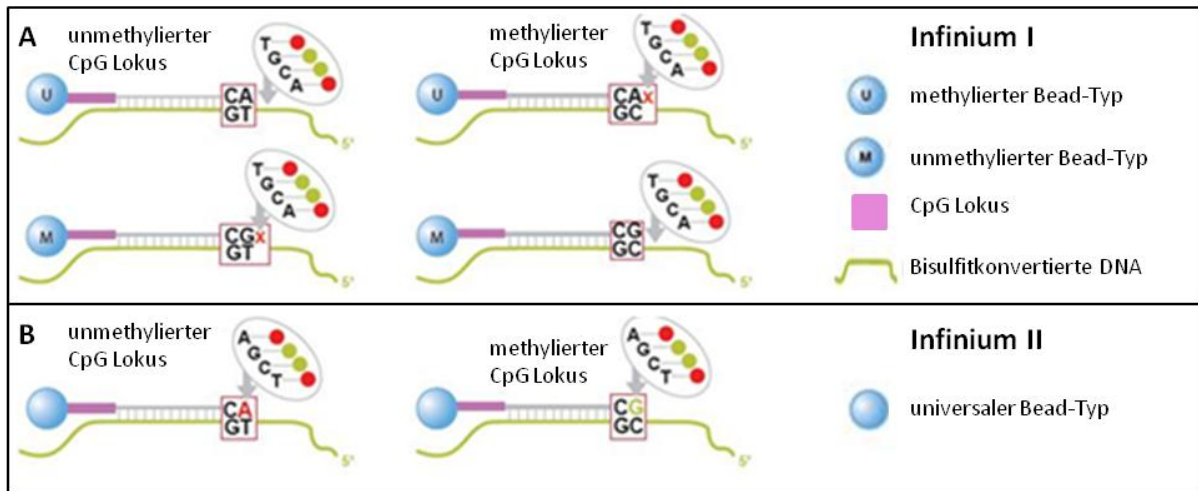


Abbildung 10: Schematische Darstellung der Assays auf dem Infinium HumanMethylation450 BeadChIP
 A: Das Infinium I Assay Design. Für jeden zu analysierenden CpG-Loci sind zwei Beadtypen auf dem Assay (methyliert und unmethyliert). U stellt in der Grafik den unmethylierten und M den methylierten Beadtypen dar (modifiziert nach²¹⁰) B: Das Infinium II Assay Design basiert auf einem universellen Beadtyp, der sowohl den methylierte als auch unmethylierten Zustand von CpG-Loci detektiert (modifiziert nach²¹⁰). Abkürzungen: C: Cytosin. G: Guanin. T: Thymin. A: Adenin.

1.8 Ziele der Arbeit

Für die Analysen der genetischen Aberrationen war es notwendig zuerst die Methode der Hochdurchsatz-Sequenzierung inkl. der Probenvorbereitung am Institut für Humangenetik zu optimieren. Anschließend sollten altersassoziierten genetische und epigenetische Aberrationen in reifzelligen Keimzentrums-B-Zell-Lymphomen identifiziert werden. Das Ziel der vorliegenden Arbeit beinhaltet somit zwei komplexe Aufgabenstellungen:

1.) Methodik – Austestung und Optimierung neuer methodischer Ansätze zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung

In Rahmen der Methodenoptimierung sollten folgende Fragestellungen beantwortet werden:

- Welche Parameter sind von zentraler Bedeutung bei der Probenvorbereitung bzw. Hochdurchsatz-Sequenzierung und wie können diese Parameter angepasst werden, um die Sequenzier-Ergebnisse zu verbessern?
- Ist es möglich FFPE-Material mittels Hochdurchsatz-Sequenzierung zu sequenzieren bzw. was ist bei der Analyse von FFPE-Material zu beachten?
- Welchen Einfluss hat die eingesetzte Polymerase in der Probenvorbereitung auf die abschließenden Sequenzier-Ergebnisse?

2.) Identifizierung von altersassoziierten Aberrationen bei Keimzentrums-B-Zell-Lymphomen

a. Genetische Aberrationen bei sehr alten Patienten

Gibt es Unterschiede zwischen jüngeren und sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom:

- In der Frequenz von chromosomalen Imbalancen der Gene *BCL2*, *BCL6*, *MYC* und *IGH* ?
- In der Expression von *BCL2*, *BCL6*, *CD10*, *Ki67* und *MUM1* ?
- In dem Mutationsprofil von Genen, die häufig in Lymphomen mutiert auftreten ?

In dieser Arbeit sollte gezielt eine neue Kohorte von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom (Ersterkrankungsalter ≥ 80 Jahre) zusammengestellt werden und diese Kohorte molekular charakterisiert werden. Dazu sollten chromosomale Imbalancen mittels Fluoreszenz-in-situ-Hybridisierung (FISH) und die Expression von Proteinen mittels Immunhistochemie (IHC) analysiert werden. Zudem sollte eine Mutationsanalyse von sehr alten Patienten mittels Target-Resequenzierung mit einem Lymphom-Panel unter Verwendung des MiSeq-Geräts der Firma Illumina erfolgen. Abschließend sollten die neu erhobenen molekulargenetischen Daten mit bereits existierenden Daten von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus verschiedenen Studien (NHL-B, RICOVER, MMML und ICGC MMML-Seq) verglichen werden. Das Ziel liegt hierbei in der Identifizierung von genetischen Aberrationen bzw. Markern, die vorwiegend in Patienten mit einem sehr hohen Ersterkrankungsalter identifiziert werden.

b. Epigenetische Aberrationen bei sehr alten Patienten

Gibt es Unterschiede in dem DNA-Methylierungsmuster zwischen:

- Verschiedenen Keimzentrums-B-Zell-Lymphom-Entitäten ?
- Jüngeren und sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom ?

Weiterhin sollten in der vorliegenden Arbeit von einem Set von sehr alten Patienten das DNA-Methylierungsmuster mittels Infinium HumanMethylation450 BeadChips erhoben werden. Anschließend sollten die DNA-Methylierungs-Daten mit bereits existierenden 450K-Daten von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den Forschungsstudien MMML und ICGC MMML-Seq verglichen werden. Das Ziel liegt hierbei in der Identifizierung von epigenetischen Modifikationen der DNA-Methylierung, die vorwiegend bei Patienten mit einem sehr hohen Ersterkrankungsalter auftreten.

2 Material und Methoden

In diesem Teil werden die im Rahmen der Arbeit verwendeten Materialien aufgelistet und anschließend die angewandten Methoden beschrieben. Allgemein übliche Labormethoden, Arbeitsanweisungen, Reagenzien und Geräte werden nicht separat angegeben. Aufgrund der Vielzahl von Fachbegriffen sind diese zum Nachschlagen separat in einem Glossar am Anfang der Arbeit zusammengetragen. Weiterhin sind einzelne Tabellen und Abbildungen im Anhang zu finden, wobei diese im Text durch ein vorgeschaltetes „A“ wie Anhang gekennzeichnet sind (z.B. Tabelle A1, erste Tabelle im Anhang). Es wird darauf hingewiesen, dass molekulargenetische Untersuchungen im Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel nach den Bestimmungen der DIN EN ISO 15189:2003 bzw. 2014 durchgeführt wurden und die entsprechenden Protokolle der Standard-Arbeitsanweisungen (SAA) hier übernommen sind. Sofern in der vorliegenden Arbeit keine weiteren Informationen zu den genannten Instituten angegeben wurden, gehören diese Institute zum Universitätsklinikum Schleswig-Holstein (UKSH), Campus Kiel.

2.1 Material

Für die Identifizierung genetischer Aberrationen- insbesondere Gen-Mutationen in Keimzentrums-B-Zell-Lymphomen mittels Hochdurchsatz-Sequenzierung- war es notwendig diese Technik am Institut für Humangenetik zu optimieren. Für die Probenvorbereitung, die Voraussetzung für die Hochdurchsatz-Sequenzierung ist, wurden verschiedene Verfahren getestet. Die im Rahmen der gesamten Methoden-Optimierung verwendeten und analysierten Proben sind im nachfolgenden Abschnitt 2.1.1 angegeben. Der Fokus dieser Arbeit lag in der altersassoziierten Analyse von genetischen und epigenetischen Aberrationen bei Patienten mit einem Keimzentrums-B-Zell-Lymphom. Hierzu wurden Patienten mit einem Alter bei Diagnose von 3 bis 97 Jahren betrachtet und in verschiedene Altersgruppen eingeteilt. Zuerst wurde eine Kohorte von Patienten mit einem Alter von mind. 80 Jahren zusammengestellt und Fluoreszenz-in-situ-Hybridisierungs (FISH)-, Immunhistochemie (IHC)-, Mutations- sowie DNA-Methylierungsdaten von den Tumoren dieser Patienten erhoben. Anschließend wurden bereits existierende molekulargenetische Daten von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus klinischen und Forschungsstudien zusammengestellt und verschiedene altersassoziierte Analysen von genetischen und epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen durchgeführt. Diese Patienten-Kohorten und Materialien sind im Abschnitt 2.1.2 zusammengestellt.

2.1.1 Probenmaterial für die Optimierung der Hochdurchsatz-Sequenzierung

Zur Optimierung der Hochdurchsatz-Sequenzierung wurden verschiedene Anreicherungsverfahren mit unterschiedlichen Probenmaterialien verwendet. Zuerst wurde das PCR-basierte, kommerziell erhältliche *TruSight Amplicon Panel* (Myeloid-Panel) getestet (siehe Abschnitt 2.2.4.1), womit Patienten mit einer T-Prolymphozyten-Leukämie (T-PLL, siehe Abschnitt 2.1.1.1) analysiert wurden. Daraufhin wurde der eigens entworfene *Two-Step-Assay* (*ID3*-Assay, siehe Abschnitt 2.2.4.1), unter Verwendung von extrahierter DNA aus dem Frischmaterial von Zelllinien (siehe Abschnitt 2.1.1.2) und dem FFPE-Material von nicht-Tumor-Proben aus dem Untersuchungskollektiv des Instituts für Pathologie/ Sektion Hämatopathologie (siehe Abschnitt 2.1.1.3), evaluiert. Es folgt die Optimierung des *TruSight Rapid Capture Panels* (Cancer-Panel), einem Hybridisierungs-basierten und kommerziell erhältlichen Panel (siehe Abschnitt 2.2.4.2). Hierzu wurde die extrahierte DNA aus dem peripheren Blut von Patienten aus den Untersuchungskollektiven des Instituts für Humangenetik und der Klinik für Gynäkologie und Geburtshilfe eingesetzt (siehe Abschnitt 2.1.1.4). Abschließend wurde das eigens entworfene Hybridisierungs-basierte *SureSelect XT Panel* (Lymphom-Panel, siehe Abschnitt 2.2.4.2) verwendet, womit sehr alte Patienten mit einem Keimzentrums-B-Zell-Lymphom analysiert wurden (siehe Abschnitt 2.1.2.1 und 2.1.2.2). Alle Proben, die im Rahmen der Methodik in der vorliegenden Arbeit verwendet wurden, sind in den nachfolgenden Kapiteln kurz zusammengestellt.

2.1.1.1 Patienten mit einer T-Prolymphozyten-Leukämie

Im Rahmen der Optimierung der Hochdurchsatz-Sequenzierung wurden 39 Patienten mit einer T-Prolymphozyten-Leukämie (T-PLL) untersucht. Die Tumor-Proben der Patienten, die innerhalb der Optimierung des *TruSight Amplicon Panel* (Myeloid-Panel) analysiert wurden, sind in dem dazugehörigen Manuskript aufgeführt und detaillierte Patienten-Informationen dort nachzulesen (Lopez et al., Genes encoding members of the JAK-STAT pathway or epigenetic regulators are recurrently mutated in T-cell prolymphocytic leukaemia, *Br. J. Haematol.*, 2016).

2.1.1.2 Keimzentrums-B-Zell-Lymphom-Zelllinien

Zur Optimierung des *Two-Step-Assays* (*ID3*-Assay) wurden Keimzentrums-B-Zell-Lymphom-Zelllinien mit bekannten Mutationen im *ID3*-Gen aus der Publikation von Richter et al., 2012 verwendet. Alle Zelllinien wurden ursprünglich von der Deutschen Sammlung von Mikroorganismen und Zellkulturen (DSMZ) vom Leibniz Institut in Braunschweig bezogen und sind mittels Mikrosatellitenanalyse (STR-Profilung) am Institut für Humangenetik auf die korrekte Identität überprüft worden. Die DNA der Zelllinien wurde bereits zuvor von Mitarbeitern am Institut für Humangenetik isoliert und bei 4°C gelagert. Für die Etablierung des *ID3*-Assays wurden die DNA-Stocks der Keimzentrums-B-Zell-Lymphom-Zelllinien zur Verfügung gestellt (siehe Tabelle 3).

Tabelle 3: Charakteristika von Keimzentrums-B-Zell-Lymphom-Zelllinien

| Zelllinie | Entität | Alter | Geschlecht | Translokation | ID3-Status | Referenz |
|------------|------------------------|-------|------------|---------------------------------------|------------|----------------------------------------|
| BL-2 | BL | 7 | m | t(8;22)(q24;q11) | ja | Bertrand et al., 1981 ²¹¹ |
| BL-41 | BL | 8 | m | t(8;14)(q24;q32) | ja | Lenoir et al., 1985 ²¹² |
| BL-70 | BL | 16 | m | t(8;14)(q24;q32) | ja | Lenoir et al., 1985 ²¹² |
| Blue-1 | BL | 29 | m | t(8;14)(q24;q32) | ja | Burmeister et al., 2006 ²¹³ |
| CA-46 | BL | 16 | m | t(8;14)(q24;q32) | ja | Magrath et al., 1980 ²¹⁴ |
| Daudi | BL | 16 | m | t(8;14)(q24;q32) | ja | Nadkarni et al., 1969 ²¹⁵ |
| DG-75 | BL | 10 | m | t(8;14)(q24;q32) | nein | Ben-Bassat et al., 1977 ²¹⁶ |
| EB-1 | BL | 9 | w | t(8;14)(q24;q32)* | nein | Epstein et al., 1964 ²¹⁷ |
| Namalwa | BL | 3 | w | t(8;14)(q24;q32)* | ja | Nadkarni et al., 1969 ²¹⁵ |
| Raji | BL | 12 | m | t(8;14)(q24;q32) | nein | Pulvertaft et al., 1964 ²¹⁸ |
| Ramos | BL | 3 | m | t(8;14)(q24;q32) | nein | nicht publiziert |
| SUDHL-5 | DLBCL-GCB [#] | 17 | w | +12, del(6)(q13) | nein | Epstein et al., 1978 ²¹⁹ |
| SUDHL-6 | DLBCL-GCB [#] | 43 | m | t(14;18)(q32;q21) | nein | Epstein et al., 1978 ²¹⁹ |
| SUDHL-10 | DLBCL-GCB [#] | 25 | m | t(8;14)(q24;q32) t(14;18)(q32;q21) | nein | Epstein et al., 1978 ²¹⁹ |
| U-698 | BL | 7 | m | t(8;22)(q24;q32) | ja | Nilsson et al., 1974 ²²⁰ |
| HT | DLBCL-GCB [#] | 70 | m | t(14;18)(q32;q21) | nein | Beckwith et al., 1990 ²²¹ |
| MC-116 | B-Zell-Lymphom | NA | m | t(8;14)(q24;q32) | ja | Dyer et al., 1990 ^{222,209} |
| Karpas-422 | DLBCL-GCB [#] | 72 | w | t(14;18)(q32;q21) | nein | Magrath et al., 1980 ²¹⁴ |

Abkürzungen und Verweise: ID3-Status: Aus der Veröffentlichung von Richter et al., 2012 ist bekannt, dass die Zelllinie eine Mutation im ID3-Gen aufweist bzw. nicht aufweist. m: männlich. w: weiblich. Alter: Patientenalter in Jahren bei Probenentnahme, die zur Etablierung der Zelllinie verwendet wurde. DLBCL-GCB: Diffus großzelliges B-Zell-Lymphom von Keimzentrumsähnlichen B-Zellen. BL: Burkitt-Lymphom. Referenz: Erste Beschreibung der Zelllinie. NA: (engl. not available) nicht verfügbar. #: Schuetz et al., 2012²²³. *: Murga et al., 2014²²⁴. Weitere Informationen zu diesen Zelllinien sind auf der Internetseite der DSMZ unter <https://www.dsmz.de/home.html> zu finden.

2.1.1.3 FFPE-Material von Nicht-Tumor-Proben

Zur Optimierung der Hochdurchsatz-Sequenzierung wurde extrahierte DNA aus Formalin-fixiertem Paraffin-eingebetteten Gewebe (FFPE)-Material benötigt, um den *Two-Step-Assay* (ID3-Assay) für die Verwendung von FFPE-Proben zu testen. Hierzu wurden vom Institut für Pathologie/ Sektion Hämatopathologie die Schnitte von drei FFPE-Proben angefertigt und zur Verfügung gestellt. Für diese FFPE-Proben waren zuvor die Diagnosen Lymphadenitis (N=2) und follikuläre Hyperplasie (N=1) gestellt worden, d.h. es waren nicht-Tumor-Gewebeproben mit reaktiven Veränderungen der entsprechenden Lymphknoten. Die FFPE-Proben wurden der Autorin vollständig anonymisiert (auch hinsichtlich Alter und Geschlecht) zur Verfügung gestellt.

2.1.1.4 Extrahierte DNA aus peripherem Blut

Zur Probenvorbereitung mittels dem *TruSight Rapid Capture Panel* (Cancer-Panel) wurden insgesamt 28 Proben verwendet. Dieses Kollektiv setzte sich zusammen aus vier Patienten des Untersuchungskollektives der Klinik für Gynäkologie und Geburtshilfe, zu denen bereits bekannte Mutationen in den Target-Genen *BRCA1* oder *BRCA2* aus der Mutationsanalyse im Rahmen der Brustkrebsvorsorgeuntersuchung vorlagen. Der *BRCA1/2*-Mutationsstatus wurde freundlicherweise zusammen mit der extrahierten DNA von der Klinik für Gynäkologie und Geburtshilfe übermittelt.

Fünf weitere Patienten des Untersuchungskollektives des Instituts für Humangenetik mit bereits bekannten Mutationen in den Target-Genen *TP53*, *NF1*, *KIT* oder *APC*, die zuvor im Rahmen der Routine-Diagnostik des Instituts für Humangenetik nachgewiesen wurden und 18 Proben von Indexpatienten aus dem Untersuchungskollektives des Instituts für Humangenetik mit einem unbekanntem Mutationsstatus in den zu analysierenden Target-Regionen wurden ebenfalls analysiert. Zusätzlich wurde eine Probe von einem gesunden Individuum (Normalkontrolle) mit dem *TruSight Rapid Capture Panel* angereichert und sequenziert. Alle Patienten, die im Rahmen der vorliegenden Arbeit unter Verwendung des *TruSight Rapid Capture Panel* analysiert wurden, sind in der nachfolgenden Tabelle 4 aufgelistet.

Tabelle 4: Analyierte Patienten und Mutationsstatus zur Optimierung des *TruSight Rapid Capture Panels*

| Patient | Gen | Transkript | Mutationsstatus vor der Austestung des <i>TruSight Rapid Capture Panels</i> |
|----------|---------------|--------------|-----------------------------------------------------------------------------|
| D12-819 | <i>TP53</i> | NM-000546.5 | c.814G>A, p.Val272Met |
| D14-397 | <i>NF1</i> | NM-001042492 | c.3831C>T, p.Gly1277Gly |
| D09-2325 | <i>KIT</i> | NM-000222.2 | c.2447A>T, p.Asp816Val |
| D11-2050 | <i>KIT</i> | NM-000222.2 | c.2446G>T, p.Asp816Tyr |
| R41-9 | <i>BRCA2</i> | NM-000059.3 | c.3G>A, p.Met1? |
| R46-24 | <i>BRCA2</i> | NM-000059.3 | c.9474delA, p.Phe3159Serfs*4 |
| R42-36 | <i>BRCA1</i> | NM-007300.3 | c.3910delG, p.Glu1304Lysfs*3 |
| R44-63 | <i>BRCA1</i> | NM-007300.3 | c.5176-5181delCTAGGA, p.Leu1726-Gly1727delinsdel |
| D14-498 | <i>APC</i> | NM-001127510 | c.4784delC, p.Gln1596Argfs*54 |
| D14-313 | <i>CHECK2</i> | NM-007194 | unbekannt |
| D14-700 | <i>CHECK2</i> | NM-007194.3 | unbekannt |
| D14-314 | <i>BRCA1</i> | NM-007294.3 | unbekannt |
| D14-726 | <i>BRCA1</i> | NM-007294.3 | unbekannt |
| D13-1852 | - | - | unbekannt |
| D13-1037 | - | - | unbekannt |
| D13-1660 | - | - | unbekannt |
| D14-1170 | - | - | unbekannt |
| D14-1171 | - | - | unbekannt |
| D14-1224 | - | - | unbekannt |
| D14-1225 | - | - | unbekannt |
| D14-1077 | - | - | unbekannt |
| D14-1165 | - | - | unbekannt |
| D14-169 | - | - | unbekannt |
| D14-637 | - | - | unbekannt |
| D14-689 | - | - | unbekannt |
| D14-733 | - | - | unbekannt |
| D14-838 | - | - | unbekannt |
| NK | - | - | unbekannt |

Abkürzungen und Verweise: NK: Normalkontrolle. Gen: In dem diesem Gen wurde eine potenziell pathogene Mutation in der analysierten Probe identifiziert. Mutation: Nomenklatur nach HGVS unter Verwendung des humanen Referenzgenoms hg19.

2.1.2 Patienten-Kohorten für die altersassoziierten Analysen an Keimzentrums-B-Zell-Lymphomen

In der vorliegenden Arbeit erfolgte die Analyse von genetischen und epigenetischen Aberrationen bei Patienten mit einem Keimzentrums-B-Zell-Lymphom. Die durchgeführten Untersuchungen wurden unter der Leitung von Frau Dr. Julia Richter im Rahmen der Förderung einer *Junior Excellence Research Group* durch die Dr. Werner Jackstädt Stiftung mit dem Titel: „Mechanismen der B-Zell-Lymphomentstehung im Senium als Grundlage für die Entwicklung altersangepasster Therapiestrategien“, für die ein positives Ethikvotum der Medizinischen Fakultät der Christian-Albrechts-Universität zu Kiel vorlag (AZ.: D 429/13), vorgenommen. Im Rahmen dieser Studie und der hier vorliegenden Arbeit werden Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem Ersterkrankungsalter von mind. 80 Jahren als senile Patienten bezeichnet. Hierbei ist insbesondere darauf zu achten, dass der Begriff „senil“ in der vorliegenden Arbeit ausschließlich auf das Ersterkrankungsalter der Patienten von mind. 80 Jahren basiert und weder mit geistiger noch körperlicher Gebrechlichkeit bzw. anderen „Altersschwächen“ der sehr alten Patienten assoziiert wird. Zum Vergleich der Altersassoziation molekulargenetischer und epigenetischer Aberrationen bei Keimzentrums-B-Zell-Lymphomen konnte auf die bereits erhobenen Daten aus verschiedenen klinischen sowie Forschungsstudien zurückgegriffen werden. Eine Übersicht der analysierten Patienten-Kohorten in Bezug auf die einzelnen Studien ist in der nachfolgenden Abbildung 11 dargestellt und ist detaillierter im Anhang in der Abbildung A1 zu finden.

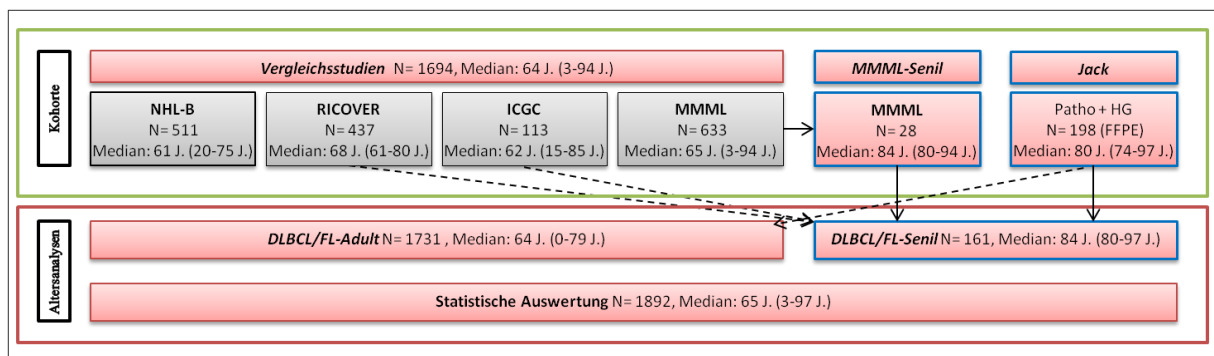


Abbildung 11: Schematische Darstellung der analysierten DLBCL- und FL-Proben

In der grafischen Darstellung der analysierten Patienten-Kohorten wurde zur Vereinfachung die Kohorte *Vergleichsstudien* inkl. der senilen Patienten aus den entsprechenden klinischen und Forschungsstudien dargestellt. Im Allgemeinen wird der Begriff *Vergleichsstudien* in dieser Arbeit für alle nicht senilen Patienten (<80 Jahre) der klinischen und Forschungsstudien verwendet. Im Rahmen der vorliegenden Arbeit wurden genetische und epigenetische Analysen von Keimzentrums-B-Zell-Lymphomen durchgeführt. Hierzu wurde eine Kohorte von 198 Patienten (FFPE-Material) mit einem Ersterkrankungsalter ≥ 74 Jahre aus dem Untersuchungsmaterial des Instituts für Humangenetik und dem Institut für Pathologie/ Sektion Hämatopathologie neu identifiziert und charakterisiert (*Jack*). Zusätzlich konnte die bereits extrahierte DNA von 28 Tumorproben (Frischmaterial) von senilen Patienten aus dem MMML-Verbundprojekt analysiert werden (*MMML-Senil*). Zum Vergleich der Altersassoziation von genetischen und epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen konnten bereits existierende Daten von 1694 Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den klinischen (NHL-B, RICOVER) sowie Forschungsstudien (MMML, ICGC) genutzt werden. Die senilen Patienten aus allen analysierten Kohorten (≥ 80 Jahre) wurden in den Analysen als eine Kohorte zusammengefasst und als *DLBCL/FL-Senil* bezeichnet (N= 161) und im Vergleich zu der Kohorte *DLBCL/FL-Adult* (<80 Jahre, N= 1731) ausgewertet. Abkürzungen: J: Jahre. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. RICOVER und NHL-B: Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom

2.1.2.1 Vergleichskohorte von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus verschiedenen klinischen und Forschungsstudien

Für die Identifizierung der altersassoziierten genetischen und epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden verschiedene Studien zu Vergleichszwecken genutzt. Die Zusammensetzungen der einzelnen Teil-Kohorten von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den verschiedenen klinischen und Forschungsstudien (DLBCL/FL-Patienten bzw. Nicht DLBCL/FL-Patienten) werden nachfolgend genauer beschrieben. Im Rahmen dieser Arbeit wurde insgesamt eine Vergleichskohorte von 1694 Patienten im Alter von 3 bis 94 Jahren (Median: 64 Jahre) mit einer Diagnose DLBCL oder FL nach WHO-Klassifikation 2008 aus verschiedenen klinischen und Forschungsstudien analysiert. Die Vergleichskohorte beinhaltet die bereits publizierten Daten von DLBCL- oder FL-Patienten aus den klinischen Studien NHL-B (N= 511) und RICOVER (N= 437) sowie den Forschungsverbunden MMML (N= 633, inkl. *MMML-Senil*) und ICGC MMML-Seq (N= 113)^{116,117,172,225-227}. Eine Übersicht der analysierten Patienten und Entitäten in Bezug auf die einzelnen Studien ist in der nachfolgenden Abbildung 12 dargestellt.

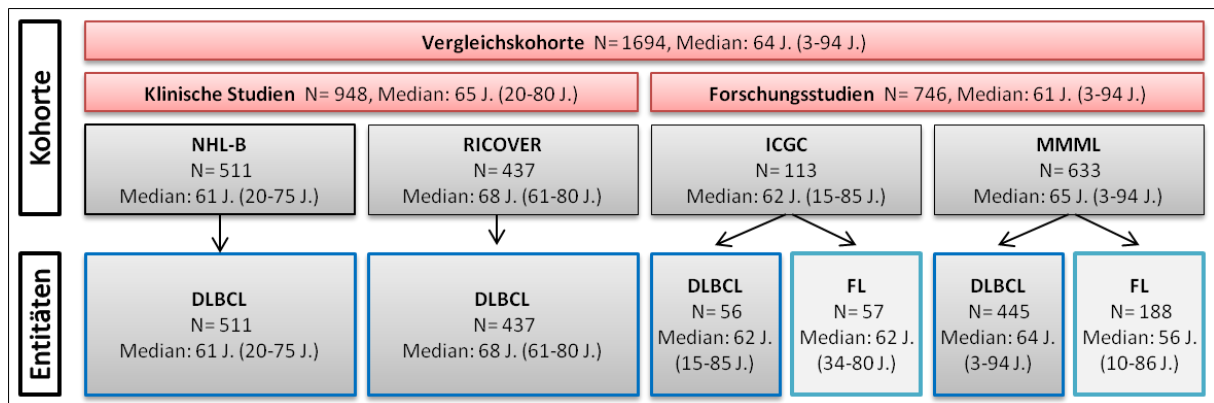


Abbildung 12: Schematische Darstellung der analysierten DLBCL- und FL-Patienten der Vergleichskohorte

In die altersassoziierten Analysen der vorliegenden Arbeit wurden bereits erhobene Daten von Patienten mit einem DLBCL oder FL aus verschiedenen vorangegangenen Studien integriert. Hierzu lagen Daten von DLBCL-Patienten aus den klinischen Studien NHL-B und RICOVER sowie aus den Forschungsstudien ICGC MMML-Seq und MMML vor. Aus den Forschungsstudien existierten zusätzlich die Patienteninformationen und Ergebnisse von molekulargenetischen Analysen zu Fällen mit einem diagnostizierten FL. Abkürzungen: DLBCL: Diffus großzelliges B-Zell-Lymphom. FL: Follikuläres Lymphom. J: Jahre. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. RICOVER und NHL-B: Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom.

Für die statistische Auswertung der altersassoziierten Analysen wurden folgende molekulargenetischen Daten von den DLBCL- und FL-Patienten der Vergleichskohorte der verschiedenen Studien zur Verfügung gestellt:

- IHC-Daten/ Genexpression der analysierten Antigene gegen BCL2, BCL6, CD10, Ki67 und MUM1/IRF4 von allen analysierten Patienten der Kohorte Vergleichskohorte der verschiedenen Studien.
- FISH-Daten zu chromosomalen Imbalancen der Gene *BCL2*, *BCL6*, *IGH* und *MYC* von allen analysierten Patienten der Studien RICOVER, MMML und ICGC MMML-Seq. Für alle analysierten Patienten der klinische Studie NHL-B lagen nur FISH-Daten von dem *MYC*-Gen vor.
- Weitere analysierte Patienten-Charakteristika: Geschlecht, DLBCL-Subtyp und *Hans-Klassifikator* (nonGCB/GCB) von allen analysierten Patienten der Vergleichskohorte der verschiedenen Studien. Zusätzlich lag für 344 Patienten aus dem MMML-Verbundprojekt und von 54 Patienten des ICGC MMML-Seq Projektes der *GEP-Klassifikator* (ABC/GCB) vor.
- Hochdurchsatz-Sequenzierungs-Daten (WGS) zu den analysierten Target-Regionen des *inhouse* erstellten *SureSelect XT Panel* (Lymphom-Panel) von 93 Patienten aus der Studie ICGC MMML-Seq.
- DNA-Methylierungsdaten der mittels Infinium HumanMethylation 450K Array (450K) analysierten CpG-Loci von 13 Patienten aus dem MMML-Verbundprojekt und von 94 Patienten des ICGC MMML-Seq Projektes.

2.1.2.1.1 Vergleichskohorte von nicht senilen DLBCL- und FL-Patienten für die genetischen und epigenetischen Analysen (*Vergleichsstudien*)

Die Vergleichskohorte der verschiedenen Studien sollte insbesondere das Altersspektrum von Patienten mit einem Keimzentrums-B-Zell-Lymphom unter 80 Jahren abdecken. Insgesamt wurden in der vorliegenden Arbeit die Daten von 1633 nicht senilen Patienten (Ersterkrankungsalter <80 Jahre) mit einem DLBCL oder FL aus der Vergleichskohorte der Studien ausgewertet (NHL-B: N= 511, RICOVER: N= 432, MMML: N= 581 und ICGC MMML-Seq: N= 109). Diese Kohorte von nicht senilen Patienten wird im Weiteren als *Vergleichsstudien* bezeichnet. Die molekulargenetischen Daten der 61 Patienten mit einem Ersterkrankungsalter von mind. 80 Jahren wurden separat betrachtet und werden im Abschnitt 2.1.2.1.3 genauer beschrieben. Insgesamt ist zu beachten, dass die Patienten-Kohorte des Verbundprojektes MMML aufgrund der damaligen Fragestellung gezielt für Kinder und junge Patienten (DLBCL insbesondere vom GCB-Subtyp) angereichert wurde.

2.1.2.1.2 Vergleichskohorte von Tumor-Proben (Nicht DLBCL/FL-Proben) für die DNA-Methylierungsanalysen

Zur Analyse von altersassoziierten Aberrationen des DNA-Methylierungsprofils von Keimzentrums-B-Zell-Lymphomen wurden zusätzlich Proben von anderen Lymphom-Entitäten (z.B. BL) analysiert, um mögliche altersassoziierte epigenetischen Aberrationen in DLBCL und FL klar von den Veränderungen in anderen Lymphom-Entitäten abzugrenzen. Hierzu wurden die 450K-Methylierungsdaten von Tumor-Proben genutzt, die bereits am Institut für Humangenetik vorlagen. Die Auswahl der Tumor-Proben erfolgte unter folgenden Selektionskriterien:

- Hybridisierungstag nach 3/2012: Die detektierten SNPs aus den 450K-Daten waren ab diesem Zeitpunkt durch Herrn Prof. Ammerpohl mit den detektierten SNPs in den Hochdurchsatz-Sequenzierungs-Daten bzw. den Sanger-Sequenzen verglichen und bestätigten die korrekte Proben-/Datenzugehörigkeit
- Rohdaten lagen vor und die Probe gehört zum Kollektiv der Vergleichskohorte: Es wurden nur Methylierungsdaten verwendet, die im Rahmen der Forschungsstudien MMML oder ICGC MMML-Seq erhoben wurden und von denen die Rohdaten der 450K-Analysen in Form von .idat-Daten vorlagen. Hierbei ist zu beachten, dass nur aus den Forschungsstudien die 450K-Daten vorlagen (keine 450K-Daten von den klinischen Studien vorhanden).
- Diagnose nach WHO-Klassifikation 2008 ist diffus großzelliges B-Zell-Lymphom (DLBCL)/follikuläres Lymphom (FL): Zusätzlich gibt es Lymphome, die nach WHO-Klassifikation 2008 nicht eindeutig einer Entität zugeordnet werden können und deren Tumore sowohl DLBCL- als auch FL-Komponenten enthalten. Diese Tumore werden als DLBCL/FL bezeichnet und wurden in die Vergleichskohorte bei den Methylierungsanalysen integriert.

- Probeninformationen lagen vollständig vor: Es wurden nur Proben mit vollständigen Angaben zum Geschlecht, Alter, histopathologischer und molekularer Diagnose in die DNA-Methylierungsanalysen eingeschlossen.

Unter Anwendung dieser Selektionskriterien wurde ein Kollektiv von 37 primären Tumor-Proben von Patienten zusammengestellt, das sich größtenteils aus Burkitt-Lymphomen (N=12), Burkitt-Leukämien (N=10) und einer Mischform von DLBCL/FL-Tumoren (N=8) zusammensetzte. Die Charakteristika und Diagnosen der Tumor-Proben sind in der Tabelle 5 zusammengefasst.

Tabelle 5: Charakteristika der Tumor-Proben für die DNA-Methylierungsanalysen

| Fall | Studie | Pathologische Diagnose | Molekulare Diagnose | Geschlecht | Alter [Jahre] |
|---------|--------|-------------------------------|---------------------|------------|---------------|
| MPI-117 | M | B-Zell-Lymphom hochmaligne | intermediate | m | 11 |
| MPI-086 | M | B-Zell-Lymphom hochmaligne | mBL | m | 14 |
| MPI-576 | M | B-Zell-Lymphom hochmaligne | mBL | w | 14 |
| MPI-845 | M | B-Zell-Lymphom hochmaligne | non-mBL | w | 85 |
| 4110996 | I | BL | mBL | m | 2 |
| 4174742 | I | BL | BL | w | 4 |
| MPI-054 | M | BL | intermediate | m | 4 |
| 4170844 | I | BL | BL | w | 5 |
| MPI-532 | M | BL | mBL | m | 5 |
| MPI-005 | M | BL | mBL | m | 5 |
| 4158769 | I | BL | BL | m | 6 |
| 4198478 | I | BL | BL | m | 7 |
| 4156551 | I | BL | BL | m | 8 |
| 4177856 | I | BL | mBL | m | 10 |
| 4162611 | I | BL | mBL | m | 16 |
| 4144633 | I | BL | mBL | w | 44 |
| 4120403 | I | FL Grad 3a/ DLBCL | non-mBL | m | 47 |
| 4113971 | I | FL Grad 3a/DLBCL | non-mBL | w | 61 |
| 4144951 | I | FL Grad 1 + 2/ DLBCL | non-mBL | w | 62 |
| 4110120 | I | FL Grad 3a, 3b/ DLBCL | non-mBL | m | 68 |
| 4113211 | I | FL 3a/ nMZL/ DLBCL | non-mBL | w | 73 |
| 4132950 | I | FL Grad 3a/ DLBCL | non-mBL | m | 76 |
| 4159421 | I | FL Grad 3a, 3b/ DLBCL | non-mBL | w | 78 |
| 4111337 | I | FL Grad 1 + 2/ DLBCL | non-mBL | w | 79 |
| 4151884 | I | BL-Leukämie | BL-Leukämie | m | 2 |
| 4190784 | I | BL-Leukämie | BL-Leukämie | m | 4 |
| 4160810 | I | BL-Leukämie | BL-Leukämie | m | 5 |
| 4108627 | I | BL-Leukämie | BL-Leukämie | w | 5 |
| 4161696 | I | BL-Leukämie | BL-Leukämie | m | 5 |
| 4164943 | I | BL-Leukämie | BL-Leukämie | w | 7 |
| 4110498 | I | BL-Leukämie | BL-Leukämie | m | 8 |
| 4178310 | I | BL-Leukämie | BL-Leukämie | m | 10 |
| 4178345 | I | BL-Leukämie | BL-Leukämie | m | 10 |
| 4152036 | I | BL-Leukämie | BL-Leukämie | m | 19 |
| 4196670 | I | nMZL | non-mBL | w | 60 |
| 4180106 | I | nMZL | non-mBL | m | 68 |
| 4196654 | I | B-Zell-Lymphom niedrigmaligne | non-mBL | m | 70 |

Abkürzungen und Verweise: m: männlich, w: weiblich; Alter: Alter bei Erstdiagnose. DLBCL: Diffus großzelliges B-Zell-Lymphom (engl. Diffuse large B-cell lymphoma). FL: Follikuläres Lymphom (engl. Follicular lymphoma). BL: Burkitt-Lymphom. nMZL: Nodales Marginalzonen-Lymphom. mBL: molekulares Burkitt-Lymphom. non-mBL: nicht-molekulares Burkitt-Lymphom. intermediate: Zwischenstufe (die molekulare Diagnose war nicht eindeutig mBL/ non-mBL zuzuordnen). I: ICGC MMML-Seq. M: MMML-Verbundprojekt.

2.1.2.1.3 Senile DLBCL- und FL-Patienten aus der Vergleichskohorte der verschiedenen klinischen und Forschungsstudien

Aus der klinischen Studie RICOVER-60 und den Forschungsstudien MMML und ICGC MMML-Seq lagen molekulargenetische Daten (u.a. IHC, FISH, Hans-Klassifikator) von insgesamt 61 senilen Patienten mit einem DLBCL bzw. FL vor, die im Rahmen der vorliegenden Arbeit in die Auswertung integriert wurden. Zu den senilen Patienten der Forschungsstudien lag zusätzlich der GEP-Klassifikator vor und von einzelnen Fällen existierten die DNA-Methylierungsdaten bzw. von einem Patienten die *Whole genome sequencing* (WGS)-Daten. Zudem existierte von 28 der 61 senilen Patienten bereits extrahierte DNA aus dem Frischmaterial der Tumore. Da alle 28 DNA-Proben von Patienten des MMML-Verbundprojektes stammten, wird die Kohorte nachfolgend als *MMML-Senil* bezeichnet. Es ist zu beachten, dass von den Patienten aus *MMML-Senil* ausschließlich extrahierte DNA aus den Tumoren von DLBCL-Patienten und keine DNA aus den Tumoren von FL-Patienten vorlag. Interessanterweise wurden in den Analysen des MMML-Verbundprojektes noch keine Mutations- oder DNA-Methylierungsanalysen an den Proben der senilen Patienten durchgeführt, was daraufhin in der vorliegenden Arbeit erfolgte. In dieser Arbeit wurden die molekulargenetischen Daten der insgesamt 61 senilen DLBCL- und FL-Patienten aus den verschiedenen klinischen und Forschungsstudien zur altersassoziierten Analyse der genetischen und epigenetischen Aberrationen in die gesamte Kohorte von senilen Patienten integriert (siehe Abschnitt 2.1.2.3). Eine Übersicht der 61 senilen Patienten aus den klinischen und Forschungsstudien (inkl. *MMML-Senil*) im Zusammenhang mit bereits existierenden bzw. in dieser Arbeit erhobenen molekulargenetischen Daten, ist in der nachfolgenden Tabelle 6 zusammengestellt.

Tabelle 6: Patienten-Charakteristika der senilen Patienten aus den analysierten klinischen und Forschungsstudien

| Fall | Studie | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans- | GEP | FISH | IHC | 450K | NGS |
|----------|--------|------------------|--------|---------------|------------|--------|-----|------|-----|------|-----|
| MPI-802* | M | DLBCL | c | 94 | w | NA | GCB | x | x | x | - |
| MPI-188* | M | DLBCL | c | 93 | w | GCB | NA | x | x | x | - |
| MPI-439 | M | DLBCL | c | 91 | m | GCB | GCB | x | x | - | - |
| MPI-688 | M | DLBCL | a | 90 | m | nonGCB | ABC | x | x | - | - |
| MPI-820* | M | DLBCL | a | 90 | w | NA | ABC | x | x | x | x |
| MPI-734 | M | DLBCL | c | 89 | w | nonGCB | ABC | x | x | - | - |
| MPI-210* | M | DLBCL | c | 89 | w | nonGCB | ABC | x | x | x | - |
| MPI-418 | M | DLBCL | c | 88 | w | nonGCB | ABC | x | x | - | - |
| MPI-164* | M | DLBCL | c | 88 | w | GCB | ABC | x | x | x | - |
| MPI-201* | M | DLBCL | c | 88 | w | nonGCB | NA | x | x | x | x |
| MPI-768* | M | DLBCL | i | 88 | w | GCB | GCB | x | x | x | x |
| MPI-557 | M | DLBCL | c | 87 | m | nonGCB | NA | x | x | - | - |
| MPI-331* | M | DLBCL | NOS | 87 | w | nonGCB | ABC | x | x | x | x |
| MPI-411* | M | DLBCL | c | 87 | w | GCB | GCB | x | x | x | x |
| MPI-042 | M | DLBCL | c | 86 | w | NA | ABC | x | x | - | - |
| MPI-440 | M | DLBCL | c | 86 | w | nonGCB | NA | x | x | - | - |
| MPI-165* | M | DLBCL | NOS | 86 | w | GCB | GCB | x | x | x | - |
| MPI-243* | M | DLBCL | c | 86 | m | nonGCB | NA | x | x | x | x |

| Fall | Studie | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans- | GEP | FISH | IHC | 450K | NGS |
|----------|--------|------------------|------------|---------------|------------|--------|-----|------|-----|----------------|-----------------|
| 4138527 | I | DLBCL | c | 85 | w | GCB | GCB | x | x | x | x |
| MPI-128 | M | DLBCL | c | 85 | w | GCB | GCB | x | x | x | - |
| MPI-745 | M | DLBCL | NOS | 85 | m | nonGCB | ABC | x | x | - | - |
| 4134434 | I | DLBCL | NOS | 84 | m | nonGCB | GCB | x | x | - | x |
| MPI-009 | M | DLBCL | c | 84 | w | nonGCB | ABC | x | x | - | - |
| MPI-295 | M | DLBCL | c | 84 | w | GCB | GCB | x | x | - | - |
| MPI-746 | M | DLBCL | NOS | 84 | w | GCB | GCB | x | x | - | - |
| MPI-309* | M | DLBCL | c | 84 | m | nonGCB | GCB | x | x | x | x |
| MPI-762* | M | DLBCL | NOS | 84 | w | NA | NA | x | x | x | x |
| MPI-851* | M | DLBCL | NOS | 84 | m | GCB | GCB | - | | x | x ^{NA} |
| MPI-857* | M | DLBCL | NOS | 84 | w | NA | ABC | x | x | x | x |
| MPI-154 | M | DLBCL | c | 83 | w | NA | ABC | x | x | - | - |
| MPI-363 | M | DLBCL | c | 82 | w | GCB | NA | x | x | - | - |
| MPI-599 | M | DLBCL | c | 82 | m | nonGCB | ABC | x | x | - | - |
| MPI-076* | M | DLBCL | c | 82 | w | nonGCB | NA | x | x | x | - |
| MPI-191* | M | DLBCL | p | 82 | m | nonGCB | GCB | x | x | x | x |
| MPI-206* | M | DLBCL | c | 82 | m | nonGCB | NA | x | x | x | x |
| MPI-749* | M | DLBCL | c | 82 | w | GCB | GCB | x | x | x | x ^{NA} |
| MPI-118* | M | DLBCL | c | 81 | m | nonGCB | ABC | x | x | x | x |
| MPI-119* | M | DLBCL | c | 81 | m | GCB | GCB | x | x | x ¹ | x |
| MPI-756* | M | DLBCL | c | 81 | w | GCB | GCB | x | x | x | x |
| MPI-841* | M | DLBCL | c | 81 | w | NA | ABC | x | x | x | x |
| MPI-859* | M | DLBCL | c | 81 | w | GCB | GCB | x | x | x | x |
| MPI-883* | M | DLBCL | c | 81 | m | NA | ABC | x | x | x | x |
| 4176325 | I | DLBCL | c | 80 | w | GCB | GCB | x | x | x | - |
| MPI-050 | M | DLBCL | c | 80 | w | nonGCB | NA | x | x | - | - |
| MPI-366 | M | DLBCL | NOS | 80 | m | nonGCB | ABC | x | x | - | - |
| MPI-047* | M | DLBCL | NOS | 80 | m | nonGCB | ABC | x | x | x | x |
| MPI-099* | M | DLBCL | c | 80 | w | NA | GCB | x | x | x | x |
| MPI-849* | M | DLBCL | c | 80 | m | GCB | NA | x | x | x | x |
| 133 | R | DLBCL | c | 80 | w | nonGCB | NA | x | x | - | - |
| 323 | R | DLBCL | c | 80 | w | GCB | NA | x | x | - | - |
| 762 | R | DLBCL | c | 80 | w | NA | NA | x | x | - | - |
| 93 | R | DLBCL | NOS | 80 | w | nonGCB | NA | x | x | - | - |
| 1327 | R | DLBCL | NOS | 80 | w | NA | NA | x | x | - | - |
| MPI-647 | M | FL | Grad 1 | 86 | m | NA | NA | x | x | - | - |
| MPI-853 | M | FL | Grad 3a | 84 | m | NA | NA | x | x | - | - |
| MPI-297 | M | FL | Grad 2 | 82 | m | NA | NA | x | x | - | - |
| 4120468 | I | FL | Grad 1 + 2 | 80 | m | NA | NA | x | x | - | - |
| MPI-097 | M | FL | Grad 3a | 80 | m | NA | NA | x | x | - | - |
| MPI-515 | M | FL | Grad 2 | 80 | m | NA | NA | x | x | - | - |
| MPI-656 | M | FL | Grad 3b | 80 | m | NA | NA | x | x | - | - |
| MPI-666 | M | FL | Grad 3a | 80 | w | NA | NA | x | x | - | - |

Abkürzungen und Verweise : DLBCL: diffus großzelliges B-Zell-Lymphom. FL: folliculäres Lymphom. c: centroblastisch. p: plasmoblastisch. i: immunoblastisch. a: anaplastisch-großzellig. NOS: nicht klassifizierbar. m: männlich. w: weiblich. NA: (engl. not available/ analyzable) nicht verfügbar oder nicht auswertbar. GCB: (engl. germinal center B-cell) Keimzentrumsähnliche B-Zellen. nonGCB: (engl.non-germinal center B-cell) Aktivierte B-Zellen. ABC: (engl. Activated B-cell) Aktivierte B-Zellen. IHC: Immunhistochemie. Hans: Klassifikator nach Hans. GEP: Klassifikator nach Genexpressionsprofil. FISH: Fluoreszenz-in-situ-Hybridisierung. NGS: Hochdurchsatz-Sequenzierung (engl. Next Generation Sequencing). 450K: Analyse der DNA-Methylierung mittels Infinium HumanMethylation 450K Array. Alter: Alter bei Erstdiagnose. x: Patienten-Probe wurde mit der Technik in dieser Studie analysiert.-: Patienten-Probe wurde mit der Technik in dieser Studie nicht analysiert. I: ICGC MMML-Seq. M: MMML-Verbundprojekt. R: RICOVER. *:Patienten-Probe der Kohorte *MMML-Senil*, wozu bereits extrahierte DNA aus dem Frischmaterial des Tumors für die Hochdurchsatz-Sequenzierung bzw.DNA-Methylierungsanalysen vorlag.

2.1.2.2 Sehr alte Patienten mit extrahierter DNA aus FFPE-Tumormaterial (*Jack*)

In den letzten Jahren wurden in verschiedenen klinischen und Forschungsstudien vorwiegend Daten von Kindern und Erwachsenen bis zu einem Alter von 80 Jahren erhoben, wodurch die Analyse der Altersgruppe der senilen Patienten bisher deutlich unterrepräsentiert ist^{100-102,212-215}. Daher sollte im Rahmen der vorliegenden Arbeit gezielt eine neue Kohorte von senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom zusammengestellt und charakterisiert werden. Dazu wurde das Untersuchungskollektiv des Instituts für Pathologie/ Sektion Hämatopathologie hinsichtlich seniler Patienten mit einem Keimzentrums-B-Zell-Lymphom selektiert. Aufgrund der geringen Anzahl an potentiell geeigneten Tumormaterialien von senilen Patienten mit einem Alter von mind. 80 Jahren, wurde das Einschlusskriterium des Ersterkrankungsalters bei Diagnose von ursprünglich mind. 80 auf 74 Jahre abgesenkt. Die Kohorte von senilen und sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den Kollektiven des Instituts für Humangenetik und des Instituts für Pathologie/ Sektion Hämatopathologie wird im Weiteren als *Jack* bezeichnet.

Die Selektion der Tumor-Proben für *Jack* erfolgte unter folgenden Kriterien:

- Eingangsdatum zwischen 03/2002 und 03/2013
- Diagnose nach WHO-Klassifikation 2008: Diffus großzelliges B-Zell-Lymphom (DLBCL) oder Follikuläres Lymphom (FL)
- Das Tumor-Material darf kein Rezidiv sein
- Ersterkrankungsalter von mind. 74 Jahren
- FFPE-Material für FISH-Analysen geeignet
- Ausreichend Material für DNA- und RNA-Extraktion und FISH-Analysen ist vorhanden

Nach dem angepassten Einschlusskriterium des Ersterkrankungsalters von mind. 80 auf 74 Jahre konnten insgesamt 198 Patienten mit der Diagnose DLBCL oder FL aus den Gesamtkollektiven des Instituts für Pathologie/ Sektion Hämatopathologie und des Instituts für Humangenetik identifiziert werden (DLBCL: N= 160, Median: 80 Jahre und FL: N=38, Median: 78 Jahre).

Das Schneiden des Tumor-Materials für die DNA-Extraktion und die Anfertigung der Schnitte für die IHC und FISH sowie die immunhistochemische Färbung der Antigene CD10, BCL2, BCL6, MUM1 und Ki67 wurden freundlicherweise vom Institut für Pathologie/ Sektion Hämatopathologie durchgeführt. Zusätzlich wurden FISH-Präparate von kleinen Gewebestücken in Form von Tissue Microarray (TMA) angefertigt. Die anschließende technische Durchführung und Auswertung der FISH-Analysen mit spezifischen Sonden für die Gene *BCL2*, *BCL6*, *MYC* und *IGH* erfolgte am Institut für Humangenetik durch Frau Dr. Inga Nagel und Frau Reina Zühlke-Jenisch. Hierzu wurden je nach Verfügbarkeit und Qualität des Patientenmaterials entweder TMAs oder Einzeltumor-Schnitte genutzt und ausgewertet. Insgesamt wurden FISH-Analysen von 141/198 Patienten durchgeführt, wovon 121/198 FISH-Daten auswertbar waren. D.h. es lagen Ergebnisse von mind. einer analysierten Sonde vor (DLBCL: N= 95, Median: 79 Jahre und FL: N= 26, Median: 77 Jahre). Von den restlichen 58 Patienten konnten keine FISH-Analysen durchgeführt werden, da nicht genügend oder kein geeignetes Material vorhanden war (nicht TMA-fähig).

Zusätzlich wurden FISH-Analysen des *CD58*-Gens durch Frau Dr. Cristina Lopez Gonzalez und Frau Reina Zühlke-Jenisch am Institut für Humangenetik durchgeführt. Insgesamt waren 45 der 55 analysierten Patienten der Kohorte *Jack* mit den FISH-Sonden des *CD58*-Gens auswertbar.

Die DLBCL-Lymphome wurden in der vorliegenden Arbeit entsprechend Ihrer vermuteten Ursprungszelle eingeteilt. Dazu erfolgte die Einteilung nach zwei verschiedenen Klassifizierungen:

- Nach Hans et al. in GCB- und nonGCB-Subtyp (*Hans-Klassifikator*¹⁴¹)
- Nach der Genexpression von Microarray-Analysen oder RNA-Sequenzierung in ABC- und GCB-Subtyp (*GEP-Klassifikator*^{142,143})

Den Patienten mit einem FL wurden keine Subtypen zugeordnet, da die Klassifikatoren für diese Entität nicht entwickelt wurden und demnach nicht anwendbar sind. Für die *Jack*-Kohorte und die klinischen Studien RICOVER-60 und NHL-B lagen keine Daten zum *GEP-Klassifikator* vor. Deshalb wurde für die *Jack*-Kohorte und alle anderen im Rahmen dieser Arbeit analysierten Fälle der *Hans-Klassifikator* verwendet. Zusätzlich wurde der *GEP-Klassifikator* verwendet, wenn dieser für die analysierten Fälle vorhanden war (Patienten der Verbundprojekte MMML und ICGC MMML-Seq, siehe Abschnitt 2.1.2.1). Eine Übersicht von der Kohorte *Jack* und den dazugehörigen Analysen, die im Rahmen der altersassoziierten Analysen erhoben wurden, ist in der Tabelle 7 zusammengestellt.

Tabelle 7: Patienten-Charakteristika der Kohorte Jack

| Fall | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans | IHC | TMA/ FISH | NGS | CD58 |
|----------|------------------|--------|---------------|------------|--------|-----|-----------|-----------------|-----------------|
| Jack_36 | DLBCL | c | 97 | w | nonGCB | x | x | x | x |
| Jack_27 | DLBCL | NOS | 93 | w | nonGCB | x | x | x ^{NA} | x |
| Jack_33 | DLBCL | c | 92 | m | GCB | x | x | x | x |
| Jack_107 | DLBCL | NOS | 90 | w | NA | x | - | - | - |
| Jack_32 | DLBCL | c | 89 | m | GCB | x | x | x ^{NA} | x |
| Jack_79 | DLBCL | NOS | 89 | m | NA | x | - | - | - |
| Jack_90 | DLBCL | NOS | 89 | m | nonGCB | x | x | - | - |
| Jack_3 | DLBCL | NOS | 88 | m | nonGCB | x | x | x ^{NA} | x |
| Jack_48 | DLBCL | NOS | 88 | m | NA | x | x | x ^{NA} | x ^{NA} |
| Jack_23 | DLBCL | i | 88 | w | GCB | x | x | x | x |
| Jack_102 | DLBCL | i | 88 | m | nonGCB | x | x | - | - |
| Jack_49 | DLBCL | NOS | 88 | w | nonGCB | x | NA | - | x ^{NA} |
| Jack_97 | DLBCL | NOS | 88 | m | NA | x | - | - | - |
| Jack_44 | DLBCL | c | 87 | w | nonGCB | x | x | x ^{NA} | x |
| Jack_35 | DLBCL | NOS | 87 | w | GCB | x | x | x | x |
| Jack_104 | DLBCL | c | 87 | m | GCB | x | - | - | - |
| Jack_2 | DLBCL | NOS | 87 | w | GCB | x | x | - | x |
| Jack_67 | DLBCL | i | 87 | m | nonGCB | x | - | - | - |
| Jack_77 | DLBCL | c | 87 | w | GCB | x | NA | - | x |
| Jack_92 | DLBCL | a | 87 | w | NA | x | - | - | - |
| Jack_96 | DLBCL | NOS | 87 | w | GCB | x | - | - | - |
| Jack_37 | DLBCL | c | 86 | m | nonGCB | x | x | x ^{NA} | x |
| Jack_42 | DLBCL | c | 86 | m | GCB | x | NA | x ^{NA} | x |
| Jack_47 | DLBCL | c | 86 | w | nonGCB | x | x | x | x |
| Jack_110 | DLBCL | i | 86 | w | GCB | x | x | - | - |
| Jack_12 | DLBCL | c | 86 | w | NA | x | x | - | - |
| Jack_14 | DLBCL | NOS | 86 | w | GCB | x | - | - | - |
| Jack_63 | DLBCL | NOS | 86 | w | GCB | x | NA | - | - |
| Jack_70 | DLBCL | c | 86 | m | GCB | x | x | - | - |
| Jack_19 | DLBCL | i | 85 | w | NA | x | x | x ^{NA} | - |
| Jack_22 | DLBCL | NOS | 85 | w | nonGCB | x | x | x ^{NA} | x |
| Jack_26 | DLBCL | c | 85 | m | nonGCB | x | NA | x ^{NA} | x ^{NA} |
| Jack_20 | DLBCL | c | 85 | m | nonGCB | x | x | x | x |
| Jack_34 | DLBCL | c | 85 | m | GCB | x | x | x | x |
| Jack_43 | DLBCL | NOS | 85 | m | nonGCB | x | x | x | x |
| Jack_113 | DLBCL | c | 85 | m | GCB | x | x | - | - |
| Jack_16 | DLBCL | NOS | 85 | w | NA | x | x | - | - |
| Jack_68 | DLBCL | NOS | 85 | w | NA | x | - | - | - |
| Jack_72 | DLBCL | NOS | 85 | m | NA | x | x | - | - |
| Jack_29 | DLBCL | c | 84 | w | GCB | x | x | x ^{NA} | x |
| Jack_103 | DLBCL | c | 84 | w | NA | x | - | - | - |
| Jack_108 | DLBCL | c | 84 | w | NA | x | - | - | - |
| Jack_69 | DLBCL | NOS | 84 | m | NA | x | - | - | - |
| Jack_73 | DLBCL | NOS | 84 | m | GCB | x | - | - | - |
| Jack_85 | DLBCL | c | 84 | w | GCB | x | - | - | - |
| Jack_86 | DLBCL | NOS | 84 | w | GCB | x | NA | - | - |
| Jack_91 | DLBCL | c | 84 | w | NA | x | NA | - | - |
| Jack_93 | DLBCL | c | 84 | w | GCB | x | - | - | - |
| Jack_94 | DLBCL | NOS | 84 | m | GCB | x | - | - | - |
| Jack_28 | DLBCL | i | 83 | w | GCB | x | x | x ^{NA} | x |
| Jack_38 | DLBCL | c | 83 | w | nonGCB | x | x | x ^{NA} | x |
| Jack_1 | DLBCL | NOS | 83 | m | nonGCB | x | x | x | x |
| Jack_10 | DLBCL | c | 83 | m | nonGCB | x | x | x | x |
| Jack_31 | DLBCL | a | 83 | w | nonGCB | x | x | x | x |
| Jack_39 | DLBCL | NOS | 83 | m | nonGCB | x | x | x | x ^{NA} |
| Jack_4 | DLBCL | c | 83 | m | nonGCB | x | x | x | x |
| Jack_64 | DLBCL | c | 83 | w | NA | x | - | - | - |
| Jack_66 | DLBCL | NOS | 83 | m | nonGCB | x | NA | - | - |
| Jack_84 | DLBCL | c | 83 | w | nonGCB | x | x | - | - |
| Jack_87 | DLBCL | NOS | 83 | w | NA | x | - | - | - |

| Fall | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans | IHC | TMA/ FISH | NGS | CD58 |
|----------|------------------|--------|---------------|------------|--------|-----|-----------|-----------------|-----------------|
| Jack_89 | DLBCL | NOS | 83 | w | GCB | x | x | - | - |
| Jack_52 | DLBCL | c | 82 | w | NA | x | x | x ^{NA} | - |
| Jack_18 | DLBCL | NOS | 82 | w | GCB | x | x | x | - |
| Jack_105 | DLBCL | NOS | 82 | w | nonGCB | x | NA | - | - |
| Jack_111 | DLBCL | NOS | 82 | w | GCB | x | x | - | - |
| Jack_13 | DLBCL | NOS | 82 | w | nonGCB | x | x | - | - |
| Jack_46 | DLBCL | c | 82 | w | nonGCB | x | NA | - | x ^{NA} |
| Jack_88 | DLBCL | c | 82 | w | GCB | x | - | - | - |
| Jack_9 | DLBCL | NOS | 82 | w | NA | NA | x | - | - |
| Jack_98 | DLBCL | i | 82 | m | NA | x | - | - | - |
| Jack_99 | DLBCL | NOS | 82 | m | NA | x | - | - | - |
| Jack_30 | DLBCL | c | 81 | m | nonGCB | x | x | x ^{NA} | x ^{NA} |
| Jack_40 | DLBCL | c | 81 | w | GCB | x | x | x ^{NA} | x |
| Jack_114 | DLBCL | NOS | 81 | w | nonGCB | x | - | - | - |
| Jack_75 | DLBCL | NOS | 81 | w | GCB | x | - | - | - |
| Jack_51 | DLBCL | c | 80 | w | GCB | x | x | x ^{NA} | - |
| Jack_54 | DLBCL | c | 80 | w | GCB | x | x | x ^{NA} | - |
| Jack_25 | DLBCL | c | 80 | m | GCB | x | x | x | x |
| Jack_57 | DLBCL | NOS | 80 | m | GCB | x | x | x | - |
| Jack_106 | DLBCL | NOS | 80 | w | nonGCB | x | NA | - | - |
| Jack_17 | DLBCL | c | 80 | m | GCB | x | - | - | - |
| Jack_62 | DLBCL | NOS | 80 | w | GCB | x | NA | - | - |
| Jack_80 | DLBCL | c | 80 | w | NA | x | x | - | - |
| Jack_81 | DLBCL | c | 80 | m | GCB | x | - | - | - |
| Jack_41 | DLBCL | c | 79 | m | GCB | x | x | x ^{NA} | x |
| Jack_24 | DLBCL | i | 79 | m | nonGCB | x | x | x | x |
| Jack_112 | DLBCL | c | 79 | w | GCB | x | x | - | - |
| Jack_119 | DLBCL | i | 79 | m | NA | x | x | - | - |
| Jack_134 | DLBCL | c | 79 | m | nonGCB | x | x | - | x |
| Jack_147 | DLBCL | NOS | 79 | m | GCB | x | x | - | x |
| Jack_15 | DLBCL | i | 79 | w | nonGCB | x | x | - | - |
| Jack_172 | DLBCL | c | 79 | m | NA | x | - | - | - |
| Jack_173 | DLBCL | NOS | 79 | w | GCB | x | - | - | - |
| Jack_184 | DLBCL | c | 79 | w | GCB | x | - | - | - |
| Jack_191 | DLBCL | NOS | 79 | m | NA | x | - | - | - |
| Jack_196 | DLBCL | NOS | 79 | w | NA | x | - | - | - |
| Jack_199 | DLBCL | i | 79 | w | GCB | x | x | - | - |
| Jack_76 | DLBCL | NOS | 79 | w | GCB | x | - | - | - |
| Jack_83 | DLBCL | NOS | 79 | m | nonGCB | x | x | - | - |
| Jack_95 | DLBCL | c | 79 | w | nonGCB | x | x | - | - |
| Jack_21 | DLBCL | NOS | 78 | w | GCB | x | x | x ^{NA} | x ^{NA} |
| Jack_139 | DLBCL | NOS | 78 | w | NA | x | x | - | x |
| Jack_164 | DLBCL | NOS | 78 | w | GCB | x | x | - | - |
| Jack_176 | DLBCL | c | 78 | w | NA | x | - | - | - |
| Jack_180 | DLBCL | NOS | 78 | m | NA | x | x | - | - |
| Jack_188 | DLBCL | i | 78 | w | nonGCB | x | x | - | - |
| Jack_82 | DLBCL | c | 78 | w | nonGCB | x | - | - | - |
| Jack_45 | DLBCL | c | 77 | w | GCB | x | x | x ^{NA} | x |
| Jack_6 | DLBCL | c | 77 | w | nonGCB | x | x | x ^{NA} | - |
| Jack_117 | DLBCL | NOS | 77 | w | GCB | x | x | - | - |
| Jack_125 | DLBCL | c | 77 | m | GCB | x | x | - | - |
| Jack_128 | DLBCL | c | 77 | m | GCB | x | x | - | x |
| Jack_129 | DLBCL | c | 77 | m | NA | x | x | - | - |
| Jack_133 | DLBCL | c | 77 | m | GCB | x | x | - | x |
| Jack_143 | DLBCL | c | 77 | w | GCB | x | x | - | x |
| Jack_149 | DLBCL | NOS | 77 | w | nonGCB | x | NA | - | - |
| Jack_159 | DLBCL | NOS | 77 | w | nonGCB | x | NA | - | - |
| Jack_165 | DLBCL | c | 77 | w | nonGCB | x | - | - | - |
| Jack_167 | DLBCL | c | 77 | w | NA | x | - | - | - |
| Jack_171 | DLBCL | c | 77 | w | NA | x | - | - | - |
| Jack_175 | DLBCL | i | 77 | w | NA | x | NA | - | - |
| Jack_183 | DLBCL | NOS | 77 | w | GCB | x | - | - | - |

| Fall | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans | IHC | TMA/ FISH | NGS | CD58 |
|----------|------------------|-----------------|---------------|------------|--------|-----|-----------|-----------------|-----------------|
| Jack_190 | DLBCL | NOS | 77 | m | GCB | x | - | - | - |
| Jack_71 | DLBCL | c | 77 | w | NA | x | - | - | - |
| Jack_120 | DLBCL | NOS | 76 | w | GCB | x | x | - | - |
| Jack_122 | DLBCL | c | 76 | w | nonGCB | x | x | - | - |
| Jack_127 | DLBCL | NOS | 76 | w | GCB | x | x | - | - |
| Jack_130 | DLBCL | NOS | 76 | w | nonGCB | x | x | - | - |
| Jack_135 | DLBCL | c | 76 | w | GCB | x | x | - | x |
| Jack_141 | DLBCL | a | 76 | m | GCB | x | x | - | x |
| Jack_145 | DLBCL | c | 76 | w | GCB | x | x | - | x |
| Jack_156 | DLBCL | NOS | 76 | m | GCB | x | x | - | - |
| Jack_170 | DLBCL | c | 76 | m | NA | x | x | - | - |
| Jack_174 | DLBCL | NOS | 76 | w | GCB | x | - | - | - |
| Jack_177 | DLBCL | NOS | 76 | m | NA | x | - | - | - |
| Jack_194 | DLBCL | NOS | 76 | w | GCB | x | - | - | - |
| Jack_115 | DLBCL | c | 75 | w | NA | x | x | - | x |
| Jack_116 | DLBCL | NOS | 75 | m | NA | x | x | - | - |
| Jack_118 | DLBCL | i | 75 | w | GCB | x | x | - | x ^{NA} |
| Jack_126 | DLBCL | c | 75 | w | GCB | x | x | - | - |
| Jack_131 | DLBCL | NOS | 75 | w | NA | x | NA | - | x ^{NA} |
| Jack_132 | DLBCL | NOS | 75 | m | nonGCB | x | - | - | - |
| Jack_136 | DLBCL | c | 75 | m | NA | x | x | - | x |
| Jack_137 | DLBCL | c | 75 | w | GCB | x | x | - | x |
| Jack_140 | DLBCL | c | 75 | m | NA | x | NA | - | x |
| Jack_168 | DLBCL | NOS | 75 | m | NA | x | - | - | - |
| Jack_169 | DLBCL | c | 75 | w | NA | x | x | - | - |
| Jack_179 | DLBCL | c | 75 | w | GCB | x | - | - | - |
| Jack_187 | DLBCL | c | 75 | w | NA | x | x | - | - |
| Jack_189 | DLBCL | NOS | 75 | m | GCB | x | - | - | - |
| Jack_138 | DLBCL | c | 74 | w | GCB | x | x | - | x |
| Jack_142 | DLBCL | i | 74 | m | GCB | x | - | - | - |
| Jack_144 | DLBCL | c | 74 | w | GCB | x | x | - | x |
| Jack_146 | DLBCL | NOS | 74 | w | nonGCB | x | x | - | x |
| Jack_148 | DLBCL | i | 74 | w | nonGCB | x | x | - | x |
| Jack_155 | DLBCL | NOS | 74 | w | nonGCB | x | x | - | - |
| Jack_166 | DLBCL | c | 74 | m | NA | x | x | - | - |
| Jack_182 | DLBCL | c | 74 | m | GCB | x | x | - | - |
| Jack_192 | DLBCL | c | 74 | w | GCB | x | - | - | - |
| Jack_193 | DLBCL | NOS | 74 | w | nonGCB | x | x | - | - |
| Jack_7 | FL | Grad 2 | 95 | w | NA | x | x | - | - |
| Jack_58 | FL | Grad 1 | 93 | m | NA | x | x | x ^{NA} | - |
| Jack_74 | FL | Grad 1 | 93 | m | NA | x | - | - | - |
| Jack_53 | FL | Grad 2 | 90 | m | NA | x | x | x ^{NA} | - |
| Jack_61 | FL | Grad 1 | 85 | w | NA | x | NA | x ^{NA} | - |
| Jack_50 | FL | Grad 2 | 84 | m | NA | x | x | x ^{NA} | - |
| Jack_55 | FL | Grad 1 | 84 | w | NA | x | x | x | - |
| Jack_100 | FL | Grad 1 | 84 | w | NA | x | - | - | - |
| Jack_109 | FL | Grad 1 + 2 | 84 | w | NA | x | x | - | - |
| Jack_55 | FL | Grad 1 | 84 | m | NA | x | x | - | - |
| Jack_60 | FL | Grad 1 + 2 | 82 | m | NA | x | x | x ^{NA} | - |
| Jack_200 | FL | Grad 1 + 2 + 3a | 81 | w | NA | x | x | x ^{NA} | - |
| Jack_65 | FL | Grad 1 | 81 | w | NA | x | - | - | - |
| Jack_78 | FL | Grad 2 | 81 | m | NA | x | - | - | - |
| Jack_5 | FL | Grad 1 + 2 + 3a | 80 | w | NA | x | x | x | - |
| Jack_8 | FL | Grad 1 | 80 | w | NA | x | - | - | - |
| Jack_101 | FL | Grad 1 | 79 | w | NA | x | - | - | - |
| Jack_59 | FL | Grad 1 + 2 | 78 | m | NA | x | x | x | - |
| Jack_163 | FL | Grad 1 + 2 + 3a | 78 | m | NA | x | x | - | - |
| Jack_56 | FL | Grad 1 | 77 | m | NA | x | x | x | - |
| Jack_11 | FL | Grad 1 | 77 | m | NA | x | x | - | - |
| Jack_121 | FL | Grad 2 | 77 | m | NA | x | x | - | - |
| Jack_154 | FL | Grad 1 + 2 | 77 | w | NA | x | x | - | - |
| Jack_181 | FL | Grad 1 + 2 | 77 | w | NA | x | - | - | - |

| Fall | Initial-diagnose | Subtyp | Alter [Jahre] | Geschlecht | Hans | IHC | TMA/ FISH | NGS | CD58 |
|----------|------------------|-------------|---------------|------------|------|-----|-----------|-----|------|
| Jack_123 | FL | Grad 2 + 3a | 76 | m | NA | x | x | - | - |
| Jack_150 | FL | Grad 1 + 2 | 76 | m | NA | x | x | - | - |
| Jack_153 | FL | Grad 1 + 2 | 76 | w | NA | x | NA | - | - |
| Jack_178 | FL | Grad 1 + 2 | 76 | w | NA | x | - | - | - |
| Jack_186 | FL | Grad 1 + 2 | 76 | w | NA | x | x | - | - |
| Jack_198 | FL | Grad 1 + 2 | 76 | m | NA | x | - | - | - |
| Jack_124 | FL | Grad 1 + 2 | 75 | w | NA | x | x | - | - |
| Jack_151 | FL | Grad 1 + 2 | 75 | w | NA | x | x | - | - |
| Jack_162 | FL | Grad 3a | 75 | m | NA | x | x | - | - |
| Jack_152 | FL | Grad 1 + 2 | 74 | w | NA | x | x | - | - |
| Jack_160 | FL | Grad 1 + 2 | 74 | m | NA | x | x | - | - |
| Jack_161 | FL | Grad 1 + 2 | 74 | m | NA | x | x | - | - |
| Jack_195 | FL | Grad 1 + 2 | 74 | m | NA | x | x | - | - |
| Jack_197 | FL | Grad 1 + 2 | 74 | m | NA | x | - | - | - |

Abkürzungen und Verweise: c: centroblastisch. p: plasmoblastisch. i: immunoblastisch. a: anaplastisch-großzellig. NOS: nicht klassifizierbar. m: männlich. w: weiblich. Alter: Alter bei Erstdiagnose. DLBCL: (engl. diffuse large B-cell lymphoma) Diffus großzelliges B-Zell-Lymphom. FL: (engl. follicular lymphoma) Follikuläres Lymphom. NA/^{NA}: (engl. not available/analyzable) nicht auswertbar. Für FL-Patienten wurden keine Subtypen bestimmt. GCB: (engl. germinal center B-cell) Keimzentrumsähnliche B-Zellen. nonGCB: (engl. non-germinal center B-cell) Aktivierte B-Zellen. IHC: Immunhistochemie. Hans: Klassifikator nach Hans. FISH: Fluoreszenz-in-situ-Hybridisierung. TMA: (engl. Tissue Microarray) Kleine Gewebestanzten. NGS: (engl. Next Generation Sequencing) Hochdurchsatz-Sequenzierung. CD58: Fluoreszenz-in-situ-Hybridisierung des CD58-Gens. x: Patienten-Probe wurde mit der Technik in dieser Studie analysiert. Patienten-Probe wurde mit der Technik in dieser Studie nicht analysiert. x^{NA}: Patienten-Probe wurde mit der Technik in dieser Studie analysiert, die Coverage über alle Gene war aber <30x.

2.1.2.3 Übersicht aller analysierten senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom (DLBCL/FL-Senil)

Die Kohorte zur Analyse von altersassoziierten genetischen und epigenetischen Aberrationen von senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem Ersterkrankungsalter von mind. 80 Jahren wird in der vorliegenden Arbeit als *DLBCL/FL-Senil* bezeichnet (siehe Abbildung 13) und setzt sich zusammen aus folgenden Teil-Kohorten:

- 1.) Existierende Daten von senilen Patienten aus anderen klinischen und Forschungsstudien (Senile Patienten inkl. *MMML-Senil*, N= 61, siehe Abschnitt 2.1.2.1)
- 2.) Tumor-Material aus FFPE-Material von senilen Patienten (*Jack*, N= 100, siehe Abschnitt 2.1.2.2)

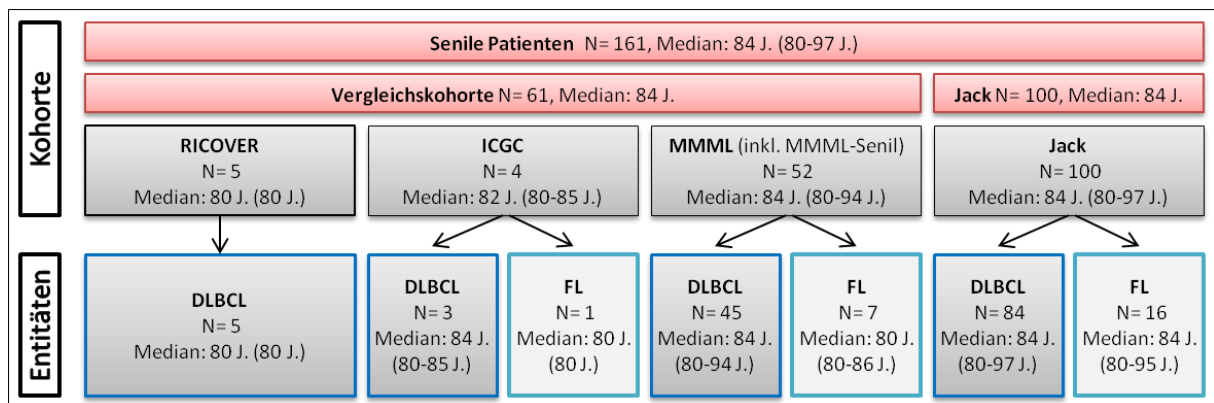


Abbildung 13: Schematische Darstellung der analysierten senilen DLBCL- und FL-Patienten

In die altersassoziierten Analysen der vorliegenden Arbeit wurden bereits erhobene Daten von senilen Patienten mit einem DLBCL oder FL aus verschiedenen vorangegangenen Studien integriert. Hierzu lagen Daten von DLBCL- und FL-Patienten aus der klinischen Studie RICOVER sowie aus den Forschungsstudien ICGC MMML-Seq und MMML vor. Von 28 DLBCL-Patienten des MMML-Verbundprojektes lag bereits extrahierte DNA von den Tumoren vor, die im Rahmen dieser Arbeit für weitere Molekulargenetische Analysen (NGS, 450K) genutzt wurde. Zusätzlich wurden 100 senile Patienten aus den Untersuchungskollektiven des Instituts für Humangenetik und dem Institut für Pathologie/ Sektion Hämatopathologie neu identifiziert und charakterisiert (Kohorte *Jack*). Aus den Forschungsstudien existierten zusätzlich die Patienteninformationen und Ergebnisse von molekulargenetischen Analysen zu Fällen mit einem diagnostizierten FL. Abkürzungen: DLBCL: Diffus großzelliges B-Zell-Lymphom. FL: Follikuläres Lymphom. J: Jahre. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. RICOVER und NHL-B: Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom. NGS: (engl. Next Generation Sequencing) Hochdurchsatz-Sequenzierung. 450K: Analyse der DNA-Methylierung mittels Infinium HumanMethylation 450K Array.

Das Gesamtkollektiv *DLBCL/FL-Senil* zur Untersuchung von senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom bestand aus 161 Tumor-Proben im Alter von 80 bis 97 Jahren, wovon die Diagnose nach WHO Klassifikation 2008 bei 137 Patienten ein DLBCL (Median: 84 Jahre) und bei 24 Patienten ein FL (Median: 83 Jahre) war.

2.1.2.4 Validierungskohorte für das Kandidatengen *CD58*

Zur Validierung der strukturellen Variationen im *CD58*-Gen wurde eine neue Kohorte aus dem Untersuchungsmaterial der tumorgenetischen Routine-Diagnostik des Instituts für Humangenetik erstellt. Die Selektion erfolgte von Frau Dr. Cristina Lopez Gonzalez unter folgenden Kriterien:

- Eine Einwilligung des Patienten zur Verwendung des Probenmaterials für Forschungszwecke lag vor
- Diagnose nach WHO-Klassifikation 2008 ist Diffus großzelliges B-Zell-Lymphom (DLBCL)
- Tumorzellen sind für die FISH-Analysen vorhanden
- Chromosomenaberrationen im *CD58*-Gen sind unbekannt bzw. der Fall wurde noch in keiner anderen Studie analysiert (z.B. MMML, ICGC MMML-Seq)

Die Validierungskohorte (N=41) wurde in der vorliegenden Arbeit ausschließlich für die Analysen der Altersassoziation von genetischen Aberrationen im *CD58*-Gen bei DLBCL-Patienten verwendet. Hierzu wurde ein Kollektiv zusammengestellt, das ein breites Altersspektrum an Patienten abdeckt (Ersterkrankungsalter: 15-83 Jahre, Median: 67 Jahre). Das Ersterkrankungsalter für alle Patienten ist in der nachfolgenden Tabelle 8 aufgelistet.

Tabelle 8: Übersicht des Ersterkrankungsalters der Patienten aus der Validierungskohorte

| Fall | Alter [Jahre] | Fall | Alter [Jahre] | Fall | Alter [Jahre] | Fall | Alter [Jahre] |
|---------|---------------|---------|---------------|---------|---------------|---------|---------------|
| 1198-11 | 15 | 294-10 | 58 | 831-10 | 67 | 1684-7 | 76 |
| 2255-7 | 15 | 618-6 | 58 | 1241-12 | 68 | 38-14 | 76 |
| 1161-8 | 44 | 931-13 | 60 | 1315-11 | 68 | 1206-12 | 77 |
| 1830-6 | 45 | 1265-10 | 61 | 117-7 | 69 | 1255-5 | 78 |
| 527-11 | 48 | 231-7 | 61 | 395-8 | 70 | 931-14 | 78 |
| 2367-8 | 53 | 75-9 | 61 | 841-14 | 70 | 1188-14 | 79 |
| 785-7 | 54 | 2272-9 | 62 | 1168-12 | 71 | 1123-6 | 80 |
| 2233-7 | 56 | 204-10 | 66 | 16-8 | 71 | 58-9 | 83 |
| 531-6 | 56 | 561-5 | 66 | 6-10 | 71 | | |
| 2591-6 | 57 | 1812-5 | 67 | 36-14 | 72 | | |
| 894-13 | 57 | 1933-5 | 67 | 1060-8 | 74 | | |

2.1.2.5 Kontrollen für die Analyse genetischer und epigenetischer Aberrationen

In der vorliegenden Arbeit wurden Gewebeproben von Patienten mit einer Lymphadenitis und FACS-sortierte Zellpopulationen von gesunden Individuen aus unterschiedlichen Altersgruppen als Kontrollen verwendet. Die analysierten Normalkontrollen wurden dazu in drei Kohorten eingeteilt und standen für die Analysen von genetischen und epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen zu Vergleichszwecken zur Verfügung. Eine Übersicht der analysierten Normalkontrollen und die damit durchgeführten Methoden ist in der nachfolgenden Abbildung 14 dargestellt (detaillierte Grafik: Abbildung A2 im Anhang).

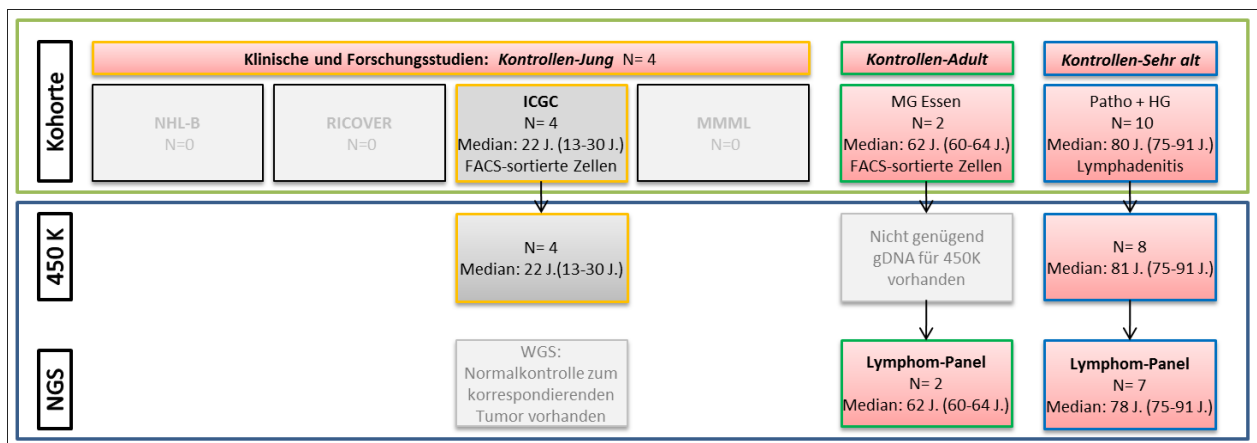


Abbildung 14: Schematische Darstellung der analysierten Normalkontrollen

Im Rahmen der vorliegenden Arbeit wurden Proben von gesunden Individuen als Normalkontrollen für die Analyse von genetischen und epigenetischen Aberrationen bei Keimzentrums-B-Zell-Lymphomen genutzt. 10 Proben von Patienten mit einer Diagnose Lymphadenitis (≥ 75 Jahre) wurden analysiert und dieses Kollektiv im nachfolgenden als *Kontrollen-Sehr alt* bezeichnet. Weiterhin wurden FACS-sortierte B-Zellen (Naive B-Zellen, IgG⁻ und IgM⁺-Gedächtnis-B-Zellen) von 2 gesunden Spendern als *Kontrollen-Adult* zusammengefasst und Hochdurchsatz sequenziert. Zusätzlich existierten Daten von 4 gesunden Spendern aus dem ICGC MMML-Seq Projekt (*Kontrollen-Jung*). Hierzu lagen bereits erhobene DNA-Methylierungsdaten (450K) vor, die innerhalb der vorliegenden Arbeit in den Methylierungsanalysen als Normalkontrollen genutzt wurden. Für die analysierten Proben aus dem ICGC MMML-Seq Projekt standen neben den Tumor-Proben auch die korrespondierenden Normalkontrollen für die Analyse der genetischen Aberrationen zur Verfügung. Abkürzungen: J: Jahre. NGS: Hochdurchsatz-Sequenzierung (engl. Next Generation Sequencing). 450K: Analyse der DNA-Methylierung mittels Infinium HumanMethylation 450K Array. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. RICOVER und NHL-B: Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom.

1.) Normalkontrollen aus dem Kollektiv des Instituts für Pathologie/ Sektion Hämatopathologie (*Kontrollen-Sehr alt*)

Als Normalkontrollen für sehr alte Patienten wurden Gewebe mit einer diagnostizierten Lymphadenitis gewählt, wobei es sich um Proben von Lymphknoten mit einer akuten oder chronischen Infektion handelte. D.h. bei den Individuen mit einer Lymphadenitis war als pathologische Diagnose sichergestellt, dass kein Lymphom sondern eine Entzündung im Gewebe vorlag. Insgesamt wurden zehn Lymphadenitis-Proben von Individuen über 75 Jahre aus dem Untersuchungskollektiv des Instituts für Pathologie ausgewählt und nachfolgend als *Kontrollen-Sehr alt* bezeichnet. Eine Übersicht zu den Charakteristika der einzelnen Lymphadenitis-Proben, die im Rahmen der Altersanalysen verwendet wurden, ist in der Tabelle 9 zusammengestellt.

Tabelle 9: Charakteristika des Untersuchungskollektives *Kontrollen-Sehr alt*

| Fall | Alter [Jahre] | Geschlecht | Lokalisation | Material | 450K | NGS |
|------------|---------------|------------|----------------------|----------|------|-----|
| Jack-NK-1 | 84 | w | inguinal LK | Kryo | x | - |
| Jack-NK-2 | 82 | m | axillär LK | Kryo | x | x |
| Jack-NK-3 | 75 | w | truncus coeliacus LK | Kryo | x | x |
| Jack-NK-4 | 77 | w | retroklavikulär LK | Kryo | x | x |
| Jack-NK-5 | 91 | m | axillär LK | Kryo | x | x |
| Jack-NK-6 | 78 | m | inguinal LK | Kryo | - | x |
| Jack-NK-7 | 81 | m | axillär LK | Kryo | x | - |
| Jack-NK-8 | 76 | m | axillär LK | Kryo | x | x |
| Jack-NK-9 | 81 | w | inguinal LK | Kryo | x | x |
| Jack-NK-10 | 91 | m | axillär LK | FFPE | - | x |

Abkürzungen und Verweise: m: männlich, w: weiblich. Alter: Alter bei Erstdiagnose. LK: Lymphknoten. Kryo: gefrorenes Frischmaterial. FFPE: Formalin-fixiertes Probenmaterial. x: Patienten-Probe wurde mit der Technik in dieser Studie analysiert.-: Patienten-Probe wurde mit der Technik in dieser Studie nicht analysiert. NGS: Hochdurchsatz-Sequenzierung (engl. Next Generation Sequencing). 450K: Analyse der DNA-Methylierung mittels Infinium HumanMethylation 450K Array.

2.) Normalkontrollen aus der AG Molekulargenetik in Essen (*Kontrollen-Adult*)

Als Normalkontrollen für adulte gesunde Individuen wurden FACS-sortierte B-Zellpopulationen (Naive B-Zellen, IgM⁺- und IgG⁺-Gedächtnis-B-Zellen) von zwei Spendern verwendet, wobei es sich um eine 64 Jahre alte Frau und einen 60 Jahre alten Mann handelte. Die Sortierung der B-Zellen aus Vollblut wurde von der Arbeitsgruppe Molekulare Genetik vom Institut für Zellbiologie (Tumorforschung) des Universitätsklinikums Essen durch Frau M.Sc. Martina Przekopowicz durchgeführt. Anschließend wurden mir die gefrorenen Zellpellets der B-Zell-Subpopulationen mit jeweils 10⁵ Zellen zur Verfügung gestellt und die DNA-Extraktion am Institut für Humangenetik durchgeführt (siehe detailliertes Protokoll unter Abschnitt 2.2.1.1). Da es sich bei diesen Normalkontrollen um Proben adulter Probanden handelte, wurde der Proben-Pool der FACS-sortierten B-Zell-Populationen beider Kontrollen als *Kontrollen-Adult* bezeichnet.

3.) Normalkontrollen des ICGC MMML-Seq Projektes (*Kontrollen-Jung*)

Als Normalkontrollen für junge gesunde Individuen wurden FACS-sortierte Zellpopulationen von vier gesunden Individuen (Naive B-Zellen und Gedächtnis-B- oder T-Zellen) verwendet, wobei es sich ausschließlich um Männer im Alter zwischen 13 und 30 Jahren handelte. Für diese gesunden Individuen standen die DNA-Methylierungsdaten von 450K für die FACS-sortierten Zell-Populationen bereits aus dem ICGC MMML-Seq Projekt zur Verfügung. Es ist zu beachten, dass diese Normalkontrollen in der vorliegenden Arbeit ausschließlich zur Analyse der DNA-Methylierung und nicht in der Mutationsanalyse verwendet wurden. Da das Patientenalter dieser gesunden Individuen zwischen 13 und 30 Jahren lag, wurde die Kohorte als *Kontrollen-Jung* bezeichnet. Informationen zum Geschlecht, dem Alter und vorhandener Zell-Population der Normalkontrollen sind in der Tabelle 10 nachzulesen.

Tabelle 10: Charakteristika des Untersuchungskollektives Kontrollen-Jung

| Fall | Geschlecht | Alter [Jahre] | Zellpopulation |
|--------------|------------|---------------|----------------|
| 4118819-B | m | 24 | Naive B-Zellen |
| 4118819-nonB | m | 24 | T-Zellen |
| 4122131-B | m | 19 | Naive B-Zellen |
| 4122131-GCB | m | 19 | GCB-Zellen |
| 4160735-B | m | 13 | Naive B-Zellen |
| 4160735-GCB | m | 13 | GCB-Zellen |
| 4174884-B | m | 30 | Naive B-Zellen |
| 4174884-GCB | m | 30 | GCB-Zellen |

Abkürzungen: m: männlich, w: weiblich. Alter: Alter bei Probenentnahme. GCB-Zellen: Keimzentrums-B-Zellen.

2.1.3 Primer

In der vorliegenden Arbeit wurden Hochdurchsatz-Sequenzierungen mit dem *Two-Step-Assay* (*ID3-Assay*, Details siehe Abschnitt 2.2.4.1) durchgeführt, wobei ein Schwerpunkt dieser Arbeit auch in der technischen Optimierung des Verfahrens lag. Innerhalb der Probenvorbereitung wurden eigene Primer verwendet, die nicht kommerziell verfügbar sind. Diese Target-spezifischen Primer wurden freundlicherweise von Frau Dr. Julia Richter erstellt und der Autorin bereitgestellt. Alle *ID3*-Primer enthalten in der Sequenz zusätzliche Illumina-spezifische Adapter am 5' Ende, die in der nachfolgenden Barcode-PCR als universelle Primerbindungsstellen dienen: Vorwärts-Primer (*5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-Target-Sequenz*) und Rückwärts-Primer (*5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-Target-Sequenz*). Die Amplifikation des *ID3*-Gens (Chr. 1) wurde für verschiedene Polymerase Systeme getestet (Details beschrieben unter Abschnitt 2.2.4.1.2). Aufgrund der Ergebnisse des Polymerasen-Tests wurde der *ID3*-Assay nachfolgend in allen Probenvorbereitungen mit dem AccuPrime Polymerase System durchgeführt. Desweiteren wurden zur Validierung des Lymphom-Panels einzelne mittels Hochdurchsatz-Sequenzierung identifizierte Mutationen mittels Sanger-Sequenzierung analysiert (*TNFRSF14*: Chr. 1, *CREBBP*: Chr. 16, *MYD88*: Chr. 3). Alle Primer, die in der vorliegenden Arbeit verwendet wurden und nicht in den kommerziellen Kits enthalten waren, wurden von der Firma Biomers hergestellt und sind in der nachfolgenden Tabelle 11 aufgelistet. Die verwendeten Primer zur Sanger-Sequenzierung stammen aus der internen Primerdatenbank des Instituts für Humangenetik und wurden von Mitarbeitern vorab bereits optimiert. Die Selektion der Primerpaare erfolgte nach folgenden Kriterien:

- Primerpaar für rekurrent mutiertes Gen liegt in der Datenbank vor und das PCR-Produkte der Primer umspannt die mittels Hochdurchsatz-Sequenzierung detektierten Mutation
- PCR-Produkt ist nicht größer als 400 bp: Es ist bekannt, dass die Fixierung der Gewebe mit Formalin häufig zur Degradierung der Nukleinsäuren führt, was mit einer schlechteren Amplifizierbarkeit von größeren DNA-Fragmenten verbunden ist. Da es sich bei den analysierten Proben um FFPE-Tumormaterial handelte, wurde davon abgesehen PCR-Produkte mit einer Fragmentgröße von über 400 bp zu generieren.

Tabelle 11: Primer, die in der vorliegenden Arbeit zur Sequenzierung verwendet wurden

| M | Gen | Transkript | Region | Sequenz 5'-3' | Start [bp] Ende [bp] | PS Temp.[°C] | Produktgröße [bp] | |
|----------------------|-----------|-----------------|--------|----------------------------|-------------------------|-----------------|----------------------|-----|
| Two-Step-Assay | ID3 | NM-002167 | Ex1 | F: *gcgttgcaaggctcactgtagc | 23885992 | Accu1 55 | 411 | |
| | | | | R: #ctgggtgttcagccctgt | 23885582 | | 350 | |
| | | | | F: *actcactccccagcatgaag | 23885931 | | | |
| | | | | R: #ctgggtgttcagccctgt | 23885582 | | | |
| | KR ID3 | | I1 | F: *tatgatgacaccgctctcca | 23884416 | Accu1 60 | 393 | |
| | | | | R: #ccacatcccctgctagaaa | 23884024 | | 398 | |
| | | | | F: *aggagggtttaaacggcaag | 23885403 | | | |
| | | | | R: #tggctgtacctccagctgat | 23885006 | | | |
| Sanger-Sequenzierung | TNFRSF14 | ENST00000451778 | Ex3 | F: atccecaatgcctgtcct | 2489336 | Accu1 60 | 204 | |
| | | | | R: agaagggggcaagagtgtct | 2489135 | | 250 | |
| | | | Ex4 | F: tagctgtgtctccctcctt | 2489926 | | | 228 |
| | | | | R: ggctgtgctggcctcttac | 2489677 | | | |
| | | | Ex5 | F: tccacgtaccctctcagc | 2491451 | | Ampli 1 55 | 300 |
| | | | | R: gaaatgggaggggtgtcc | 3786584 | | | |
| | CREBBP | NM-004380 | Ex27 | F: taactcctaaaggcagggc | 2491451 | 315 | | |
| | | | | R: aaaaggcacacaaatacctcc | 3788475 | | | |
| | | | Ex26 | F: aatgacagagcaagaccctg | 3786883 | 306 | | |
| | | | | R: ttaaaataccattatttcagg | 3788475 | | | |
| | | | Ex25 | F: ggtgtgcagaagcaccttg | 3789789 | Ampli 2 65 | 329 | |
| | | | | R: gaaggctcacaggctctc | 3789484 | | | |
| | | | Ex16 | F: ctteccgtgaggttgeg | 3817979 | Ampli 2 55 | | |
| | | | | R: tectccatggaatcctaac | 3817651 | | | |
| | MYD88 | NM-002468 | Ex3 | F: acacctgagcatgtgtgat | 38182796 | Ampli 2 60 | 336 | |
| | | | | R: caggcctcagaacagtct | 38182461 | | | |

Abkürzungen und Verweise: Details zur Zusammensetzung der Reagenzien für die unterschiedlichen Reaktionsansätze der Polymerase Systeme sind im Abschnitt 2.2.3.1 nachzulesen M: Verwendete Methode bzw. Assay. Region: Lokalisation der Target-Region im Gen-Kontext. Start/Ende: Start- und Endpunkt der Target-Region in der genomischen DNA (hg19). Temp: Annealing-Temperatur. KR: Kontrollregion. Ex: Exon. I: Intron. F: Vorwärts-Primer. R: Rückwärts-Primer. Accu 1/ 2: Ansatz AccuPrime Polymerase 1 bzw. 2. Ampli 1/ 2: Ansatz AmpliTaq Polymerase 1 bzw.2.

*Vorwärts-Primer: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-Target- Sequenz.

#Rückwärts-Primer: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-Target- Sequenz.

2.1.4 Reagenzien

In der nachfolgenden Tabelle 12 sind die in dieser Arbeit verwendeten Reagenzien und deren Hersteller zusammengetragen.

Tabelle 12: Reagenzien, die in der vorliegenden Arbeit verwendet wurden

| Reagenzien | Hersteller |
|-----------------------------------|---------------------------------------------------|
| 10x TBE Puffer | Life Technologies, Paisley, Scotland |
| 1x Low TE Buffer | Life Technologies, Paisley, Scotland |
| Agarose | Eurogentec, Seraing, Belgien |
| Aqua bidest (dd H ₂ O) | Fresenius Kabi AG, Bad Homburg, Deutschland |
| Betain Lösung | Geyer, Hamburg, Deutschland |
| Bromphenol Blue | Serva, Heidelberg, Deutschland |
| EB Buffer | Qiagen, Hilden, Deutschland |
| Ethanol | Merck KG, Darmstadt, Deutschland |
| Ethidiumbromid | Sigma-Aldrich Chemie GmbH, Steinheim, Deutschland |
| Glycerol | Sigma-Aldrich Chemie GmbH, Steinheim, Deutschland |
| Isopropanol | Geyer, Hamburg, Deutschland |
| Nukleasefreies Wasser | Thermo Fisher, Waltham, MA, USA |
| Roti.-Histol | Carl ROTH, Karlsruhe, Deutschland |
| Tween 20 | Sigma-Aldrich Chemie GmbH, Steinheim, Deutschland |

2.1.5 Enzyme und Kits

In der nachfolgenden Tabelle 13 sind die in dieser Arbeit verwendeten Enzyme und Kits sowie deren Hersteller zusammengetragen. Es ist zu beachten, dass die Produktnamen nicht in die deutsche Sprache übersetzt wurden und in Englischer Sprache den Herstellerangaben entsprechend aufgelistet sind.

Tabelle 13: Enzyme und Kits, die in der vorliegenden Arbeit verwendete wurden

| Enzym/Kit | Hersteller |
|-----------------------------------------------|-----------------------------------------------|
| AccuPrime Polymerase System | Life Technologies, Paisley, Scotland |
| Agencourt AMPure XP Kit | Agencourt, Beverly, Massachusetts, MA, USA |
| AmpliTaq Gold Polymerase | Life Technologies, Paisley, Scotland |
| Big Dye Terminator. V3.1 Cycle Sequencing Kit | Applied Biosystems, Foster City, USA |
| Covaris Mikro-Tubes | Covaris, Woburn, MA, USA |
| Desoxynukleosidtriphosphat Set | Roche Diagnostics GmbH, Mannheim, Deutschland |
| Dynabeads MyOne Streptavidin T1 | Life Technologies, Paisley, Scotland |
| EZ DNA Methylation Kit | Zymo Research, Freiburg, Deutschland |
| FastStart High Fidelity PCR System | Roche Diagnostics GmbH, Mannheim, Deutschland |
| Herculase II Fusion DNA Polymerase | Agilent, Santa Clara, CA, USA |
| High Sensitivity DNA Kit | Agilent, Santa Clara, CA, USA |
| Infinium HD FFPE QC Assay | Illumina, San Diego, USA |
| Infinium HumanMethylation450 BeadChip Kit | Illumina, San Diego, USA |
| MiniElute 96 UF PCR Purification Kit | Qiagen, Hilden, Germany |
| MiSeq Reagent Kit v2, 300 cycles (Standard) | Illumina, San Diego, USA |

| Enzym/Kit | Hersteller |
|-------------------------------------------------|-----------------------------------------------------|
| MiSeq Reagent Kit v2, 500 cycles (Standard) | Illumina, San Diego, USA |
| MiSeq Reagent Kit v3, 600 cycles | Illumina, San Diego, USA |
| MiSeq Reagent Nano Kit v2, 500 cycles | Illumina, San Diego, USA |
| Montage SEQ 96 Sequencing reaction Clean up Kit | Merck KG, Darmstadt, Deutschland |
| Nextera Index Kit | Illumina, San Diego, USA |
| PhiX Control v3 | Illumina, San Diego, USA |
| Phusion High-Fidelity DNA Polymerase | New England Biolabs, Frankfurt am Main, Deutschland |
| PUREGENE DNA Isolation Kit (Gentra Systems) | Qiagen, Hilden, Germany |
| Q5 High –Fidelity DNA Polymerase | New England Biolabs, Frankfurt am Main, Deutschland |
| QIAmp DNA FFPE tissue kit | Qiagen, Hilden, Germany |
| QuantiTect.SYBR.Green PCR Kit | Qiagen, Hilden, Germany |
| Qubit ds DNA HS Assay Kit | Life Technologies, Paisley, Scotland |
| SureSelectXT Custom 0.5 Mb up to 2.9 Mb | Agilent, Santa Clara, CA, USA |
| SureSelectXT Reagent Kit | Agilent, Santa Clara, CA, USA |
| TriDye 100 bp DNA-Längenstandard | New England Biolabs, Frankfurt am Main, Deutschland |
| TruSight Cancer Sequencing Panel | Illumina, San Diego, USA |
| TruSight Rapid Capture Kit | Illumina, San Diego, USA |

2.1.6 Puffer und Lösungen

Innerhalb der vorliegenden Arbeit wurden die nachfolgende Puffer und Lösungen hergestellt.

- Ladepuffer für Agarosegelelektrophorese: Für 100 ml Probenladepuffer wurden 125 mg Bromphenolblau und 28,7 g Glycerol mit Aqua bidest (dd H₂O) auf 100 ml aufgefüllt.
- 1x TBE Puffer: 10x TBE Puffer wurde mit dest. H₂O 1:10 verdünnt.
- 2% Agarosegel: Ein 2%iges Agarosegels wurde aus 4 g Agarose und 200 ml 1x TBE hergestellt und abschließend 10 µl Ethidiumbromid hinzugeben.

2.1.7 Geräte

In der nachfolgenden Tabelle 14 sind die in dieser Arbeit verwendeten Geräte und deren Hersteller zusammengetragen. Es ist zu beachten, dass einige Gerätenamen nicht in die deutsche Sprache übersetzt wurden und in englischer Sprache den Herstellerangaben entsprechend aufgelistet sind.

Tabelle 14: Geräte, die in der vorliegenden Arbeit verwendet wurden

| Geräte | Hersteller |
|----------------------------------------|------------------------------------------------|
| Bioanalyser 2100 | Agilent, Santa Clara, CA, USA |
| BioDoc Analyser | Biometra GmbH, Göttingen, Deutschland |
| Covaris S220 | Covaris, Woburn, MA, USA |
| DynaMag 2 Magnetständer | ThermoScientific, Erlangen, Deutschland |
| Gelkammer | PeqLab Biotechnologie, Erlangen, Deutschland |
| Gradienten-Thermocycler | BioRad, München, Deutschland |
| Heizplatte Hotplate | VWR International GmbH, Darmstadt, Deutschland |
| Hochdurchsatz-Sequenzierer MiSeq | Illumina, San Diego, USA |
| Hybridisierungsöfen | Illumina, San Diego, USA |
| iScan | Illumina, San Diego, USA |
| Kapillarsequenzierer ABI PRISM 3130 | Applied Biosystems, Foster City, USA |
| Laborzentrifuge Megafuge 1.OR | Heraeus Kulzer GmbH, Dormagen, Deutschland |
| Laborzentrifuge Megafuge 40R | Heraeus Kulzer GmbH, Dormagen, Deutschland |
| Laborzentrifuge Universal 16 | Hettich, Tuttlingen, Deutschland |
| Light Cycler 480 II | Roche Diagnostics GmbH, Mannheim, Deutschland |
| Plattenschüttler High-Speed Microplate | Illumina, San Diego, USA |
| Plattenschüttler MS3 basic | IKA, Staufen, Deutschland |
| Qubit 2.0 | Life Technologies, Paisley, Scotland |
| Speedvac RC1010 | Jouan, Saint Herblain, Frankreich |
| Thermocycler T3000 | Biometra GmbH, Göttingen, Deutschland |
| Thermomixer comfort | Eppendorf, Hamburg, Deutschland |
| Vakuumpumpe Vacuum Pump | ThermoScientific, Erlangen, Deutschland |

2.1.8 Software und Datenbanken

In der Tabelle 15 sind die in dieser Arbeit verwendeten Datenbanken und Software-Programme sowie deren Hersteller aufgelistet.

Tabelle 15: Software und Datenbanken, die in der vorliegenden Arbeit verwendet wurde

| Software und Datenbanken | Hersteller/Homepage |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| 1000 Genomes | http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/ |
| Agilent 2100 Expert Software B.02.08 | Agilent, Santa Clara, CA, USA |
| BioDoc Analyzer 2.1 | Applied Biosystems, Foster City, USA |
| CCLE | http://www.broadinstitute.org/ccle/home |
| COSMIC v69 | http://cancer.sanger.ac.uk/cancergenome/projects/cosmic |
| dbSNP (build 141) | http://www.ncbi.nlm.nih.gov/projects/SNP/ |
| DSMZ | http://www.dsmz.de/home.html |
| Ensemble Genome Browser Release 81 | http://www.ensembl.org/index.html |
| GeneCards | http://www.genecards.org/ |
| GenomeStudio Software v2011.1 | Illumina, San Diego, USA |
| Human Genome Variation Society | http://www.hgvs.org/mutnomen/ |
| Illumina Experiment Manager 1.7. | Illumina, San Diego, USA |
| Illumina Variant Studio 2.2 | Illumina, San Diego, USA |
| Integrative Genomics Viewer 2.3 | Broad Institut, Cambridge, Boston, USA |
| intOGen | https://www.intogen.org |
| iScan Control Software 3.3.28 | Illumina, San Diego, USA |
| Lightcycler 480 SW 1.5.1 | Roche Diagnostics GmbH, Mannheim, Deutschland |
| MiSeq-Reporter 2.4.60.9 | Illumina, San Diego, USA |
| Mutalyzer 2.0.11 | https://mutalyzer.nl/ |
| Mutation Mapper | http://www.cbioportal.org/mutation-mapper.jsp |
| MutationTaster | http://www.mutationtaster.org/ |
| Pathway-Analyse | http://string-db.org/ |
| PolyPhen-2 | http://genetics.bwh.harvard.edu/pph2/index.shtml |
| Qlucore Omics Explorer 3.0 | Qlucore, Lund, Schweden |
| R version 3.0.1 | http://r-project.org/ |
| RCSB Protein Data Bank | http://www.rcsb.org |
| Reverse Complement | http://www.bioinformatics.org/sms/rev-comp.html |
| RnBeads (V.0.99.10) | http://rnbeads.mpi-inf.mpg.de/ |
| R-Studio 0.99.473 | http://rstudio.com |
| Seq Pilot 4.2.1 (Tool: SeqNext) | JSI medical systems, Ettenheim, Deutschland |
| Sequencing Analysis Software 6.0 | Applied Biosystems, Foster City, USA |
| Sequencing Analysis Viewer 6.0 | Illumina, San Diego, USA |
| SonoLab 7 | Covaris, Woburn, MA, USA |
| UCSC Genome Bioinformatics | http://genome.ucsc.edu/ |
| Venn-Diagramm | http://bioinfogp.cnb.csic.es/tools/venny/index.html |

2.2 Methoden

Im nachfolgenden Abschnitt werden die in der vorliegenden Arbeit angewandten Methoden erläutert. Das Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel führt molekulargenetische Untersuchungen nach den Bestimmungen der DIN ES ISO 15189:2007 bzw. 2014 durch, sodass die meisten hier verwendeten Protokolle als Standardarbeitsanweisungen (SAA) vorliegen und entsprechend genutzt wurden. Es ist zu beachten, dass in diesem Kapitel die Namen der verwendeten Lösungen/ Reagenzien aus den Kits und spezielle Software- oder Fach-Begriffe nicht in die deutsche Sprache übersetzt wurden und in englischer Sprache entsprechend den Herstellerangaben bzw. der allgemeingültigen Fachsprache benannt sind. Zur Vereinfachung wurden diese Begriffe kursiv dargestellt. Weiterhin sind die zuvor unter Abschnitt 2.1.2 beschriebenen Kohorten (*Jack*, *MMML-Senil* und *Vergleichsstudien*), die im Rahmen dieser Arbeit zusammengestellt wurden, kursiv geschrieben. Eine Übersicht der angewandten Methoden stellt das Fließschema in Abbildung 15 dar, worin beispielhaft der Ablauf der verschiedenen Arbeitsschritte im Rahmen der altersassoziierten Analysen von genetischen und epigenetischen Aberrationen für die Kollektive *Jack* und *MMML-Senil* gezeigt sind. Eine detaillierte Darstellung der verwendeten Methoden und Analysen im Rahmen der vorliegenden Arbeit ist im Anhang in der Abbildung A3 zusammengestellt.

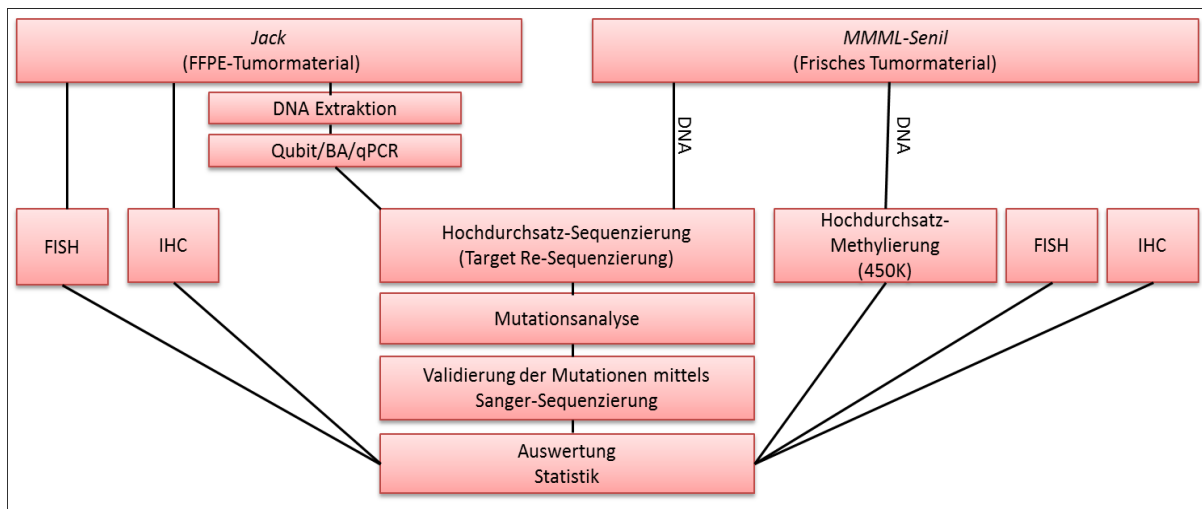


Abbildung 15: Schematische Darstellung des Arbeitsablaufs zur Analyse genetischer und epigenetischer Aberrationen
 Zu allen Patienten der Kohorten *Jack* oder *MMML-Senil* wurden FISH- und IHC-Daten erhoben (sofern ausreichend Tumormaterial in geeigneter Qualität vorhanden war) oder existierten bereits aus vorherigen Studien. Alle molekulargenetischen Daten wurden abschließend mittels Statistik analysiert. Aus dem FFPE-Tumormaterial der Patienten-Kohorte *Jack* wurde DNA extrahiert und die Qualität mittels quantitativer PCR ermittelt. Die extrahierte DNA wurde mit dem Qubit quantifiziert und die Produktgröße der spezifischen DNA-Fragmente nach der Probenanreicherung mit dem Bioanalyzer überprüft. Anschließend wurde ein Set dieser Patienten-Kohorte sowie einige Fälle aus *MMML-Senil* zur Analyse von genetischen Aberrationen Hochdurchsatz sequenziert. Einzelne Mutationen wurden mittels Sanger-Sequenzierung validiert. Parallel dazu wurden Analysen an extrahierter DNA aus frischem Tumormaterial mittels Methylierungsarrays (450K) durchgeführt, um altersassoziierte epigenetische Aberrationen in Keimzentrums-B-Zell-Lymphomen zu identifizieren. Abkürzungen: qPCR: quantitative Realtime PCR. BA: Bioanalyzer. FISH: Fluoreszenz-in-situ-Hybridisierung. IHC: Immunhistochemie, BA: Bioanalyzer, qPCR: Quantitative Realtime PCR. 450K: Infinium HumanMethylation 450K Array *Jack*: Kohorte von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphome, die im Rahmen der vorliegenden Arbeit neu identifiziert wurden (Details siehe unter 2.1.2.2). *MMML-Senil*: Kohorte von senilen Patienten aus dem MMML-Verbundprojekt (Details siehe unter 2.1.2.1).

2.2.1 DNA Extraktion

Die genomische DNA (gDNA) wurde aus Zellpellets und Formalin-fixiertem Gewebe (FFPE) extrahiert. Die dafür verwendeten Protokolle werden im folgenden Abschnitt genauer beschrieben.

2.2.1.1 DNA Extraktion aus Frischmaterial

Die DNA aus den FACS-sortierten B-Zellen wurde mit Hilfe des *Genra PUREgene Kits* (Qiagen) am Institut für Humangenetik in Kiel extrahiert. Dabei wurden folgende Schritte durchgeführt:

Das Tube mit dem Zellpellet wurde aus dem -80°C Gefrierschrank genommen, sofort auf Eis gestellt und 1 ml *Cell Lysis Solution* hinzugegeben. Um eine möglichst gute Zellyse zu gewährleisten, wurde das Tube mehrfach für 10 sec gevortext. Anschließend wurden 5 µl Proteinase K zugegeben, erneut gevortext und die Probe bei 55°C über Nacht bei 200 rpm in einem Thermoschüttler inkubiert. Nach dem Abkühlen der Probe auf RT wurden 5 µl RNase-Lösung hinzugegeben, gevortext und die Probe für 1 h bei 37°C im Thermoschüttler inkubiert. Zum Entfernen der ungewünschten Proteine und Zellbestandteile wurde die Probe erneut auf RT abgekühlt, 350 µl *Protein Precipitation Solution* hinzugegeben und das Tube sofort für mind. 30 min auf Eis gestellt. Anschließend wurde die Probe bei 4500 rpm für 30 min zentrifugiert und nur der klare Überstand in ein neues Tube überführt. Zum Fällen der DNA wurde zum klaren Überstand die gleiche Menge Isopropanol gegeben und das Tube leicht geschwenkt. Die Probe wurde bei 4500 rpm für 30 min zentrifugiert, der Überstand verworfen und das Pellet mit 500 µl 70%igem Ethanol resuspendiert. Nach erneutem zentrifugieren unter 4500 rpm für 30 min wurde der Überstand entfernt und das Pellet an der Luft getrocknet. Das Pellet wurde in 10 µl dest. H₂O aufgenommen und die DNA über Nacht quellen gelassen

2.2.1.2 DNA Extraktion aus FFPE-Tumormaterial

In der Sektion Hämatopathologie des Instituts für Pathologie wurden Paraffinblöcke von Lymphknoten geschnitten. Die Tumorgewebeschnitte (20 µm, 5 Schnitte je Probe) wurden in Eppendorf-Tubes überführt und bei -80°C gelagert. Um das Auftauen des FFPE-Tumormaterials zu verhindern, wurden die Proben von mir auf Trockeneis in das Institut für Humangenetik transportiert und bei -80°C gelagert. Anschließend wurde unter Durchführung der nachfolgenden Arbeitsschritte die DNA mittels *QIAmp DNA FFPE tissue kit* (Qiagen) aus dem FFPE Material isoliert:

Zum Tube mit den Tumorgewebeschnitten wurde 1 ml Roti-Histol hinzugegeben und ca. 10 sec gevortext, um das Paraffin vom Gewebe zu entfernen. Um eine möglichst gute Trennung des Zellpellets vom Paraffin zu gewährleisten, wurde das Tube 2 min bei RT und 13.000 rpm zentrifugiert und der Überstand abgenommen. Es wurde erneut 1 ml Roti-Histol hinzugegeben, der Zentrifugationsschritt wiederholt und der Überstand abgenommen.

Das Zellpellet wurde mit 1 ml 100% Ethanol gewaschen, um das Roti-Histol von der Probe zu entfernen und nachfolgend 1 min bei RT und 13.000 rpm zentrifugiert. Anschließend wurde der Überstand abgenommen. Das Pellet wurde erneut mit 1 ml 100% Ethanol gewaschen und der Überstand restlos abgenommen und anschließend das Zellpellet bei geöffnetem Tube und RT stehen gelassen bis der Alkohol verdampft war. Das Pellet wurde in 360 µl ATL Puffer resuspendiert, 40 µl Proteinase K hinzugeben und über Nacht bei 56°C und leichtem schütteln im Thermoblock inkubiert. 200 µl AL Puffer wurden hinzugeben, gevortext, 200 µl 100% Ethanol hinzugeben, gevortext und alles auf eine *MinElute-Säule* aus dem Kit pipettiert. Es wurde 1 min bei RT und 8.000 rpm zentrifugiert, der Durchfluss verworfen und die Säule in ein neues Tube gestellt. 500 µl AW1 Puffer wurden auf die Säule geben, 1 min bei RT und 8.000 rpm zentrifugiert, der Durchfluss verworfen und die Säule in ein neues Tube stellt. 500 µl AW2 Puffer wurden auf die Säule geben, 1 min bei RT und 8.000 rpm zentrifugiert, der Durchfluss verworfen und die Säule in ein neues Tube gestellt. Die Säule wurde für ca. 3 min bei 14.000 rpm zentrifugiert, um die Membran zu trocknen und anschließend in ein neues Tube gesetzt. 30 µl dest. H₂O wurden auf die Säule geben, 1 min bei RT inkubiert und für 1 min bei ca. 14.000 rpm zentrifugieren. Die Säule wurde auf ein neues Tube gestellt und eine 2. Elution mit 30 µl dest. H₂O durchgeführt

2.2.2 Qualitätskontrolle der DNA- und PCR-Fragmente

Innerhalb der vorliegenden Arbeit wurden verschiedene Methoden zur Quantifizierung und Qualitätskontrolle der DNA- und PCR-Fragmente genutzt. Die dafür angewandten Protokolle sind im nachfolgenden Abschnitt dargestellt.

2.2.2.1 Quantifizierung der DNA und PCR-Fragmente mittels Qubit

Die extrahierte DNA und verschiedene PCR-Produkte wurden mittels Qubit 2.0 (Life Technologies) quantifiziert. Der Qubit ist ein Fluorometer zur Bestimmung der Nukleinsäuren-Konzentration von Lösungen mit Hilfe von Fluoreszenz. Der Farbstoff bindet an die Nukleinsäure in der Lösung, wodurch direkt ein Fluoreszenzsignal ausgestrahlt wird. Zur Konzentrationsbestimmung wird die detektierte Fluoreszenz mit einer Standardkurve abgeglichen. Zur Quantifizierung der DNA wurde in der vorliegenden Arbeit das *Qubit DNA HS Assay Kit* (Life Technologies) entsprechend den Herstellerangaben verwendet.

2.2.2.2 Qualitätskontrolle der extrahierten DNA mittels quantitativer Real Time PCR

Die Qualität der extrahierten DNA aus dem FFPE-Tumormaterial (Kohorte *Jack*, Details siehe unter Abschnitt 2.1.2.2) wurde mit Hilfe des *FFPE QC Kit* (Illumina) unter Verwendung des Verfahrens der quantitativen Realtime PCR (qPCR) geprüft. Das Verfahren basiert auf der Detektion von Fluoreszenzsignalen wobei Fluorophore an die DNA binden und anschließend gemessen werden. Das Fluoreszenzsignal verhält sich hierbei proportional zur DNA-Konzentration, was eine Quantifizierung der DNA ermöglicht. Hierbei wurde die Amplifikation der DNA mit einem Illumina-spezifischen Primermix (QCP) gemessen, wobei sich nach Herstellerangaben die Konzentration an PCR-Produkt proportional zur DNA-Qualität verhalten soll. Zur Bestimmung der Qualität wurde im Kit zusätzlich ein Illumina-Standard (QCT) mitgeliefert und analysiert. Die abschließende Qualitätsbewertung der DNA wurde über die ermittelten ct-Werte der qPCR vorgenommen, wobei der ct-Wert den Zyklus beschreibt, an dem die Fluoreszenz erstmals signifikant über dem Hintergrundsignal detektiert wird. Die Differenz der ct-Werte des Standards und der zu analysierenden DNA-Probe ergab den Qualitätswert. Die Durchführung der qPCR erfolgte in Triplikaten entsprechend der Herstellerangaben des *FFPE QC Kits*. Die wichtigsten Arbeitsschritte sind im nachfolgenden kurz beschrieben:

In einem Eppendorf-Tube wurden 10 µl unverdünnter QCT Reaktionsmix (*Infinium HD FFPE QC Assay*, Illumina) mit 990 µl dest. H₂O gemischt und das Tube bei 280 xg an zentrifugiert. Jeweils 4 µl von QCP Kontroll-DNA (*FFPE QC Kit*, Illumina), gDNA (von jeder zu analysierenden Probe) und dest. H₂O wurden in 3 Wells einer 96-Well-Platte pipettiert. Anschließend wurden zu jeder Probe 16 µl qPCR-Mastermix hinzugegeben. Der qPCR-Mastermix für 1 Amplifikation wurde wie folgt erstellt:

| Reagenzien | Volumen |
|-----------------------------------------|------------------------|
| Dest. H ₂ O | 4 µl |
| 2x QuantiTect SYBR Green PCR Master Mix | 10 µl |
| QCT (<i>FFPE QC Kit</i> , Illumina) | 2 µl |
| Gesamt | 16 µl/ Reaktion |

Die Platte wurde mit einer Klebefolie verschlossen und für 1 min bei 280 xg zentrifugiert. Anschließend wurde die Platte in den *LightCycler 480 II* der Firma Roche gestellt und das nachfolgende qPCR-Programm gestartet.

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|--------|------------|--------|
| Denaturierung | 2 min | 50°C | 1x |
| Annealing | 10 min | 95°C | 40x |
| | 10 sec | 95°C | |
| Elongation | 5 sec | 57°C | |
| Abkühlung | 4 min | 72°C | 1x |
| | ∞ | 4°C | |

Zur groben Abschätzung der Degradierung wurde der ct-Wert jeder Probe betrachtet, der einen linearen Zusammenhang zum Qualitätswert darstellt. D.h. je kleiner der gemessene ct-Wert war desto besser war die DNA-Qualität, da die Primer schneller an die gDNA gebunden haben und die Konzentration an PCR-Produkt höher war. Bei stark degradiertem DNA ist davon auszugehen, dass es nicht möglich ist in der qPCR die gDNA zu amplifizieren (kein ct-Wert messbar). Die Firma Illumina sowie die aktuelle Literatur gaben keine konkreten Grenzwerte und Angaben zum Zusammenhang zwischen dem gemessenen Qualitätswert und der Qualität der erhobenen Hochdurchsatz-Sequenzierungs-Daten der extrahierten Proben. Daraufhin wurde ein willkürlicher Grenzwert von 15 als Einschlusskriterium für die FFPE-Probe in die weiteren Analysen festgelegt.

2.2.2.3 Qualitätskontrolle und Quantifizierung von PCR-Fragmenten mittels Bioanalyzer

Der Agilent 2100 Bioanalyzer ist ein Gelelektrophorese System zur Bestimmung von Produktgröße, Qualität und Konzentration von Nucleinsäuren. Basierend auf der Mikrofluidtechnik werden die Proben und der Standard (Leiter) in einen Chip pipettiert und in Mikrokanälen elektrophoretisch der Größe nach aufgetrennt. Die Nucleinsäuren werden durch eingelagerte Fluoreszenzfarbstoffe in der Gelmatrix detektiert und in jedem Lauf gegen einen mitgeführten Standard abgeglichen. Für jeden Typ von Molekül wird das zugehörige Kit inkl. eines spezifischen Standards verwendet. Abschließend wird in der Expert Software der Firma Agilent die Produktgröße sowie die Konzentration der gemessenen Nucleinsäuren grafisch in Form von Elektropherogrammen und Gelbildern ausgegeben. In der vorliegenden Arbeit wurde zur Qualitätsbestimmung und Quantifizierung der PCR-Produkte das *HS DNA Kit* (Agilent) entsprechend der Herstellerangaben verwendet. Mit Hilfe dieses Kits ist es möglich zeitgleich 11 Proben zu analysieren und die DNA-Fragmentgrößen im Bereich von 50 bis 7000 bp sowie die Nucleinsäuren-Konzentrationen zwischen 5 bis 500 pg/µl DNA zu messen. Die wesentlichen Arbeitsschritte sind nachfolgend kurz beschrieben. Zur Herstellung der Gelmatrix wurden 15 µl Farbstoff-Konzentrat auf die Gelmatrixsäule pipettiert. Das Tube mit der Gelmatrixsäule wurde für 15 min bei 2240 xg und RT zentrifugiert. Die Säule wurde verworfen und das Tube mit Aluminiumfolie ummanteln, um den Fluoreszenzfarbstoff vor Lichteinstrahlungen zu schützen. Der Chip wurde in die Auflageplatte der *Chip priming station* gelegt und 9 µl Gelmatrix in das fett markierte G-Well pipettiert. Die Injektionsspritze wurde heruntergedrückt, die *Chip priming station* für 1 min geschlossen und danach die Injektionsspritze wieder gelöst. Es wurde gewartet bis die Spritze bei etwa bei 0,7 ml war und dann langsam bis 1 ml hochgezogen. Die *Chip priming station* wurde geöffnet und der Chip aus der Auflageplatte genommen. 9 µl Gelmatrix wurden in die 2 weiteren G-Wells pipettiert. Danach wurden 5 µl Marker (grüner Tubendeckel) in alle Proben-Wells und das Leiter-Well überführt. Ggf. nicht zu benutzende Proben-Wells wurden mit 6 µl Marker befüllt. 1 µl Leiter (gelber Tubendeckel) wurde in das Leiter-Well und je 1 µl Probe in die Proben-Wells pipettiert. Der Chip wurde in den Vortexer gelegt und für 1 min bei 2400 rpm geschüttelt.

Der Bioanalyser wurde geöffnet, der Chip einlegt und der Deckel wieder geschlossen. Parallel wurde die Software *Expert 2100* am Computer geöffnet und die Kategorie *Instrument* geöffnet. Unter der Kategorie *Assay* → *Electrophoreses* wurde das analysierte Kit *DNA HS* ausgewählt und die Probenamen unter *Sample Name* in die Tabelle eintragen. Abschließend wurde der Lauf mit der Bestätigung des grünen (Pfeil-)Buttons gestartet. Die Messergebnisse wurden in Form eines Reports (.pdf-Datei) gespeichert und enthielten die Informationen zu den DNA-Fragmentgrößen und der gemessenen DNA-Konzentration von den analysierten Proben.

2.2.3 Klassisches Verfahren der Sanger-Sequenzierung

Zur Validierung der mittels Hochdurchsatz-Sequenzierung identifizierten Mutationen (detailliert unter Abschnitt 2.2.4 beschrieben) wurde eine Sanger-Sequenzierung durchgeführt²²⁸. Dieses Verfahren basiert auf dem Kettenabbruch nach dem Einbau von Desoxynukleotiden, die heutzutage häufig fluoreszenzmarkiert sind. Die Methode funktioniert wie eine herkömmliche PCR, wobei eine DNA-Matrix, mit Hilfe einer DNA-Polymerase, durch den Einbau von Desoxynukleotiden (dNTP) erzeugt wird. Zusätzlich werden bei der Sanger-Sequenzierung fluoreszenzmarkierte Didesoxynukleotide (ddNTP) in die DNA-Matrix eingebaut, die bei der Elongation die Ausbildung weiterer Phosphodiesterbindungen verhindern und einen Abbruch der Kettenreaktion hervorrufen. Ein Überschuss an dNTP im Verhältnis zu ddNTP in der Lösung führt statistisch gesehen dazu, dass die Abbruchreaktionen an allen Basenpositionen der zu sequenzierenden DNA-Fragmente auftreten. Es entstehen somit Sequenzen verschiedener Länge, die jeweils mit einem spezifischen ddNTP enden. Aufgrund der Verwendung von vier unterschiedlichen Fluoreszenzmarkierungen der ddNTP (ddATP, ddCTP, ddGTP, ddTTP) können die einzelnen Basen mit einem Laser angeregt und anschließend detektiert werden. Zu Beginn der Mutationsanalyse mittels Sanger-Sequenzierung war es notwendig die Target-spezifischen Region in der zu analysierenden Probe anzureichern. Hierzu wurde eine Polymerasekettenreaktion (PCR) mit Target-Spezifischen Primern vorgenommen und eine Gelelektrophorese durchgeführt, um die Fragmentgröße der Amplifikate zu überprüfen. Anschließend wurden die Target-spezifischen DNA-Fragmente in einer Sequenzierungs-PCR amplifiziert und nachfolgend sequenziert. Die durchgeführten Arbeitsschritte zur Sanger-Sequenzierung einer DNA-Probe werden im nachfolgenden kurz beschrieben.

2.2.3.1 Polymerasekettenreaktion und Agarose-Gelelektrophorese

Zur Amplifikation von genomischer DNA wurde das *AmpliTaq Gold Polymerase* bzw. *AccuPrime Polymerase System* verwendet.

1.) Anreicherung der Target-spezifischen DNA-Fragmente mittels PCR

Für jede Amplifikation wurde 1 µl von 100 ng gDNA mit 24 µl PCR-Mastermix gemischt. Die Zuordnung der Polymerasen/Ansätze zu den Primerpaaren wurde im Material-Teil der vorliegenden Arbeit unter Abschnitt 2.1.3 zusammengestellt. Der Mastermix für eine Reaktion wurde wie folgt erstellt:

| Reagenzien | AmpliTaq Gold | | AccuPrime | |
|-------------------------------|-------------------|-------------------|-------------------|-------------------|
| | Ansatz 1 | Ansatz 2 | Ansatz 1 | Ansatz 2 |
| H ₂ O | 17,3 µl | 16,8 µl | 20,0 µl | 18,7 µl |
| Puffer (10x/ 5x) | 2,5 µl | 2,5 µl | 2,5 µl | 2,5 µl |
| MgCl ₂ (25mM) | 2,0 µl | 2,5 µl | - | - |
| Betain | - | - | - | 1,3 µl |
| dNTPs (5 mM pro NTP) | 1,0 µl | 1,0 µl | - | - |
| Vorwärts-Primer (20 pmol/µl) | 0,5 µl | 0,5 µl | 0,5 µl | 0,5 µl |
| Rückwärts-Primer (20 pmol/µl) | 0,5 µl | 0,5 µl | 0,5 µl | 0,5 µl |
| Polymerase | 0,2 µl | 0,2 µl | 0,5 µl | 0,5 µl |
| Gesamt | 24 µl/ Rkt | 24 µl/ Rkt | 24 µl/ Rkt | 24 µl/ Rkt |

Rkt: Reaktion

Anschließend wurden die PCR-Tubes in den Thermocycler (Biometra) gestellt und das Polymerase zugehörige PCR-Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|--------|---------------------------------------|--------|
| Denaturierung | 5 min | 95°C | 1x |
| Annealing | 30 sec | 95°C | 36x |
| | 30 sec | 55° C / 60 C °/ 65°C* | |
| Elongation | 30 sec | 68°C ¹ / 72°C ² | |
| Abkühlung | 5 min | 68°C ¹ / 72°C ² | 1x |
| | ∞ | 4°C | |

*Abhängig von der Annealing-Temperatur der eingesetzten Primer (siehe unter Abschnitt 2.1.3, Tabelle 12). ¹: Elongationstemperatur für das *AccuPrime Polymerase System*. ² Elongationstemperatur für das *AmpliTaq Gold Polymerase System*.

Die PCR-Produkte wurden vor der Sequenzierungs-PCR mit dem *MiniElute 96 UF PCR Purification Kit* (Qiagen) aufgereinigt. Hierzu wurde das PCR-Produkt mit 40 µl dest. H₂O gemischt und auf die Filtermembran pipettiert. Mit Hilfe einer Vakuumpumpe wurde ein Unterdruck erzeugt (23-25 mmHG) und die Flüssigkeit durch die Filtermembran gezogen. 50 µl dest. H₂O wurden auf die Membran mit den gebundenen PCR-Produkten pipettiert und die Flüssigkeit erneut unter Vakuum durch die Membran gezogen. Abschließend wurde das aufgereinigte Amplifikat durch schütteln der Platte mit jeweils 20 µl H₂O gelöst, abpipettiert und jeweils in ein neues Tube überführt.

2.) Agarose-Gelelektrophorese

Zur Überprüfung der erfolgreichen Amplifikation der DNA wurde eine Gelelektrophorese, eine Methode zur Auftrennung von DNA-Molekülen unterschiedlicher Größe und Ladung in einem elektrischen Feld, durchgeführt. Ein 2% iges Agarosegel wurde mit 1x TBE-Puffer angefertigt (siehe Abschnitt 2.1.6). 3,5 µl Ladepuffer wurden mit jeweils 5 µl PCR-Produkt gemischt und in eine Geltasche pipettiert. Zur anschließenden Bestimmung der PCR-Produktlänge wurde ein 100 bp DNA-Leiter (NewEngland Biolabs) mit aufgetragen. Die Gelelektrophoresekammer wurde mit 1x TBE-Puffer befüllt und an ein Spannungsgerät bei 180 V angeschlossen. Die abschließende Geldokumentation erfolgte mit dem *BioDoc Analyzer* der Firma Biometra.

2.2.3.2 DNA Sequenzierung nach Sanger

Die aufgereinigten PCR-Produkte wurden mittels Sequenzierungs-PCR amplifiziert, aufgereinigt und abschließend mit einem Vier-Kapillarsequenzierer (ABI Prism 3130 Genetic Analyzer) analysiert. Die dazu durchgeführten Arbeitsschritte werden in den nachfolgenden Abschnitten kurz erläutert.

1.) Sequenzierungs-PCR

Zur Sanger-Sequenzierung der PCR-Produkte wurde zuerst eine Sequenzierungs-PCR der aufgereinigten PCR-Produkte mit dem *BigDye Terminator v3.1 Cycle Sequencing Kit* der Firma Applied Biosystems nach Herstellerangaben durchgeführt. Die Sequenzierung erfolgte für jede Probe in zwei Ansätzen wobei jeweils eine Reaktion mit dem Vorwärts- bzw. dem Rückwärts-Primer durchgeführt wurde. Hierzu wurden jeweils 1 µl aufgereinigtes PCR-Produkt und 19 µl Mastermix pro Sequenzierreaktion gemischt. Für den Mastermix wurden die Reagenzien für 1 Reaktion wie folgt zusammenpipettiert:

| Reagenzien | Volumen |
|-----------------------------------------------------|----------------------------------|
| Dest. H ₂ O | 15 µl |
| Vorwärts- <u>oder</u> Rückwärts-Primer (20 pmol/µl) | 1 µl |
| <i>Big Dye Terminator (BDT)</i> | 1 µl |
| 5x Puffer | 2 µl |
| Gesamt | 19 µl/ Sequenzierreaktion |

Die PCR-Tubes wurden in den Thermocycler (Biometra) gestellt und das PCR-Programm zur Amplifikation mit den nachfolgenden Parametern gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|------------|-------------|--------|
| Denaturierung | 2 min | 95°C | 1x |
| Annealing | 10 sec | 96°C | 25x |
| Elongation | 5 sec | 57°C | |
| Abkühlung | 4 min ∞ | 60°C 4°C | 1x |

Nachdem die Proben auf 4°C abgekühlt waren, wurden die PCR-Tubes aus dem Cycler genommen und bis zum nächsten Schritt bei 4°C im Kühlschrank gelagert.

2.) Aufreinigung der Sequenzierungs-PCR-Produkte

Die Sequenzierungs-PCR-Produkte wurden vor der Sanger-Sequenzierung mit dem *Montage SEQ 96 Sequencing Reaction Cleanup Kit* (Merck) nach Herstellerangaben aufgereinigt. Hierzu wurde das Sequenzier-PCR-Produkt mit 20 µl Injektionslösung aus dem Kit gemischt und auf die Filterplatte pipettiert. Die Flüssigkeit wurde unter Einsatz eines Vakuums (23-25 mmHG) durch die Filtermembran gezogen, die gebundenen PCR-Produkte mit 20 µl Injektionslösung gewaschen und abschließend durch schütteln der Platte in 20 µl dest. H₂O gelöst, abpipettiert und jeweils in ein neues Tube überführt.

3.) Sanger-Sequenzierung und Datenanalyse

Zur Sanger-Sequenzierung wurden 3 µl aufgereinigtes Sequenzierungs-PCR-Produkt mit 30 µl dest. H₂O verdünnt und in eine 96-Well-Sequenzier-Platte pipettiert. Anschließend wurde die Sequenzier-Platte in den Vier-Kapillarsequenzierer (ABI Prism 3130 Genetic Analyzer) gestellt und ein Standard-Programm mit einer Injektionszeit von 10 sec und einer Laufzeit von 18 min ausgewählt. Die Analyse der Sequenzierungs-Rohdaten erfolgte mit der *Sequencing Analysis Software Version 6.0*. Anschließend wurden die Sequenzen manuell gegen das Referenzgenom hg19 abgeglichen und die identifizierten Mutationen entsprechend der HGVS-Mutationsnomenklatur angegeben²¹⁷. Die mittels Sanger-Sequenzierung identifizierten Mutationen wurden mit den detektierten Aberrationen aus der Hochdurchsatz-Sequenzierung verglichen. Alle Proben, die die gleiche Mutation in der Hochdurchsatz- und Sanger-Sequenzierung zeigten, wurden als erfolgreich validiert angesehen.

2.2.4 Hochdurchsatz-Sequenzierung

Die Hochdurchsatz-Sequenzierung ermöglicht es, gleichzeitig tausende Nukleinsäure-Fragmente parallel zu sequenzieren, um Sequenzveränderungen (z.B. Mutationen und strukturelle Aberrationen wie z.B. Deletionen) zu identifizieren. Mit Hilfe dieser modernen Technologie können mittlerweile ganze Genome, Transkriptome, Epigenome, Exome oder auch nur bestimmte genetische Regionen bzw. Gene analysiert werden. Letzteres wird als Panel-Analyse bezeichnet und findet durch die gezielte Analyse genetischer Regionen sowie Zeit- und Kostenersparnis immer mehr Anwendung in der Diagnostik. Teil der vorliegenden Arbeit war die Optimierung der Hochdurchsatz-Sequenzierung am Institut für Humangenetik, um diese neue Technologie anschließend für die Fragestellung der altersassoziierten Analysen der genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen nutzen zu können. Da das Institut bereits über einen Hochdurchsatz-Sequenzierer (MiSeq) der Firma Illumina verfügte, wurde die Optimierung unter Anwendung des MiSeq Geräts vorgenommen. Dabei wurden verschiedene Anreicherungsverfahren zur Analyse von spezifischen genetischen Target-Regionen ausgetestet. Hierbei gibt es verschiedene Möglichkeiten eine Anreicherung von Target-Regionen vorzunehmen¹⁷⁹, wobei vornehmlich in zwei Anreicherungsprozesse unterschieden wird: erstens.) PCR-basiert und zweitens.) Hybridisierungs-basiert. Das Prinzip der einzelnen Anreicherungsverfahren ist in der Einleitung genauer beschrieben worden und kann dort unter Abschnitt 1.6.1 nachgelesen werden. Beide Verfahren wurden im Rahmen der vorliegenden Arbeit am Institut für Humangenetik getestet und sind in der Abbildung 16 als schematischer Überblick im Zusammenhang mit den angewandten Panel/Assay dargestellt. Hierbei ist zu beachten, dass zur Probenvorbereitung neben kommerziell erhältlichen Panel der Firma Illumina auch eigens entworfene Panel/Assay verwendet wurden. Diese eigens entworfenen Panel/Assay werden im Folgenden als *Custom Panel* bezeichnet und wurden im Rahmen der vorliegenden Arbeit neu entworfen bzw. die Probenvorbereitung optimiert.

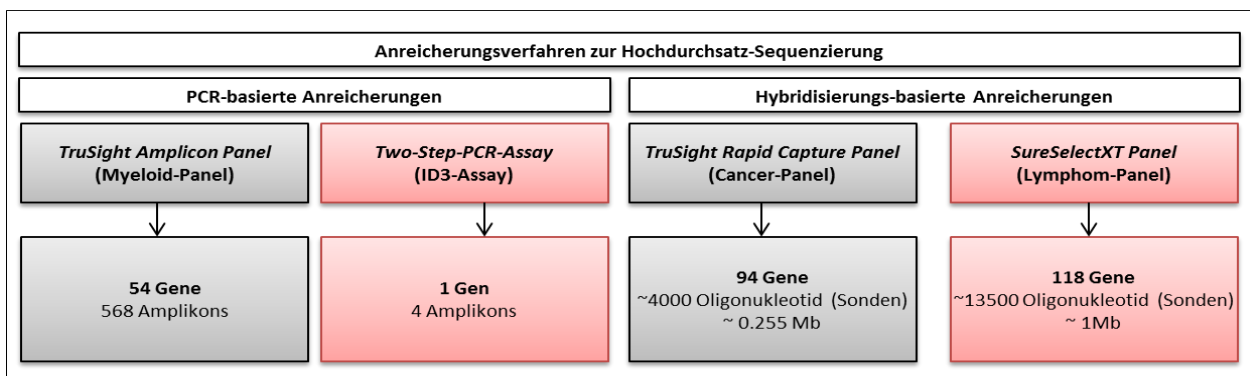


Abbildung 16: Darstellung zu den verwendeten Anreicherungsverfahren und Panel/Assay

In der vorliegenden Arbeit wurden PCR- und Hybridisierungs-basierte Anreicherungsverfahren zur Herstellung der Library im Rahmen der Methoden-Optimierung der Hochdurchsatz-Sequenzierung durchgeführt. Dazu wurde das kommerziell erhältliche *TruSight Amplicon Panel* (Myeloid-Panel), das eigens entworfene *Two-Step-PCR-Assay* (ID3-Panel), das *TruSight Rapid Capture Panels* (Cancer-Panel) sowie ein *inhouse* entworfenes *SureSelectXT Panel* (Lymphom-Panel) getestet und optimiert. In der Grafik sind die kommerziell erhältlichen Panel grau und die eigens entworfenen Panel (*Custom Panel*) rot gekennzeichnet.

Die PCR-basierten Methoden wurden für die Mutationsanalysen mittels *TruSight Amplicon Panel* (Myeloid-Panel, kommerziell erhältliches Panel) und *Two-Step-PCR-Assay* (*ID3-Assay*, Custom Assay) verwendet. Das Hybridisierungs-basierte Anreicherungsverfahren wurde zur Analyse genetischer Veränderungen mit dem *TruSight Rapid Capture Panel* (Cancer-Panel, kommerziell erhältliches Panel) und *SureSelect XT Panel* (Lymphom-Panel, Custom Panel) genutzt. Detaillierte Beschreibungen zu den verwendeten Panel/Assay sind in den nachfolgenden Kapiteln separat aufgeführt. Zur Vereinfachung der Thematik wurde versucht möglichst analoge Begriffe in der Erläuterung der Probenvorbereitung der verschiedenen Panel/Assay für die hergestellten DNA-Lösungen (*Target-Spezifische DNA-Fragmente*, *Proben-Pool von Target-Spezifischen DNA-Fragmenten*, *Barcode-Proben-Pool*, *Library*) und die durchgeführten Arbeitsschritte (*Amplifikation spezifischer genomischer Target-Regionen*, *Amplifikation mit Barcode und Sequenzier-Adapter*, *Herstellung der Library*, *Sequenzierung mit dem MiSeq*) zu verwenden. Aufgrund der Komplexität und der unterschiedlichen Panel und Anreicherungsverfahren wurden teilweise weitere Begriffe benötigt, um die Probenvorbereitung der Panel/Assay detailliert zu beschreiben (z.B. *Fragmentierung der genomischen DNA*). Zur Übersicht wurden die verwendeten Begriffe sowohl für PCR- als auch Hybridisierungs-basierte Verfahren beispielhaft an die eigens entworfenen Custom Panels dargestellt und sind in den jeweiligen Kapiteln zur Beschreibung der einzelnen Panel/Assays wiederzufinden (PCR-basierte Verfahren siehe Abbildung 17 unter Abschnitt 2.2.4.1, Hybridisierungs-basierte Verfahren siehe Abbildung 18 unter Abschnitt 2.2.4.2).

2.2.4.1 PCR-basierte Anreicherungsverfahren

In der vorliegenden Arbeit wurden zwei verschiedene PCR-basierte Anreicherungsverfahren zur Probenvorbereitung verwendet. Hierzu wurden am Institut für Humangenetik zum einen das *TruSight Amplicon Panel* (Myeloid-Panel, kommerziell erhältliches Panel) und zum anderen der *Two-Step-PCR-Assay (ID3-Panel, Custom Assay)* getestet und optimiert. Die einzelnen PCR-basierten Panel/Assay werden nachfolgend genauer beschrieben.

2.2.4.1.1 *TruSight Amplicon Panel (Myeloid-Panel)*

Das *TruSight Amplicon Panel* (Myeloid-Panel) beruht auf einem PCR-basierten Anreicherungsverfahren und ermöglicht es somatische Mutationen in Genen, die bekanntermaßen häufig in hämatologischen Neoplasien mutiert vorliegen, zu identifizieren. Das Panel besteht aus 568 Amplikons die komplette Exone oder teilweise nur spezifische Regionen von insgesamt 54 Genen abdecken, die rekurrent mutiert in Leukämien und myeloproliferativen Erkrankungen vorliegen¹⁸⁰. Insgesamt wurde im Rahmen dieser Arbeit das Mutationsprofil von 39 T-PLL-Patienten mit dem *TruSight Amplicon Panel* (Myeloid-Panel) der Firma Illumina analysiert. Bei dem PCR-basierten Verfahren der Firma Illumina erfolgt zunächst eine Hybridisierung mit einem Pool von Oligonukleotiden, die die gewünschte Target-Region in der genomischen DNA der Patienten umspannen. Es erfolgt eine lineare Amplifikation zur Vervielfältigung der Target-Regionen (*Proben-Pool von Target-spezifischen DNA-Fragmenten*). Daraufhin wird eine Amplifikation durchgeführt bei dem jedem *Proben-Pool* ein individueller Barcode und die Sequenzier-Adapter zugeordnet werden. Abschließend werden alle *Barcode-Proben-Pools* zur *Library* zusammenpipettiert und mit dem MiSeq Gerät von Illumina sequenziert. Die Probenvorbereitung wurde nach Herstellerangaben mit dem Protokoll *TruSight DNA Amplicon Library Preparation Guide* durchgeführt. Die durchgeführten Arbeitsschritte sind nachfolgend genauer beschrieben. Sofern keine Herstellerangaben angegeben wurden, stammten die verwendeten Reagenzien aus dem *TruSight Myeloid Kit* (Illumina).

1.) *Amplifikation spezifischer genomischer Target-Regionen*

Neben der genomischen DNA (gDNA) von den zu analysierenden Patienten-Proben wurde eine Illumina-spezifische Kontroll-DNA (ACD1) bei der Library Präparation mitgeführt. Bei dieser Kontroll-DNA-Probe handelte es sich um gDNA, die zu Kontrollzwecken als Positivkontrolle mitgeführt wurde und die Fehleranalyse bei eventuellen Problemen in der Probenvorbereitung erleichtert. 5 µl Kontroll-DNA (ACD1) und 5 µl dest. H₂O wurden in ein Well einer 96-Well-Platte pipettiert. Jeweils 50 ng Patienten-DNA wurden auf 10 µl mit dest. H₂O aufgefüllt und jede verdünnte Patienten-DNA-Probe in ein Well der 96-Well-Platte überführt.

In der weiteren Probenvorbereitung wird die Kontroll-DNA-Probe analog den zu analysierenden Patienten-DNA-Proben durchgeführt. Nachfolgend umfasst der Begriff "Probe" sowohl die Kontroll- als auch Patienten-DNA-Proben.

Zu jeder verdünnten DNA-Probe wurden 5 µl TSO hinzugegeben, wobei es sich um die Target-spezifischen Oligonukleotide handelte. Die Platte wurde mit einer Klebefolie verschlossen und für 1 min bei 20°C und 1000 xg zentrifugiert. Die Folie wurde entfernt und 35 µl OHS2-Lösung zu jeder Probe pipettiert. Die Proben wurden erneut für 1 min bei 20°C und 1000 xg zentrifugiert und dann in einem vorgeheizten Heizblock bei 95°C für 1 min inkubiert. Die Temperatur vom Heizblock wurde auf 40°C reduziert und die Proben für weitere 80 min inkubiert. Innerhalb dieser Zeit kühlte der Heizblock kontinuierlich von 95°C auf 40°C herunter. Nach der Inkubationszeit wurde sichergestellt, dass der Heizblock eine Temperatur von 40°C hatte. Danach wurde die 96-Well-Platte aus dem Heizblock genommen, 1 min bei 20°C und 1000 xg zentrifugiert. Die jeweilige Proben-DNA-Oligo-Lösung wird im Weiteren als *Target-spezifische DNA-Fragmente* bezeichnet. Da die erste Amplifikation in der Probenvorbereitung mittels *TruSight Amplicon Panel* (Myeloid-Panel) eine Multiplex-PCR beinhaltet, müssen die Target-spezifischen DNA-Fragmente der einzelnen Proben nicht mehr zusammengeführt werden und die Target-spezifischen DNA-Fragmente liegen bereits zu diesem Zeitpunkt der Probenvorbereitung als *Proben-Pool* vor.

45 µl SW1-Lösung wurden in jedes Wells der Illumina-spezifischen FPU Filterplatte pipettiert, die Filterplatte mit einem Filterplattendeckel verschlossen und für 5 min bei 2400 xg zentrifugiert. Anschließend wurden von jeder Probe 65 µl des *Proben-Pools von Target-spezifischen DNA-Fragmenten* in ein Well hinzugegeben, die Filterplatte mit dem Deckel verschlossen und für 5 min bei 20°C und 2400 xg zentrifugiert. Zum Entfernen der ungebundenen Oligonukleotide aus der DNA-Lösung wurde jede Probe 2x mit Puffer gewaschen. Dazu wurden jeweils 45 µl SW1-Lösung zu jeder Probe ins Well pipettiert, die Filterplatte abgedeckt und für 5 min bei 2400 xg zentrifugiert. Anschließend wurde aus jedem Well der restliche Überstand mit einer Pipette abgenommen und verworfen. Danach wurden in jedes Well 45 µl UB1-Lösung gegeben. Die Filterplatte wurde mit dem Deckel verschlossen und für 5 min bei 2400 xg zentrifugiert. 45 µl ELM4-Lösung wurden in jedes Well der Filterplatte zu den aufgereinigten *Proben-Pools von Target-spezifischen DNA-Fragmenten* pipettiert und die Filterplatte mit einer Klebefolie verschlossen. Es folgte eine Inkubation der Proben für 45 min bei 37°C im vorgeheizten Hybridisierungssofen (Illumina). Anschließend wurde die Platte für 5 min bei 2400 xg zentrifugiert und dann 25 µl 50 mM NaOH zu jeder Probe gegeben und für 5 min bei RT inkubiert.

2.) *Amplifikation mit Barcode und Sequenzier-Adapter*

Jeweils 4 µl Index-Primer i5 und i7 wurden für jede Probe in ein Well einer neuen 96-Well-Platte pipettiert. 56 µl TDP1 wurden zu 2,8 ml PMM2 gegeben, gemischt und 22 µl TDP1/PMM2-Reaktionsmix zu jedem i5/i7-Primermix in das Well überführt. Weiterhin wurden jeweils 22 µl aufgereinigter *Proben-Pool von Target-spezifischen DNA-Fragmenten* von jeder Probe in ein Well der Platte pipettiert. Die Platte wurde mit einer Folie verschlossen und für 1 min bei 20°C und 1000 xg zentrifugiert und zur Amplifikation in einen Thermocycler (BioRad) gestellt. Anschließend wurde folgendes PCR-Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|--------|------------|--------|
| Denaturierung | 3 min | 95°C | 1x |
| Annealing | 30 sec | 95°C | 27x |
| | 30 sec | 66°C | |
| Elongation | 60 sec | 72°C | |
| | 5 min | 72°C | 1x |
| Abkühlung | ∞ | 10°C | |

Nachfolgend wird für diese amplifizierten Proben der Begriff *Barcode-Proben-Pool* verwendet. Zur Überprüfung der Amplifikation und der Produktgrößen wurde ein Bioanalyser-Lauf mit 1 µl PCR-Produkt von jedem *Barcode-Proben-Pool* durchgeführt (Details siehe unter Abschnitt 2.2.2.3).

Anschließend wurden 45 µl AMPure XP Beads zu jedem *Barcode-Proben-Pool* in die 96-Well-Platte pipettiert, die Platte mit einer Klebefolie verschlossen und für 2 min bei 1800 rpm zentrifugiert. Die Platte wurde für 10 min bei RT inkubiert und danach auf einen Magnetständer gestellt. Nachdem der Überstand klar war, wurde dieser abgenommen und zu jeder Probe 200 µl 80% EtOH gegeben. Nach etwa 20 sec Inkubationszeit wurde der Überstand abgenommen und jede Probe erneut unter den gleichen Bedingungen mit EtOH gewaschen. Die Magnet-Beads wurden für 10 min bei RT getrocknet und abschließend 30 µl EBT in jedes Well pipettiert. Die Platte wurde verschlossen, 2 min bei 1800 rpm zentrifugiert, 2 min bei RT inkubiert und in den Magnetständer gestellt. Der Überstand (aufgereinigter *Barcode-Proben-Pool*) wurde abgenommen und in eine neue 96-Well-Platte überführt.

3.) *Herstellung der Library*

800 µl LNB1 (Magnetische Normalisierungs-Beads) wurden zu dem 15 ml Tube mit LNA1 (Normalisierungs-Puffer) hinzugegeben und gut gemischt. 45 µl LNA1/LNB1 wurden zu jedem aufgereinigten *Barcode-Proben-Pool* in die 96-Well-Platte pipettiert. Die Platte wurde mit einer Klebefolie geschlossen und für 30 min bei 1800 rpm zentrifugiert. Die Platte wurde auf den Magnetständer gestellt und der klare Überstand abgenommen und verworfen. 45 µl LNB1 wurden in jedes Well pipettiert, die Platte verschlossen, 5 min bei 1800 rpm zentrifugiert, dann wurde die Platte erneut auf den Magnetständer gestellt und der klare Überstand abgenommen. Die Platte wurde vom Magnetständer genommen, 30 µl 0,1 N NaOH zu jeder Probe gegeben, 5 min bei 1800 rpm zentrifugiert und die Platte wieder in den Magnetständer gestellt. Der Überstand (normalisierter *Barcode-Proben-Pool*) wurde abgenommen und in eine neue 96-Well-Platte überführt. Anschließend wurden 30 µl LNS2 zu jeder Probe hinzugegeben, die Platte verschlossen und für 1 min bei 1000 xg zentrifugiert. 5 µl jeder Probe der normalisierten *Barcode-Proben-Pools* wurde zur *Library* zusammenpipettiert.

4.) *Sequenzierung mit dem MiSeq*

20 µl der *Library* wurden mit 580 µl HT1 (MiSeq Reagent Kit v3) gemixt und das Tube für 2 min bei 96°C im Heizblock inkubiert. Anschließend wurde das Tube kurz gevortext, die Lösung herunterzentrifugiert und die *Library* in das Proben-Reservoir der Sequenzier-Kartusche geladen. Die Hochdurchsatz-Sequenzierung des *TruSight Amplicon Panels* (Myeloid-Panel) erfolgte mit einem 2x 150 bp (*Paired-End*) MiSeq-Lauf. Die weiteren Arbeitsschritte zum Starten der Hochdurchsatz-Sequenzierung wurden, wie ab dem Kapitel 2.2.4.4 beschrieben, durchgeführt. Details zu den einzelnen Sequenzier-Reagenzien und optimalen Parametern in der Hochdurchsatz-Sequenzierung sind in dem Abschnitt 2.2.5 detaillierter beschrieben.

2.2.4.1.2 Two-Step-PCR-Assay (ID3-Assay)

Das Prinzip der *Two-Step-PCR* basiert auf zwei nacheinander folgenden Amplifikationen ohne zusätzliche Hybridisierungen oder andere komplexe Methodik. Hierbei wird zuerst eine herkömmliche PCR mit Primern inkl. Adaptoren durchgeführt, wodurch die spezifischen genomischen Target-Regionen amplifizieren werden. Anschließend wird eine weitere Amplifikation zur Vervielfältigung sowie zum Anfügen von Barcodes und Sequenzier-Adaptoren vorgenommen (Abbildung 17).

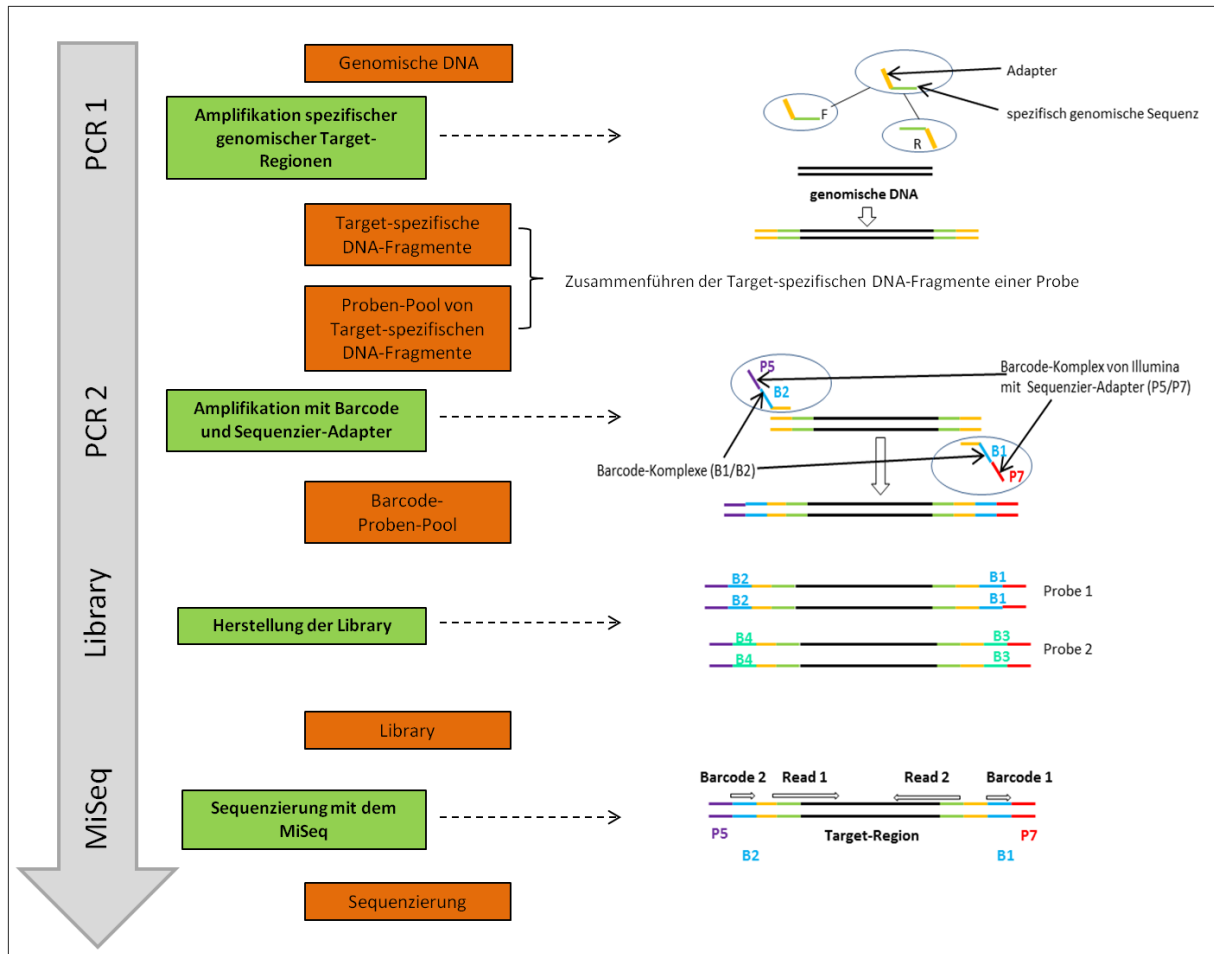


Abbildung 17: Schematische Darstellung der Probenvorbereitung mit dem Verfahren der *Two-Step-PCR*

Bei dem Verfahren der *Two-Step-PCR* erfolgt zuerst eine herkömmliche PCR der genomischen DNA, wobei die verwendeten Target-spezifischen Primer zusätzlich am 5' Ende Illumina-spezifische Oligonukleotide (Adapter) besitzen. Die *Target-spezifischen DNA-Fragmente* eines Patienten werden gepoolt (*Proben-Pool von Target-spezifischen DNA-Fragmenten*) und eine zweite PCR durchgeführt, worin jedem *Proben-Pool von Target-spezifischen DNA-Fragmenten* ein spezifischer Barcode-Sequenzier-Adapter-Komplex zugeordnet wird (*Barcode-Proben-Pool*). Es folgt die Herstellung der *Library* indem von allen zu analysierenden Proben die *Barcode-Proben-Pool* zusammengeführt werden. Abschließend wird die Hochdurchsatz-Sequenzierung der *Library* auf dem MiSeq durchgeführt (modifiziert nach Illumina²¹⁸). In der Grafik sind die verwendeten Begriffe während der Probenvorbereitung zur Beschreibung der Arbeitsschritte in grün und zur Beschreibung der DNA-Lösungen in orange gekennzeichnet. *: Zur Vereinfachung der schematischen Darstellung wurden der Arbeitsschritt beispielhaft für nur eine Probe dargestellt.

Insgesamt ähnelt das Verfahren des *Two-Step-PCR-Assay (ID3-Assay)* der zuvor beschriebenen Probenvorbereitung mittels *TruSight Amplicon Panel (Myeloid-Panel)*. Jedoch ist zu beachten, dass die erste Amplifikation des *Two-Step-PCR-Assay (ID3-Assay)* keine Multi-Plex-PCR ist und die Target-spezifischen DNA-Fragmente einer Probe separat zu einem Proben-Pool von Target-spezifischen DNA-Fragmenten zusammengeführt werden. Die weitere Probenvorbereitung verläuft analog zum *TruSight Amplicon Panel (Myeloid-Panel)*. Die Anreicherung der Proben mittels *Two-Step-PCR* wurde in dieser Arbeit zur Mutationsanalyse des *ID3*-Gens genutzt. Die durchgeführten Arbeitsschritte sind analog zu der Abbildung 17 nachfolgend kurz beschrieben:

1.) *Amplifikation spezifischer genomischer Target-Regionen*

Für den *Two-Step-PCR-Assay (ID3-Assay)* wurden zwei Primer-Paare zur Amplifikation von Exon 1 mit unterschiedlichen Produktgrößen generiert, um möglichst alle genetischen Aberrationen inkl. Deletionen und Insertionen im *ID3*-Gen (Exon 1) zu detektieren. Zusätzlich wurden zwei *ID3*-Kontrollregionen (KR) ausgewählt, die in den *Whole genome sequencing (WGS)*-Daten des ICGC MMML-Seq Projektes in diesen Bereichen keine Mutationen aufweisen. Diese Negativregionen wurden verwendet um potenzielle Sequenzier-Artefakte und-Fehler besser identifizieren zu können. Es ist bekannt, dass Proofreading-Polymerasen eine höhere Prozessivität und bessere Korrekturlesefunktionen besitzen als herkömmliche Polymerasen. Da es sich bei der Hochdurchsatz-Sequenzierung um eine sehr sensitive Methode handelt, sollten während der Library Präparation möglichst wenig Amplifikationsfehler in die Nukleotid-Sequenz eingebaut werden. Um die Hochdurchsatz-Sequenzierungs-Daten hinsichtlich Amplifikations- und Sequenzier-Fehlern zu beurteilen, wurden im Rahmen dieser Arbeit vier verschiedene Polymerase Systeme zur Probenvorbereitung mit dem *ID3*-Assay verwendet. Für die Testversuche wurden unterschiedliche Polymerasen ausgewählt, die zum damaligen Zeitpunkt am Institut für Humangenetik zur Verfügung standen und vom Hersteller als Hochdurchsatz- und Multiplex-PCR-geeignet beschrieben waren: AccuPrime (Life Technologies), Roche Fast Start High Fidelity (Roche), Q5 High-Fidelity (New England Biolabs) und Phusion High-Fidelity (New England Biolabs). In der Austestung der Polymerasen wurden die standardmäßigen Amplifikationsbedingungen nach Herstellerangaben verwendet. Der Reaktionsmix für eine Amplifikation wurde standardmäßig wie folgt erstellt:

| Reagenzien | AccuPrime* | Roche | Q5 | Phusion |
|-------------------------------|------------------------|------------------------|------------------------|------------------------|
| genomische DNA (50 ng) | 1,0 µl | 1,0 µl | 1,0 µl | 1,0 µl |
| H ₂ O | 20,0 µl | 11,8 µl | 15,5 µl | 11,8 µl |
| Puffer | 2,5 µl | 4,0 µl | 5,0 µl | 4,0 µl |
| dNTPs | - | 2,0 µl | 2,0 µl | 2,0 µl |
| Vorwärts-Primer (20 pmol/µl) | 0,5 µl | 0,5 µl | 0,6 µl | 0,5 µl |
| Rückwärts-Primer (20 pmol/µl) | 0,5 µl | 0,5 µl | 0,6 µl | 0,5 µl |
| Polymerase | 0,5 µl | 0,2 µl | 0,3 µl | 0,2 µl |
| Gesamt | 25 µl/ Reaktion | 20 µl/ Reaktion | 25 µl/ Reaktion | 20 µl/ Reaktion |

* dNTPs sind im Reaktionspuffer des AccuPrime Polymerase Systems enthalten.

Die Reaktionsmische wurden zusammenpipettiert, kurz gevortext, herunterzentrifugiert und in einem Thermocycler der Firma Biometra das Programm für die jeweils verwendete Polymerase gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|--------|---------------------------------------|-------------------------------------|
| Denaturierung | 5 min | 98°C | 1x |
| Annealing | 30 sec | 98°C | 25x ¹ / 36x ² |
| | 30 sec | 55°C/ 60°C* | |
| Elongation | 30 sec | 72°C ¹ / 68°C ² | |
| Abkühlung | 5 min | 72°C ¹ / 68°C ² | 1x |
| | ∞ | 10°C | |

*Abhängig von der Annealing-Temperatur der eingesetzten Primer (2.1.3, Tabelle 12), ¹: Elongationstemperatur und Zyklenzahl für die Amplifikation mit den Polymerase Systemen Roche Fast Start, Q5 und Phusion. ²: Elongationstemperatur und Zyklenzahl für die Amplifikation mit dem Polymerase Systemen AccuPrime.

Die jeweilige Proben-DNA-Oligo-Lösung nach der Amplifikation wird im Weiteren als *Target-spezifische DNA-Fragmente* bezeichnet. Anschließend wurde dieser Target-spezifische DNA-Reaktionsmix über eine Filterplatte unter Verwendung von Vakuum (Mini Elute 96 PCR Purification, Qiagen) aufgereinigt. Die Durchführung der Aufreinigung der PCR-Produkte erfolgte analog zu den zuvor beschriebenen Arbeitsschritten des Verfahrens in „Polymerasekettenreaktion und Agarose-Gelelektrophorese“ (Details siehe unter Abschnitt 2.2.3.1). Abschließend wurden von jeder Probe jeweils 5 µl aufgereinigte *Target-spezifische DNA-Fragmente* zum *Proben-Pool von Target-spezifischen DNA-Fragmenten* zusammengeführt (*ID3-Assay*: 4x *Target-spezifische DNA-Fragmente* pro Probe).

2.) Amplifikation mit Barcode und Sequenzier-Adapter

Es folgte eine zweite PCR, wobei als Input der *Proben-Pools* von *Target-spezifischen DNA-Fragmenten* jeder Probe verwendet wurde. Innerhalb der einzelnen Amplifikationen wurden spezifische Barcodes und die Illumina-spezifischen Sequenzier-Adapter an die aufgereinigten Target-spezifischen DNA-Fragmente jedes *Proben-Pools* angefügt. Der Reaktionsmix von 50 µl für eine Amplifikation wurde standardmäßig wie folgt erstellt:

| Reagenzien | Volumen |
|--------------------------------------------------------|------------------------|
| aufgereinigte Target-spezifische DNA-Fragmente (50 ng) | 10,0 µl |
| H ₂ O | 23,0 µl |
| 10x Puffer des AccuPrime Polymerase Systems | 6,0 µl |
| Index 1 Nextera Kit (Barcode 1) | 5,0 µl |
| Index 2 Nextera Kit (Barcode 2) | 5,0 µl |
| AccuPrime Polymerase | 1,0 µl |
| Gesamt | 50 µl/ Reaktion |

Die zusammenpipettierten Reaktionsmische wurden zur Barcode-PCR in den Thermocycler gestellt und das nachfolgende Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|---------------|--------|------------|--------|
| Denaturierung | 2 min | 95°C | 1x |
| Annealing | 10 sec | 95°C | 5x |
| | 30 sec | 63°C* | |
| Elongation | 5 min | 68°C | |
| Abkühlung | ∞ | 10°C | 1x |

Die amplifizierten *Proben-Pools von Target-spezifischen DNA-Fragmenten* wurden aus dem Cycler genommen und erneut aufgereinigt (Mini Elute 96 PCR Purification, siehe Abschnitt 2.2.3.1). Die aufgereinigten *Proben-Pools von Target-spezifischen DNA-Fragmenten* nach der Barcode-PCR werden nachfolgend als *Barcode-Proben-Pool* bezeichnet. Der *Barcode-Proben-Pool* jeder Probe wurde in ein PCR-Tube überführt und die DNA-Konzentration mittels Qubit (siehe Abschnitt 2.2.2.1) und die Produktgröße mittels Bioanalyzer bestimmt (siehe Abschnitt 2.2.2.3).

3.) Herstellung der Library

Die DNA-Konzentration jedes aufgereinigten *Barcode-Proben-Pools* wurde mit dest. H₂O auf 2 nM eingestellt und jeweils 5 µl verdünnter *Barcode-Proben-Pool* (2 nM) zur *Library* zusammenpipettiert.

4.) Sequenzierung mit dem MiSeq

Die *Library* wurde nach dem Standard-Protokoll (siehe Abschnitt 2.2.4.3) auf 10 pM normalisiert und mit 10% PhiX versetzt. Die Arbeitsschritte zur Normalisierung sowie das weitere Vorgehen zum Starten der Hochdurchsatz-Sequenzierung sind ab dem Kapitel 2.2.4.3 genauer beschrieben und dort nachzulesen. Für den *Two-Step-PCR-Assay (ID3-Assay)* wurde ein *MiSeq Reagent Kit v2* mit 500 Zyklen verwendet und eine Hochdurchsatz-Sequenzierung von 2x 250 bp (*Paired-End*) durchgeführt. Details zu den einzelnen Sequenzier-Kit-Komponenten und den optimalen Parametern in der Hochdurchsatz-Sequenzierung sind in dem Abschnitt 2.2.5 detaillierter beschrieben.

2.2.4.2 Hybridisierungs-basierte Anreicherungsverfahren

In der vorliegenden Arbeit wurden zwei verschiedene Hybridisierungs-basierte Anreicherungsverfahren zur Probenvorbereitung verwendet. Hierzu wurden am Institut für Humangenetik zum einen das *TruSight Rapid Capture Panel* (Cancer-Panel, kommerziell erhältliches Panel) und zum anderen das *SureSelect XT Panel* (Lymphom-Panel, Custom Panel) getestet und optimiert. Die einzelnen Hybridisierungs-basierten Panel werden nachfolgend genauer beschrieben.

2.2.4.2.1 *TruSight Rapid Capture Panel* (Cancer-Panel)

Zur Analyse von Indexpatienten mit einem Verdacht auf ein familiäres Mamma- oder Ovarialkarzinom wurde das *TruSight Rapid Capture Panel* (Cancer-Panel) der Firma Illumina am Institut für Humangenetik getestet. Dieses kommerziell erhältliche Panel umfasst 94 Gene (~0,255 Mb), die in malignen Tumoren rekurrent mutierte sind und deckt über 1700 Exons durch die Verwendung von etwa 4000 Target-spezifischen Oligonukleotiden ab²²⁹.

Bei dem Hybridisierungs-basierten Verfahren der Firma Illumina wird zunächst die genomische DNA enzymatisch mittels Transposase auf eine DNA-Fragmentgröße von ca. 150 bis 200 bp fragmentiert und anschließend universelle Adapter an die genomischen DNA-Fragmente ligiert. Es folgte eine Amplifikation bei dem jedem *Proben-Pool* ein individueller Barcode zugeordnet wird. Nachfolgend wurden alle *Barcode-Proben-Pools* zur *Library* zusammenpipettiert und eine Hybridisierung der *Library* mit Biotin-markierten Oligonukleotiden vorgenommen, die spezifisch an die Target-Regionen binden. Mit Hilfe von magnetischen Streptavidin-Beads werden die Target-spezifischen DNA-Fragmente aus der *Library* isoliert und abschließend mit dem MiSeq Gerät von Illumina sequenziert.

Die Probenvorbereitung wurde nach Herstellerangaben mit dem Protokoll „*TruSight Rapid Capture Library Preparation Guide*, Stand 2014“ durchgeführt. Die durchgeführten Arbeitsschritte sind nachfolgend genauer beschrieben. Sofern keine Herstellerangaben angegeben sind, stammten die verwendeten Reagenzien aus dem *TruSight Rapid Capture Kit* (Illumina).

1.) Fragmentierung genomischer DNA

Die genomische DNA wurde für jede zu analysierende Probe auf 5 ng/µl in einem Endvolumen von 10 µl mit Tris-Cl (10 mM) verdünnt und in ein Well einer 96-Well-Platte überführt. Anschließend wurden zu jeder Probe 25 µl *Tagment Buffer* und 15 µl *Tagment DNA Enzym 1* pipettiert, die 96-Well-Platte wurde mit einer Klebefolie verschlossen, 1 min bei 1800 rpm auf den Plattschüttler gestellt und für 1 min bei 300 xg zentrifugiert. Die Proben wurden in den Cycler gestellt (BioRad) und für 10 min bei 58°C inkubiert. Abschließend wurde der Fragmentierungsprozess durch die Zugabe von 15 µl *Stop Tagment Buffer* gestoppt. Die 96-Well-Platte wurde erneut für 1 min bei 1800 rpm auf den Plattschüttler gestellt, für 1 min bei 300 xg zentrifugiert und abschließend 4 min bei RT inkubiert.

Zur Aufreinigung der DNA-Fragmente wurden 65 µl *Sample Purification Beads* in jedes Proben-Well pipettiert, die Platte mit einer Klebefolie verschlossen und für 1 min bei 1800 rpm zentrifugiert. Die Platte wurde für 8 min bei RT inkubiert, 1 min bei 300 xg zentrifugiert und danach auf einen Magnetständer gestellt. Nachdem der Überstand klar war, wurde dieser abgenommen und in jedes Well 200 µl 80% EtOH gegeben. Nach 20 sec Inkubationszeit wurde der Überstand abgenommen und jede Probe erneut unter den gleichen Bedingungen mit EtOH gewaschen. Die Magnet-Beads wurden für 10 min bei RT getrocknet und abschließend 22,5 µl *Resuspensions Buffer* in jedes Well pipettiert. Die 96-Well-Platte wurde für 2 min bei RT inkubiert, der Überstand (DNA-Fragmente) abgenommen und in eine neue Platte überführt.

2.) Amplifikation mit Barcode und Sequenzier-Adapter

Für jede Probe wurden jeweils 5 µl i5- und i7-Primer (Barcodes) in unterschiedlichen Kombinationen in ein Well einer neuen 96-Well-Platte pipettiert. Jeweils 20 µl *Nextera Library Amplification Mix* und 20 µl aufgereinigte DNA-Fragmente wurden zu jedem i5/i7-Primermix in das Well pipettiert. Die Platte wurde mit einer Folie verschlossen und für 1 min bei 1200 rpm zentrifugiert und zur Amplifikation in einen Thermocycler (BioRad) gestellt. Anschließend wurde folgendes PCR-Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|--------|------------|--------|
| Denaturierung | 3 min | 72°C | 1x |
| Annealing | 30 sec | 98°C | 5x |
| | 30 sec | 60°C* | |
| Elongation | 30 sec | 72°C | |
| Abkühlung | 5 min | 72°C | 1x |
| | ∞ | 10°C | |

Danach wurde die Platte aus dem Cycler genommen und die DNA-Lösung aufgereinigt. Dazu wurden die Arbeitsschritte analog zu der zuvor beschriebenen Aufreinigung im Abschnitt Fragmentierung genomischer DNA durchgeführt. Die aufgereinigten Barcode-DNA-Fragmente wurden in 27 µl *Resuspensions Buffer* eluiert und abschließend die DNA-Konzentration mittels Qubit bestimmt (siehe Abschnitt 2.2.2.1). Die aufgereinigten Barcode-DNA-Fragmente werden nachfolgend als *Barcode-DNA-Pool* bezeichnet.

3.) Herstellung der Library

Von jeder Probe wurden jeweils 500 ng *Barcode-DNA-Pool* in einem neuen PCR-Tube zusammengeführt und das Volumen auf 40 µl mit *Resuspensions Buffer* eingestellt. Der zusammengeführte DNA-Pool wird nachfolgend als *Library* bezeichnet.

4.) *Anreicherung spezifischer genomischer Target-Regionen*

Es wurden 50 µl *Enrichment Hybridization Buffer* und 10 µl *Custom Selected Oligos (TruSight Cancer Sequencing Kit, Illumina)* zur *Library* in das PCR-Tube hinzugegeben und das Tube für 1 min bei 1200 rpm auf den Plattenschüttler gestellt. Anschließend wurden die Proben in den Thermocycler gestellt (BioRad) und folgendes Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|--------------------------|---------------|------------|------------------------------------|
| Denaturierung | 10 min | 95°C | 1x |
| Annealing und Elongation | 60 sec | 94- 58°C | 18x (je Zyklus 2 Grad abkühlen) |
| Hybridisierung | 90 min – 24 h | 58°C | |

Die Probe wurde aus dem Cyclyer genommen und mittels magnetischer Streptavidin-Beads wurden die Target-spezifischen DNA-Fragmente von den unspezifischen DNA-Fragmenten der *Library* selektiert. Dazu wurde die *Library* nach der Hybridisierung in eine 96-Well-Platte überführt und 250 µl *Streptavidin Magnetic Beads* hinzugegeben. Die Platte wurde mit einer Klebefolie verschlossen, 5 min bei 1200 rpm geschüttelt, 25 min bei RT inkubiert und 1 min bei 300 xg zentrifugiert. Danach wurde die Platte auf einen Magnetständer gestellt und der klare Überstand verworfen. 200 µl *Enrichment Wash Solution* wurden zu jeder Probe gegeben, gemixt und 30 min bei 50°C im Cyclyer (BioRad) inkubiert. Die Platte wurde erneut auf den Magnetständer gesetzt und der Überstand verworfen. Anschließend wurde die *Library* ein zweites Mal mit 200 µl *Enrichment Wash Solution* gewaschen. Nachfolgend wurde 28,5 µl *Enrichment Elution Buffer* mit 1,5 µl NaOH gemixt und 23 µl dieser Lösung zur *Library* hinzugegeben. Die NaOH-*Library*-Bead-Lösung wurde 2 min bei 1800 rpm zentrifugiert, 2 min bei RT inkubiert und 1 min bei 300 xg zentrifugiert. Danach wurde die Platte erneut auf den Magnetständer gestellt und 21 µl klarer Überstand in ein neues PCR-Tube überführt. Abschließend wurden 4 µl *Elute Target Buffer 2* zur *Library von Target-Spezifischen DNA-Fragmenten* hinzupipettiert und das Tube für 1 min bei 1200 rpm zentrifugiert. Nachfolgend wurde eine zweite Hybridisierung mit den gleichen Target-spezifischen Oligonukleotiden vorgenommen, um die Spezifität der Anreicherung des Panels zu erhöhen. Hierzu wurden 15 µl *Resuspensions Buffer*, 50 µl *Enrichment Hybridization Buffer* und 10 µl *Custom Selected Oligos (TruSight Cancer Sequencing Kit)* zu den 25 µl aufgereinigte *Library* aus der ersten Hybridisierung hinzugegeben. Die Lösung wurde 1 min bei 1200 rpm zentrifugiert und erneut das zuvor beschriebene Thermocycler-Programm zur Hybridisierung verwendet. Nach der Hybridisierung wurde erneut eine Selektion der Target-spezifischen DNA-Fragmente von den unspezifischen DNA-Fragmenten der *Library* mittels Streptavidin-Beads vorgenommen. Die Durchführung der Arbeitsschritte der Selektion erfolgte analog zu den Arbeitsschritten nach der ersten Hybridisierung. Abschließend wurden 25 µl *Library von Target-spezifischen DNA-Fragmenten* in eine Platte überführt und zur Aufreinigung 45 µl *Sample Purification Beads* hinzupipettiert, die Platte für 1 min bei 1800 rpm auf den Plattenschüttler gestellt und danach 10 min bei RT inkubiert.

Anschließend wurde die Platte 1 min bei 300 xg zentrifugiert, auf den Magnetständer gestellt und der klarer Überstand verworfen. 200 µl 80% EtOH wurden zur Bead-Lösung hinzupipettiert, die Proben für 30 sec auf dem Magnetständer inkubiert und der klare Überstand abgenommen. Der Waschschrift wurde erneut mit 200 µl 80% EtOH wiederholt. Die Beads wurden anschließend für etwa 15 min bei RT getrocknet, die Library in 27,5 µl *Resuspensions Buffer* eluiert und 25 µl Überstand (*Library von Target-spezifischen DNA-Fragmenten*) in ein PCR-Tube überführt.

5.) Amplifikation

5 µl *Primer Cocktail* und 20 µl *Nextera Enrichment Amplification Mix* wurden zu den 25 µl aufgereinigte *Library von Target-spezifischen DNA-Fragmenten* in das PCR-Tube pipettiert und 1 min bei 1200 rpm zentrifugiert. Die Probe wurde in den Thermocycler gestellt (BioRad) und folgendes Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|------------|--------------|--------|
| Denaturierung | 30 sec | 98°C | 1x |
| Annealing | 10 sec | 98°C | 12x |
| | 30 sec | 60°C | |
| Elongation | 30 sec | 72°C | |
| Abkühlung | 5 min ∞ | 72°C 10°C | 1x |

Nach der Amplifikation wurde die *Library von Target-spezifischen DNA-Fragmenten* erneut in eine 96-Well-Platte überführt und aufgereinigt. Hierzu wurden 90 µl *Sample Purification Beads* hinzugegeben, 1 min bei 1800 rpm zentrifugiert, 10 min bei RT inkubiert und für 1 min bei 300 xg zentrifugiert. Die Platte wurde auf den Magnetständer gestellt und der klare Überstand verworfen. Anschließend wurden die Beads 2x mit 200 µl 80% EtOH gewaschen. Die Durchführung der Waschschriffe erfolgt analog dem bereits zuvor beschriebenen Vorgehen bei der Aufreinigung der Library mit den *Sample Purification Beads*. Die DNA wurde abschließend in 32 µl *Resuspensions Buffer* eluiert und 30 µl Überstand (*Library von Target-spezifischen DNA-Fragmenten*) in ein neues PCR-Tube überführt. Die *Library von Target-spezifischen DNA-Fragmenten* wird nachfolgend nur noch als *Library* bezeichnet. Die DNA-Konzentration der Library wurde mittels Qubit (siehe Abschnitt 2.2.2.1) und die Produktgröße mittels Bioanalyzer bestimmt (siehe Abschnitt 2.2.2.3). Anschließend wurde die Library wurde mit *Resuspensions Buffer* auf 2 nM verdünnt.

6.) Sequenzierung mit dem MiSeq

Die *Library* wurde nach dem Standard-Protokoll der Firma Illumina auf 10 pM normalisiert und mit 1% PhiX versetzt. Die Arbeitsschritte zur Normalisierung sowie das weitere Vorgehen zum Starten der Hochdurchsatz-Sequenzierung sind ab dem Kapitel 2.2.4.3 genauer beschrieben und dort nachzulesen. Für das *TruSight Rapid Capture Panel* (Cancer-Panel) wurde ein *MiSeq Reagent Kit v2* mit 300 Zyklen verwendet und eine Hochdurchsatz-Sequenzierung von 2x 150 bp (*Paired-End*) durchgeführt. Details zu den einzelnen Illumina-spezifischen Sequenzier-Kit-Komponenten und den optimalen Parametern in der Hochdurchsatz-Sequenzierung sind in dem Abschnitt 2.2.5. genauer beschrieben.

In Abhängigkeit von der Konzentration der *Library* wurde die Normalisierung in zwei Verfahren durchgeführt:

a) Für *Libraries* mit einer Molarität > 2 nM:

Dieses Verfahren entspricht der Normalisierung wie das Vorgehen im vorherigen Abschnitt „Sequenzierung mit dem MiSeq“ beschrieben wurde. Die *Library* wurde mit dem *Resuspensions Buffer* auf 2 nM verdünnt und nach dem Standard-Protokoll (Details siehe Abschnitt 2.2.4.3.1) für die Hochdurchsatz-Sequenzierung auf 10 pM normalisiert.

b) Für *Libraries* mit einer Molarität < 2nM

In Testversuchen entsprach die gemessene DNA-Molarität nach der *Library*-Präparation von 4 Proben mit dem Standard-Protokoll nicht den erfordernten 2 nM, die für eine Normalisierung notwendig waren. Da häufig weniger als 12 zu analysierende Proben vorliegen, war es notwendig ein Protokoll zu optimieren. Dazu wurde ein *Small-Scale-Protokoll* optimiert, dass es ermöglicht auch niedriger konzentrierte *Libraries* von unter 2 nM mittels Hochdurchsatz-Sequenzierung zu analysieren. Eine detaillierte Beschreibung und Durchführung zur Normalisierung von *Libraries* mit dem *Small-Scale-Protokoll* ist unter dem Abschnitt 2.2.4.3.2 zu finden. Die Sequenzierung der mittels *Small-Scale-Protokoll* normalisierten Cancer-Panel-*Libraries* erfolgte mit einem *MiSeq Reagent Nano Kit* der Firma Illumina (Details zu den Sequenzier-Kit-Komponenten siehe Abschnitt 2.2.5). Die weiteren Parameter in der Sequenzierung der mittels *Small-Scale-Protokoll* normalisiert Cancer-Panel-*Libraries* sind analog zu dem Standard-Protokoll (1% PhiX, 2x 150 bp).

2.2.4.2.2 *SureSelectXT Panel (Lymphom-Panel)*

Zur Hochdurchsatz-Sequenzierung der Keimzentrums-B-Zell-Lymphome von sehr alten Patienten wurde ein *SureSelect XT Panel* (Lymphom-Panel, Custom Panel) in Zusammenarbeit mit der Firma Agilent neu erstellt und 71 Patienten-Proben (*Jack*: N= 49 siehe Abschnitt 2.1.2.2 und *MMML-Senil*: N= 22 siehe Abschnitt 2.1.2.1) und 9 Kontroll-Proben (*Kontrollen-Sehr alt*: N= 7 und *Kontrollen-Adult*: N= 2 siehe Abschnitt 2.1.2.5) analysiert. Das Panel enthält ausgewählte Target-Regionen, die in Lymphomen rekurrent mutiert aufgetreten sind. Die Auswahl der spezifischen Genorte und Target-Regionen erfolgte vor dem Beginn dieser Arbeit durch Frau Dr. Julia Richter am Institut für Humangenetik und basierte auf den Ergebnissen der Verbundprojekte MMML und ICGC MMML-Seq sowie bereits publizierter Literatur^{31,62,127,135,140,230-232}. Eine Liste mit entsprechenden Target-Regionen (inkl. Start- und Endposition der gDNA) wurde der Firma Agilent Technologies übermittelt. Anschließend erfolgte das Design des Lymphom-Panels bei der Firma Agilent durch ein *in-siliko* Software-Programm, durch das die spezifischen Oligonukleotide zur Abdeckung unserer Target-Regionen ermittelt wurden. Es folgte nun in mehreren Schritten eine Überprüfung des Panel-Designs um sicherzustellen, dass *in-siliko* alle gewünschten Target-Regionen von dem konzipierten Lymphom-Panel abgedeckt wurden. Die entsprechend nicht abgedeckten Bereiche wurden an Agilent zurückgespiegelt und das Oligonukleotid-Design angepasst. Abschließend wurde das finale Design des Lymphom-Panels von Frau Dr. Julia Richter und mir bestätigt und eine Bestellung ausgelöst, wonach die Lymphom-spezifischen Oligonukleotide von der Firma Agilent ausschließlich für uns synthetisiert wurden. Dieses Panel wird im Weiteren als Lymphom-Panel bezeichnet deckt insgesamt 118 Lymphom-spezifische Gene mit einer Größe von etwa 1 Megabase (Mb) ab. Hierbei sind sowohl die kompletten kodierenden Regionen von Genen inkl. untranslatierter Region (UTR) (z.B. *TP53*, *NOTCH1*, *TNFRSF14*), potentiell somatisch hypermutierte Genorte (z.B. *BCL2*, *BCL6*, *SGK1*) als auch einzelne Mutations-Hotspots (z.B. *MYD88*, *BRAF*, *SGK1*) abgedeckt. Eine Liste aller Gene und Regionen, die durch das Lymphom-Panel abgedeckt sind, ist im Anhang in der Tabelle A1 zu finden.

Beim *SureSelectXT Panel* wurde die genomische DNA mittels Ultraschallen auf eine Größe von ca. 150 bis 200 bp fragmentiert. Die Library Präparation erfolgte mittels *SureSelect XT Library Prep. Kit*, wobei universelle Adapter an die genomischen DNA-Fragmente ligiert werden (*DNA-Adapter-Fragmente*). Anschließend wird eine Hybridisierung mit Biotin-markierten Oligonukleotiden vorgenommen, die spezifisch an die Target-Regionen der *Adapter-DNA-Fragmente* binden. Mit Hilfe von magnetischen Streptavidin-Beads werden die Target-spezifischen *Adapter-DNA-Fragmente* von den unspezifischen DNA-Fragmenten isoliert, wonach der *Proben-Pool von Target-spezifischen DNA-Fragmenten* vorliegen. Es folgt eine Amplifikation bei der jedem Probe-Pool ein individueller Barcode zugeordnet wurde (*Barcode-Proben-Pool*).

Abschließend wurde die Konzentration der einzelnen *Barcode-Proben-Pools* mittels Qubit bestimmt, eine Normalisierung jeder Probe vorgenommen und die *Library* erstellt. Die Sequenzierung der *Library* wurde mit dem MiSeq-Gerät der Firma Illumina durchgeführt. Eine schematische Darstellung der durchgeführten Arbeitsschritte zur Probenvorbereitung mittels *SureSelectXT Panel* ist in der nachfolgenden Abbildung 18 zu sehen.

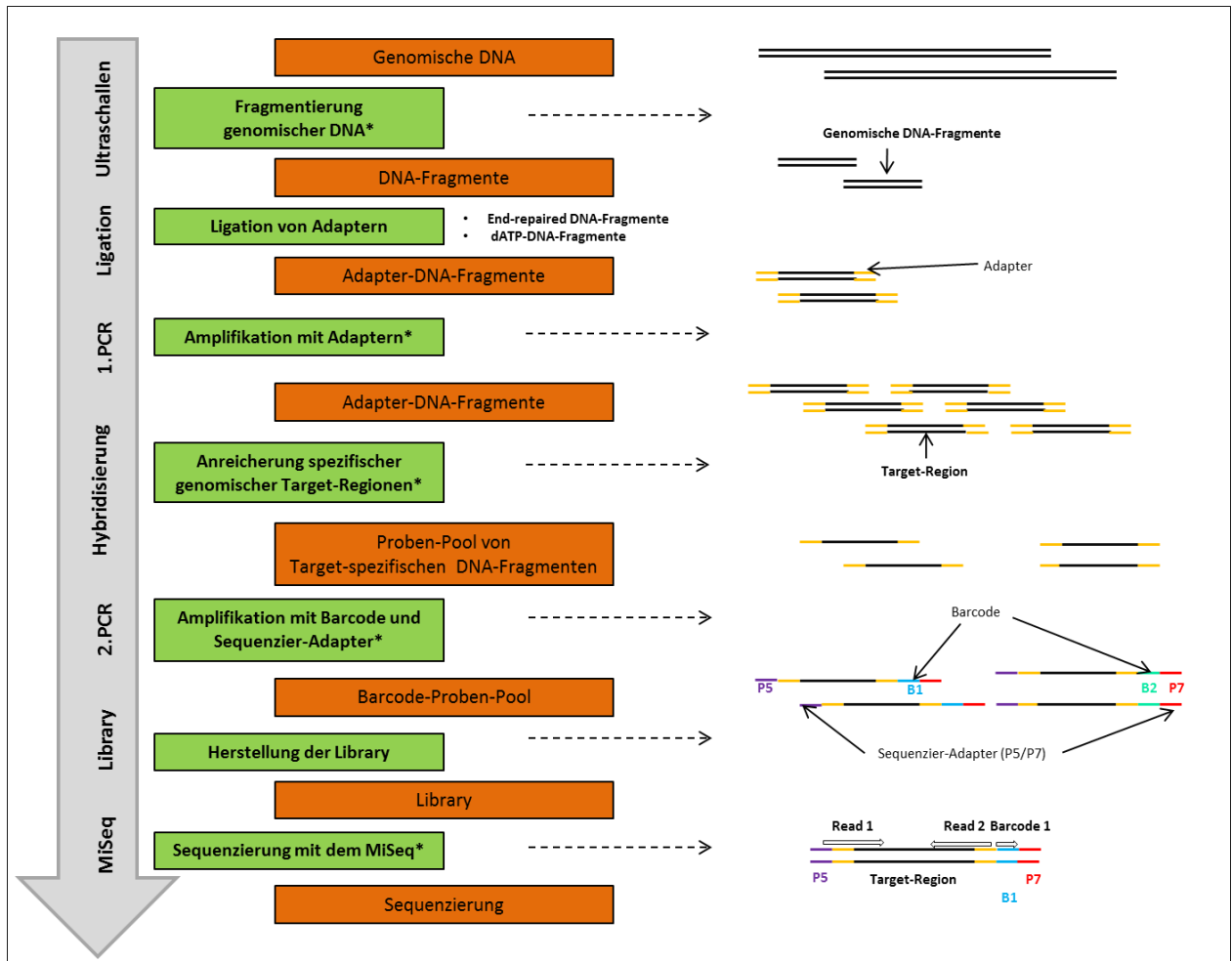


Abbildung 18: Schematische Darstellung der Probenvorbereitung mit dem SureSelectXT Verfahren von Agilent

Die Anreicherung der gDNA mit dem SureSelectXT Kit der Firma Agilent erfolgt mittels Hybridisierungsverfahren. Die genomische DNA wird durch Ultraschall fragmentiert. Anschließend erfolgt ein End-repair-Prozess der DNA-Fragmente und an die 3'-DNA-Fragmentenden werden Desoxyadenosinmonophosphate (dATPs) ligiert. Nachfolgend werden Adapter an die *DNA-Fragmente* ligiert und eine Amplifikation der *DNA-Adapter-Fragmente* vorgenommen. Durch die Hybridisierung von spezifischen Oligonukleotiden an die *Adapter-DNA-Fragmente* erfolgt eine Anreicherung spezifischer genomischer Target-Regionen (*Proben-Pool von Target-spezifischen DNA-Fragmenten*). Die spezifischen DNA-Fragmente werden amplifiziert und sequenziert (modifiziert nach Agilent Technologies). Nachfolgend wird jedem *Proben-Pool von Target-spezifischen DNA-Fragmenten* ein spezifischer Barcode zugeordnet (*Barcode-Proben-Pool*) und die verschiedenen *Barcode-Proben-Pools* zu einer *Library* zusammengeführt. Abschließend wird die *Library* mit dem MiSeq-Gerät Hochdurchsatzsequenziert.

Insgesamt ähnelt der Ablauf des Hybridisierungsverfahrens der Firma Agilent dem zuvor beschriebenen Verfahren des *TruSight Rapid Capture Panel* (Cancer-Panel) der Firma Illumina, wobei die einzelne Arbeitsschritte bzw. die Reihenfolge der Arbeitsschritte variieren. Insbesondere ist zu beachten, dass beim *SureSelectXT Panel* die Anreicherung der Target-spezifischen DNA-Fragmente für jede Probe separat durchgeführt und erst zum Ende der Probenvorbereitung die Library erstellt wird. Hingegen wurde bei der zuvor beschriebenen Probenvorbereitung mittels *TruSight Rapid Capture Panel* zuerst die Library erstellt und dann die Anreicherung der Target-spezifischen DNA-Fragmente vorgenommen.

Die Probenvorbereitung wurde mit dem Protokoll „*SureSelect XT Target Enrichment System for Illumina Paired-End Sequencing Library- Illumina HiSeq and MiSeq Multiplexed Sequencing Platforms (2014)*“ durchgeführt. Einzelne Optimierungen des Protokolls wurden bei der Etablierung des Lymphom-Panels in Anlehnung an die Angaben von Agilent vorgenommen. Die durchgeführten Arbeitsschritte sind analog zu der Abbildung 18 nachfolgend kurz beschrieben. Sofern keine Herstellerangaben angegeben wurden, stammten die verwendeten Reagenzien aus dem *SureSelectXT Library Prep. Kit* (Agilent).

1.) *Fragmentierung genomischer DNA*

Jede zu analysierende genomische DNA-Probe wurde mit 1x Low TE Puffer auf 200 ng in 50 µl eingestellt und die Lösung in ein spezielles Mikro-Tube der Firma Covaris pipettiert. Anschließend wurde das Tube in dem zugehörigen Adapter befestigt und dieser in den Sonifizierer S220 (Covaris) gesetzt. Zum Starten des Ultraschall-Prozesses wurde der zugehörige Laptop mit der Software Sonolab 7 hochgefahren und das Ultraschall-Programm mit den nachfolgenden Parametern eingestellt und gestartet:

| Parameter | Werte |
|---------------------------|-----------|
| Duty Factor | 10% |
| Peak Incident Power (PIP) | 175 |
| Cycles per Burst | 200 |
| Treatment Time | 360 sec |
| Bath Temperature | 4°C – 8°C |

Nach dem Sonifizieren wurde die DNA in ein Eppendorf-Tube überführt und die Fragmentgröße der DNA von etwa 120 bis 150 bp mittels Bioanalyzer überprüft (siehe Abschnitt 2.2.2.3).

2.) Ligation von Adaptern

Eine End-Repair-Reaktion der *DNA-Fragmente* wurde vorgenommen, um die DNA-Fragmentenden für die anschließende Ligation von Desoxyadenosinmonophosphaten (dATP) zu präparieren. Für den End-Repair-Mastermix wurden folgende Reagenzien aus dem *SureSelect Library Kit* zusammenpipettiert:

| Reagenzien | Volumen |
|--------------------------|------------------------|
| Dest. H ₂ O | 35,2 µl |
| 10x End Repair Buffer | 10,0 µl |
| dNTP Mix | 1,6 µl |
| T4 DNA Polymerase | 1,0 µl |
| Klenow DNA Polymerase | 2,0 µl |
| T4 Polynucleotide Kinase | 2,2 µl |
| Gesamt | 52 µl/ Reaktion |

Für jede Probe wurden 48 µl fragmentierte DNA und 52 µl des End-Repair-Mastermixes in ein PCR-Tube gegeben, das Tube kurz gevortext, zentrifugiert und für 30 min bei 20°C im Thermocycler (BioRad) inkubiert. Nach Ablauf der Inkubationszeit wurden die Tubes aus dem Cycler genommen und alle weiteren Arbeitsschritte bei RT durchgeführt. Nach dem erfolgreichen End-Repair-Prozess erfolgte eine Aufreinigung der *DNA-Fragmente*. Dazu wurden jeweils 180 µl magnetische *AmPure XP Beads* (Agencourt) in ein Well einer 96-Well Mikrotiterplatte pipettiert und anschließend 100 µl *End-repaired DNA-Fragmente* hinzugegeben. Die Platte wurden mit einem Deckel verschlossen, kurz gevortext und zentrifugiert. Nach einer Inkubationszeit von 5 min bei RT wurde die Platte auf einen Magnetständer gestellt und der klare Überstand mit der Pipette abgenommen. Die Proben wurden anschließend 2x mit 70% EtOH gewaschen. Hierzu wurden jeweils 500 µl 70% EtOH in das Well pipettiert, 1 min inkubiert und der klare Überstand verworfen. Das restliche EtOH wurde möglichst vollständig aus den Wells entfernt, die Platte vom Magnetständer genommen und die Beads für 15 min bei RT getrocknet. Anschließend wurden die *End-repaired DNA-Fragmente* in 32 µl dest. H₂O resuspendiert. Abschließend wurde die 96-Well-Platte erneut auf den Magnetständer gestellt, etwa 30 µl klarer Überstand (*End-repaired DNA-Fragmente*) mit einer Pipette abgenommen und in ein neues PCR-Tube überführt.

Daraufhin folgte die Ligation von Desoxyadenosinmonophosphate (dATP's) an das 3'Ende der aufgereinigten *End-repaired DNA-Fragmente*. Dazu wurde ein dATP-Mastermix aus den Reagenzien wie folgt zusammenpipettiert:

| Reagenzien | Volumen |
|------------------------------|------------------------|
| Dest. H ₂ O | 11,0 µl |
| 10x Klenow Polymerase Buffer | 5,0 µl |
| dATP | 1,0 µl |
| Exo(-) Klenow | 3,0 µl |
| Gesamt | 20 µl/ Reaktion |

20 µl dATP-Mastermix wurden zu jeder Probe von *End-repaired DNA-Fragmenten* in das PCR-Tube hinzugegeben. Die Tubes wurden kurz gevortext, zentrifugiert und für 30 min bei 37°C in den Thermocycler (BioRad) gestellt. Anschließend wurden die *dATP-DNA-Fragmente* erneut mit den AMPure XP Beads (Magnet-Beads) aufgereinigt. Es wurden 90 µl Magnet-Beads zu jeder Probe mit *dATP-DNA-Fragmente* gegeben, 2x mit 70% EtOH gewaschen, die *dATP-DNA-Fragmente* in 15 µl H₂O eluiert und in ein neues PCR-Tube überführt. Die Durchführung der Aufreinigung mit den Magnet-Beads erfolgte analog zu dem zuvor beschriebenen Arbeitsschritten des Verfahrens nach dem End-Repair-Prozess. An die aufgereinigten *dATP-DNA-Fragmente* wurden Adapter ligiert, die als Bindungsstelle für die Primer in der nachfolgenden Amplifikation benötigt werden. Für die Ligation von einer Reaktion wurde ein Adapter-Mastermix wie folgt hergestellt:

| Reagenzien | Volumen |
|------------------------------|------------------------|
| Dest. H ₂ O | 15,5 µl |
| 5x T4 DNA Ligase Buffer | 10,0 µl |
| SureSelect Adapter Oligo Mix | 10,0 µl |
| T4 DNA Ligase | 1,5 µl |
| Gesamt | 37 µl/ Reaktion |

37 µl Adapter-Mastermix wurden zu den aufgereinigten *dATP-DNA-Fragmenten* in das PCR-Tube pipettiert. Die Tubes wurden kurz gevortext, zentrifugiert und für 15 min bei 20°C in den Thermocycler (BioRad) gestellt. Zur Aufreinigung der *Adapter-DNA-Fragmente* wurden 90 µl Magnet-Beads zu jeder Probe gegeben, 2x mit 70% EtOH gewaschen und die *Adapter-DNA-Fragmente* in 32,5 µl dest. H₂O eluiert. Die Durchführung der Aufreinigung mit den Magnet-Beads erfolgte analog zu den zuvor beschriebenen Arbeitsschritten des Verfahrens nach der Ligation der dATPs.

3.) Amplifikation mit Adaptern

Zur Amplifikation der *Adapter-DNA-Fragmente* wurde ein Mastermix aus den Reagenzien des *SureSelect Library Prep. Kits*, *SureSelect Target Enrichment Kits* und der *Herculase II Fusion DNA Polymerase* (Agilent) hergestellt und wie folgt zusammenpipettiert:

| Reagenzien | Volumen |
|---------------------------------------------------------------------|------------------------|
| Dest. H ₂ O | 6,0 µl |
| SureSelect Primer* | 1,25 µl |
| SureSelect ILM Indexing Pre Capture PCR Reverse Primer [#] | 1,25 µl |
| 5x Herculase II Rxn Buffer* | 10,0 µl |
| 100 mM dNTP Mix* | 0,5 µl |
| Herculase II Fusion DNA Polymerase* | 1,0 µl |
| Gesamt | 20 µl/ Reaktion |

*Reagens aus dem *SureSelect Library Prep. Kit*, [#] Reagens aus dem *SureSelect Target Enrichment Kit ILM Indexing Hyb Module Box 2*, *Reagens aus dem *Herculase II Fusion DNA Polymerase Kit*

20 µl Mastermix und 30 µl Adapter- DNA-Fragmente wurden in einem PCR-Tube gemixt, in den Thermocycler (Biometra) gestellt und das nachfolgende Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|----------------------|------------|-------------|--------|
| Denaturierung | 2 min | 98°C | 1x |
| Annealing | 30 sec | 98°C | 8x |
| | 30 sec | 65°C | |
| Elongation | 60 sec | 72°C | |
| Abkühlung | 5 min ∞ | 72°C 4°C | 1x |

Die *Adapter-DNA-Fragmente* wurden mit 90 µl Magnet-Beads aufgereinigt, in 30 µl dest. H₂O eluiert und in ein Eppendorf-Tube überführt. 30 µl aufgereinigte Adapter-DNA-Fragmente wurden mit einer SpeedVac unter Vakuum vollständig eingedampft, 3,5 µl H₂O in das Tube pipettiert und für mind. 1 h bei 700 rpm in den Thermoschüttler gestellt. Die DNA wurde zum Quellen über Nacht bei RT stehen gelassen.

4.) Anreicherung spezifischer genomischer Target-Regionen

Für die Hybridisierung der Target-spezifischen DNA-Fragmente wurden am Folgetag verschiedene Mastermixe zusammenpipettiert:

| Reagenzien für <i>Hybridization Buffer</i> | Volumen |
|--------------------------------------------|------------------------|
| SureSelect Hyb1 | 6,6 µl |
| SureSelect Hyb2 | 0,3 µl |
| SureSelect Hyb3 | 2,7 µl |
| SureSelect Hyb4 | 3,4 µl |
| Gesamt | 13 µl/ Reaktion |

Der *Hybridization Buffer* wurde im Wasserbad für 5 min bei 65°C erwärmt. Weiterhin wurde ein *SureSelect Block Mix* aus folgenden Reagenzien zusammenpipettiert:

| Reagenzien für <i>SureSelect Block Mix</i> | Volumen |
|--------------------------------------------|-------------------------|
| SureSelect Indexing Block1 | 2,5 µl |
| SureSelect Block 2 | 2,5 µl |
| SureSelect Indexing Block 3 | 0,6 µl |
| Gesamt | 5,6 µl/ Reaktion |

Anschließend wurde die *SureSelect Capture Library* präpariert wobei alle Arbeitsschritte auf Eis durchgeführt wurden. Zuerst wurde die RNase Block Dilution mit H₂O 1:10 (10%) verdünnt und anschließend für jede Probe ein *SureSelect Capture Library Mix* wie folgt präpariert:

| Reagenzien für <i>SureSelect Capture Library Mix</i> | Volumen |
|------------------------------------------------------|-----------------------|
| Verdünnte RNase Block Dilution (1:10) | 2 µl |
| SureSelect Capture Library | 5 µl |
| Gesamt | 7 µl/ Reaktion |

Für die Hybridisierung erfolgte eine Kombination verschiedener Arbeitsschritte worin die oben beschriebenen Mastermixe und Lösungen verwendet wurden. Zur Vereinfachung wurden die einzelnen durchgeführten Arbeitsschritte in der Abbildung 19 schematisch dargestellt und werden nachfolgend genauer beschrieben.

Jeweils 3,5 µl aufkonzentrierte *Adapter-DNA-Fragmente* und 5,6 µl *SureSelect Block Mix* wurden in die Reihe B einer 96-Well-Platte pipettiert (Abbildung 19, rot) und gut durchmischt. Die Platte wurde in einen Thermocycler (BioRad) gestellt und alle nachfolgenden Pipettierschritte am Thermocycler durchgeführt ohne die Platte herauszunehmen. Zuerst wurde der *Adapter-DNA-Fragmente* + *SureSelect Block Mix* für 5 min für 95°C inkubiert. Anschließend wurde die Temperatur des Thermocyclers auf 65°C heruntergestellt. Nachdem der Thermocycler auf 65°C temperiert war, wurden jeweils 40 µl *Hybridization Buffer* in die Wells der Reihe A der 96-Well-Platte pipettiert (Abbildung 19, blau). Die Platte wurde mit einer Klebefolie verschlossen und für 5 min bei 65°C inkubiert. Danach wurden 7 µl *SureSelect Capture Library Mix* in die Reihe C der Platte überführt (Abbildung 19, grün) und für weitere 2 min bei 65°C inkubiert. 13 µl *Hybridization Buffer* wurden aus der Reihe A zur *SureSelect Capture Library Mix* in Reihe C überführt und der Reaktionsmix im Well gut durchmischt (Abbildung 19, Pfeil 1). Abschließend wurde der *Adapter-DNA-Fragmente* + *SureSelect Block Mix* aus Reihe B vollständig zu dem *Hybridization Buffer* + *SureSelect Capture Library Mix* in Reihe C pipettiert und alles gut durchmischt (Abbildung 19, Pfeil 2). Die 96-Well-Platte wurde verschlossen und für mind. 16 h bei 65°C inkubiert. Währenddessen erfolgte die Bindung der Target-spezifischen Oligonukleotide aus der *Sure Select Capture Library* an die *Adapter-DNA-Fragmente*.

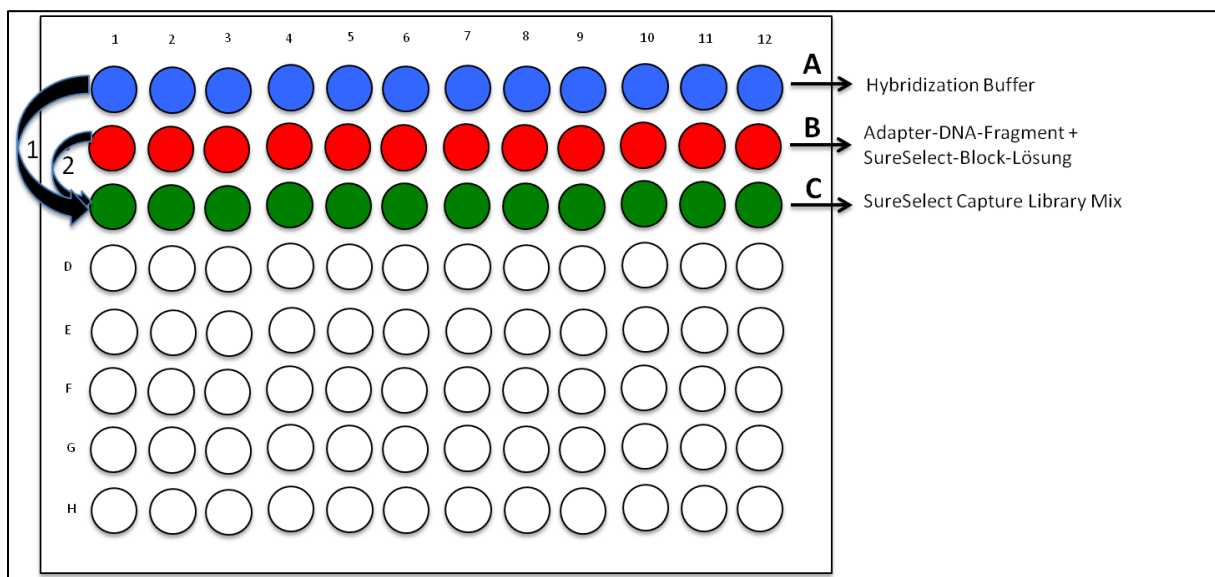


Abbildung 19: Schematische Darstellung des Pipettierschemas zur Hybridisierung von 12 Proben

Reihe A: *Hybridization Buffer* (blau). Reihe B: *Adapter-DNA-Fragmente* + *SureSelect Block Mix* (rot). Reihe C: *Sure Select Capture Library Mix* (grün). 1: *Hybridization Buffer* wird zum *Sure Select Capture Library Mix* gegeben. 2: *Adapter-DNA-Fragmente* + *SureSelect Block Mix* wird zum *Sure Select Capture Library Mix* pipettiert (modifiziert nach Agilent).

Zur Aufreinigung der *Adapter-DNA-Fragmenten* nach der Hybridisierung mussten die magnetischen Streptavidin-Beads (*Dynabeads MyOne Streptavidin T1*, Life Technologies) in mehreren Schritten vorbereitet werden. Der *SureSelect Wash 2* wurde im zirkulierenden Wasserbad von 65°C erwärmt. Für jede zu analysierende Probe wurden 50 µl Streptavidin-Beads in ein Eppendorf-Tube pipettiert und 200 µl *SureSelect Binding Buffer* pro 50 µl Streptavidin-Beads hinzugegeben, gevortext, das Tube in einen Magnetständer gestellt und der klare Überstand verworfen. Analog wurden 2 weitere Waschschritte durchgeführt. Abschließend wurden 200 µl *Sure Select Binding Buffer* für jede zu analysierende Probe zu den Streptavidin-Beads gegeben und die Lösung mit der Pipette gut durchmischt. Jeweils 200 µl gewaschene Streptavidin-Bead-Lösung wurde in ein Well einer 96-Well-Platte überführt.

Die Lösung der hybridisierten *Adapter-DNA-Fragmente* (Abbildung 20, Reihe C) wurden nach mind. 16 h aus dem Termocycler in die Wells der 96-Well-Platte zu den 200 µl gewaschenen Streptavidin-Beads gegeben, die Platte mit einer Klebefolie verschlossen und für 30 min bei 1000 xg auf einen Plattenschüttler gestellt. Anschließend wurde die Platte kurz zentrifugiert, auf den Magnetständer gestellt und der klare Überstand verworfen. 500 µl *SureSelect Wash 1* wurden zu jeder Probe gegeben, gemixt und 15 min bei RT inkubiert. Die Platte wurde erneut auf den Magnetständer gesetzt und der Überstand verworfen. Anschließend wurden 500 µl des vorgewärmten *SureSelect Wash 2* zu jeder Probe pipettiert und die Platte für 10 min bei 65°C inkubiert. Nachfolgend wurde die Platte auf den Magnetständer gestellt und der Überstand verworfen. Dieser Waschschritte mit dem *SureSelect Wash 2* wurde insgesamt 3x. Abschließend wurden die Streptavidin-Bead-Adapter-DNA-Komplexe in 30 µl H₂O resuspendiert. Zur Vereinfachung werden die hybridisierten *Adapter-DNA-Fragmente* einer Probe nach der Selektion mittels Streptavidin-Beads als *Proben-Pool von Target-spezifischen DNA-Fragmenten* bezeichnet.

5.) Amplifikation mit Barcode und Sequenzier-Adapter

Zur Amplifikation der *Proben-Pools von Target-spezifischen DNA-Fragmenten* wurde eine Barcode-PCR durchgeführt und für eine Reaktion ein Mastermix wie folgt zusammenpipettiert:

| Reagenzien | Volumen |
|----------------------------------------------------------------------|------------------------|
| Dest. H ₂ O | 22,5 µl |
| 5x Herculase II Rxn Buffer [★] | 10,0 µl |
| Herculase II Fusion DNA Polymerase [★] | 0,5 µl |
| 100 mM dNTP Mix [★] | 1,0 µl |
| SureSelect ILM indexing Post Capture Forward PCR Primer [#] | 1,0 µl |
| Gesamt | 35 µl/ Reaktion |

[#] Reagenz aus dem *SureSelect Target Enrichment Kit ILM Indexing Hyb Module Box 2*, [★] Reagenz aus dem *Herculase II Fusion DNA Polymerase Kit*

Je 35 µl Mastermix, 15 µl *Proben-Pool von Target-spezifischen DNA-Fragmenten* und 1 µl PCR Primer Index wurden in ein PCR-Tube pipettiert und gut gemixt (für jede Probe wurde ein andere Primer Index verwendet). Die Tubes wurden in den Thermocycler (Biometra) gestellt und das nachfolgende PCR-Programm gestartet:

| Schritt | Zeit | Temperatur | Zyklen |
|---------------|-------------|-------------|--------|
| Denaturierung | 2 min | 98°C | 1x |
| Annealing | 30 sec | 98°C | 14x |
| | 30 sec | 57°C | |
| Elongation | 60 sec | 72°C | |
| Abkühlung | 10 min ∞ | 72°C 4°C | 1x |

Die PCR-Tubes wurden aus dem Thermocycler genommen und der *Proben-Pool von Target-spezifischen DNA-Fragmenten* mit 90 µl AMPure XP Beads (Magnet-Beads) aufgereinigt. Die Durchführung der Aufreinigung mit den Magnet-Beads erfolgte analog zu den in diesem Kapitel beschriebenen Arbeitsschritten des Verfahrens nach der Ligation der dATPs. Die aufgereinigten *Proben-Pools von Target-spezifischen DNA-Fragmenten* wurden abschließend in jeweils 7 µl dest. H₂O eluiert und werden im Folgenden als *Barcode-Proben-Pool* bezeichnet.

Abschließend wurde von jedem *Barcode-Proben-Pool* die DNA-Konzentration mittels Qubit (siehe Abschnitt 2.2.2.1) bestimmt und die Produktgröße mittels Bioanalyzer (siehe Abschnitt 2.2.2.3) von 250 bis 350 bp überprüft (Target-spezifischen DNA-Fragmenten+Adapter+Barcode+Sequenz-Adapter).

6.) Herstellung der Library

Die DNA-Konzentration jedes *Barcode-Proben-Pools* wurde mit dest. H₂O auf 2 nM eingestellt und jeweils 5 µl verdünnter *Barcode-Proben-Pool* (2 nM) zur *Library* zusammenpipettiert.

7.) Sequenzierung mit dem MiSeq

Die ersten mittels *SureSelectXT Panel* (Lymphom-Panel) präparierten *Libraries* wurden nach dem Standard-Protokoll (siehe Abschnitt 2.2.4.3.1) auf 10 bzw. 8 pM normalisiert und mit 5% PhiX versetzt. Es wurde ein *MiSeq Reagent Kit v2* mit 300 Zyklen verwendet und eine Hochdurchsatz-Sequenzierung von 2x 150 bp (*Paired-End*) mit dem MiSeq-Gerät der Firma Illumina durchgeführt. Details zu den einzelnen Illumina-spezifischen Sequenzier-Komponenten und die optimalen Parametern in der Hochdurchsatz-Sequenzierung sind in dem Abschnitt 2.2.5 detaillierter beschrieben.

Im Rahmen der Optimierung des Lymphom-Panels wurde die Normalisierung der präparierten *Libraries* zusätzlich mit dem *Small-Scale-Protokoll* (siehe Abschnitt 2.2.4.3) getestet. Die mittels *Small-Scale-Protokoll* normalisierten *Libraries* wurden mit 1% PhiX versetzt. Die Sequenzier-Kit-Komponenten und-Parameter waren analog zur Hochdurchsatz-Sequenzierung der mittels Standard-Protokoll normalisierten *Libraries* (*MiSeq Reagent Kit v2* mit 300 Zyklen, 2x 150 bp). Die Ergebnisse zeigten, dass der Output der Hochdurchsatz-Sequenzierung nach der Normalisierung der *Library* mittels *Small-Scale-Protokoll* deutlich höher war. Daraufhin wurden alle nachfolgenden (Lymphom-Panel)-*Libraries* ausschließlich mit dem *Small-Scale-Protokoll* normalisiert. Die weiteren Arbeitsschritte zum Starten der Hochdurchsatz-Sequenzierung wurden durchgeführt wie ab dem Kapitel 2.2.4.4 beschrieben.

2.2.4.3 Normalisierung der Library und PhiX-Kontrolle

Die präparierten *Libraries* wurden in dieser Arbeit nach zwei Verfahren normalisiert: erstens, Standard-Protokoll der Firma Illumina und zweitens, *Small-Scale-Protokoll* für *Libraries* mit einer geringeren Konzentration als 2 nM. Zusätzlich wurde häufig die PhiX-Kontrolle verwendet, um technische Artefakte der Sequenzierung besser beurteilen zu können. Die durchgeführten Arbeitsschritte der Normalisierungen sind in den nachfolgenden Abschnitten kurz erläutert.

2.2.4.3.1 Standard-Protokoll

Zur Normalisierung der *Library* erfolgte zuerst die Denaturierung indem 5 µl 2 nM *Library* und 5 µl 0,2 N NaOH zusammenpipettiert und für 5 min bei RT inkubiert wurden. Zur denaturierten *Library* wurden 990 µl gekühlter HT1 Puffer (*MiSeq Reagent Kit v2/v3*) gegeben, sodass abschließend eine denaturierte *Library* von 10 pM vorlag. Die Probenvorbereitung und Sequenzierung jedes Panel bzw. Assay wird zu Beginn getestet und der Input an *Library* für die Hochdurchsatz-Sequenzierung optimiert. Für den optimalen Output an Sequenzen ist es wichtig die richtige Cluster Density bei der Hochdurchsatz-Sequenzierung zu erzielen (Details zu den Sequenzier-Parametern siehe Abschnitt 2.2.5). Für die kommerziellen Panel wurden die Optimierungen der finalen Konzentration an *Library* zuvor von der Firma Illumina durchgeführt und in dieser Arbeit die Angaben des Herstellers eingesetzt. Für das *SureSelectXt Panel* (Lymphom-Panel) erfolgte die Optimierung der finalen Konzentration der *Library* im Rahmen dieser Arbeit. Zu Beginn wurden 10 pM denaturierte *Library* (standardmäßiger Input von vielen kommerziellen *Libraries*) zur Sequenzierung eingesetzt. Diese Konzentration wurde während der Optimierung des *SureSelectXt Panel* (Lymphom-Panel) auf 8 pM denaturierte *Library* (480 µl 10 pM denaturierte *Library* + 120 µl HT1 Puffer) reduziert. Da nur eine mittels *Two-Step-PCR-Assay* (ID3-Assay) präparierte *Library* Hochdurchsatz-sequenziert wurde, wurde keine Optimierung der finalen Konzentration der *Library* des *Two-Step-PCR-Assay* (ID3-Assay) vorgenommen und ausschließlich 10 pM denaturierte *Library* sequenziert.

Zur Herstellung der PhiX-Kontrolle wurden 2 µl 10 nM PhiX (Illumina) und 3 µl 10 nM Tris-Cl gut gemischt und 5 µl 0,2 nM NaOH hinzugegeben. Die verdünnte PhiX-Kontrolle (4 nM) wurde zur Denaturierung für 5 min bei RT inkubiert. Anschließend wurden 990 µl gekühlter HT1 Puffer (*MiSeq Reagent Kit v2/v3*) zur denaturieren PhiX-Kontrolle pipettiert (20 pM). Zur Herstellung der 12,5 pM PhiX-Kontrolle wurden 375 µl 20 pM denaturierte PhiX-Kontrolle mit 225 µl gekühltem HT1 Puffer gemischt und entsprechend den Herstellerempfehlungen zur *Library* hinzugeben *Two-Step-PCR-Assay*: 10%, *TruSight Rapid Capture Panel*: 1%, *SureSelectXT Panel*: 5%. Für das *TruSight Amplicon Panel* sind im Herstellerprotokoll keine Angaben zur Zugabe von der PhiX-Kontrolle vermerkt, weshalb bei diesem Panel auf PhiX verzichtet wurde. Für die weiteren Panel/Assay wurde die denaturierte 12,5 pM PhiX-Kontrolle mit der im vorherigen Abschnitt beschriebenen denaturierten *Library* (10 bzw. 8 pM) nach folgendem Prinzip verdünnt:

| Reagenzien | 1% | 5% | 10% |
|--------------------------------------------|---------------------------|--------|--------|
| 12,5 pM denaturierte PhiX-Kontrolle | 6 µl | 30 µl | 60 µl |
| Denaturierte <i>Library</i> (10 bzw. 8 pM) | 594 µl | 570 µl | 540 µl |
| Gesamt | 600 µl/ MiSeq-Lauf | | |

Abschließend wurde den 600 µl Lösung aus der denaturierten *Library* (+ PhiX-Kontrolle) in die Sequenzier-Kartusche geladen. Dazu wurde mit einer Pipettenspitze das *Sample Port* der Sequenzier-Kartusche vorsichtig durchstoßen und 600 µl der *Library*-Lösung in das Probenreservoir pipettiert.

2.2.4.3.2 *Small-Scale-Protokoll zur Normalisierung der Library*

Die *Library* bestehend aus den Barcode-Proben-Pools wurde auf 0,65 ng/µl mit dest. H₂O eingestellt. Anschließend wurden 5 µl *Library* (0,65 ng/µl) mit 5 µl 0,2 N NaOH denaturiert und die *Library*-NaOH-Lösung für 5 min bei RT inkubiert. Es wurden 584 µl HT1 Puffer (*MiSeq Reagent Kit v2/v3*) und 6 µl 12,5 pM denaturierte PhiX-Kontrolle (Details zur Herstellung siehe unter Abschnitt 2.2.4.3.1) zur denaturierten *Library* hinzugegeben und 600 µl Lösung aus der denaturierten *Library* + PhiX-Kontrolle vollständig in das Probenreservoir der Sequenzier-Kartusche geladen. Im Rahmen der vorliegenden Arbeit wurde das *Small-Scale-Protokoll* für die Normalisierung von Libraries verwendet, die mittels folgender Panel hergestellt wurden:

- *TruSight Rapid Capture Panel* (Cancer-Panel, kommerziell erhältliches Panel)
- *SureSelect XT Panel* (Lymphom-Panel, Custom Panel)

2.2.4.4 Erstellen von *Sample Sheet* und *Manifest* für die Sequenzierung

Für die Hochdurchsatz-Sequenzierung der Library ist es notwendig ein *Sample Sheet* für jeden Lauf zu erstellen, worin die Parameter der Sequenzierung und die Zuordnung der Barcodes zu den jeweiligen Proben aufgelistet werden. Das *Sample Sheet* wurden für das MiSeq Gerät mit dem *Illumina Experiment Manager* (IEM) erstellt. Für die Auswertung von Target-Regionen ist es zusätzlich notwendig ein *Manifest* für das zu analysierende Panel bzw. Assay zu haben, worin die genomischen Target-Regionen (Chromosom, Start- und Endpunkt) klar definiert sind. Das Vorgehen zur Erstellung eines *Sample Sheets* und *Manifests* ist im nachfolgenden kurz beschrieben:

Die IEM wurde geöffnet, das Feld *Create Sample Sheet* auswählen und anschließend das Instrument MiSeq durch das Feld *Next* bestätigt. Nachfolgend wurden die Kategorie *Targeted Resequencing* ausgewählt und für jeden Panel/Assay die Anwendung eingegeben. Für das *TruSight Amplicon Panel* (Myeloid-Panel) und *Two-Step-PCR-Assay (ID3-Assay)* wurde die Kategorie *Amplicon* verwendet und für das *TruSight Rapid Capture Panel* (Cancer-Panel) und das *SureSelectXT Panel* (Lymphom-Panel) die Kategorie *Enrichment*. Die spezifische Kartuschen Nummer des verwendeten *MiSeq Reagent Kits* (MS-Nummer) wurde im Feld *Reagent Cartridge Barcode* eingetragen und das *Sample Prep. Kit* ausgewählt (Myeloid-Panel: *TruSight Amplicon*, ID3-Assay: *TruSeq Amplicon*; Cancer- und Lymphom-Panel: *TruSight Enrichment*). Zusätzlich wurden weitere Informationen wie z.B. *User*, *Experiment* und *Date* eingetragen. Als *Read Type* wurde *Paired-End* angeklickt und die entsprechende Leselänge angeben (Myeloid-Panel: *Cycles Read 1/2 je 150*, ID3-Assay: *Cycles Read 1/2 je 250*, Cancer-Panel: *Cycles Read 1/2 je 150*, Lymphom-Panel: *Cycles Read 1/2 je 100*). Danach wurden unter *workflow* die spezifischen Parameter für die Analyse der Rohdaten mit dem MiSeq-Reporter eingestellt: *Indel Realignment GATK*, *Use Somatic Variant Caller*, *Variant Quality Filter = 30*, *Use Adapter Trimming*. Abschließend wurde für jede zu analysierende Probe der Probenname (*Sample ID*) eintragen und der zugehörige *Barcode* sowie das zu analysierende *Manifest* ausgewählt. Das *Manifest* für das Myeloid- und Cancer-Panel (Illumina) sowie das Lymphom-Panel (Agilent Technologies) wurden von den Herstellern mitgeliefert.

Für den ID3-Assay wurde mittels IEM ein neues *Manifest* angelegt. Dazu wurde das Feld *New* unter *Manifest* im IEM ausgewählt und bei dem Feld *Genome* die Kategorie *Homo-sapiens\UCSC\hg19\Sequence\WholeGenomeFASTA* angeklickt. Die zu analysierenden Target-Regionen (Chromosom, Amplikon-Start, Amplikon-Ende) wurden für den Assay eingetragen, das Gen als Name des *Manifests* verwendet und der Vorgang der Erstellung des Manifests mit *OK* beendet. Die Daten wurden als Text-basierte Datei gespeichert und konnten im Weiteren Prozess der Erstellung des *Sample Sheets* als *Manifest* für den ID3-Assay ausgewählt werden. Abschließend wurde das *Sample Sheet* als .csv-Datei abgespeichert und stand für die Hochdurchsatz-Sequenzierung mittels MiSeq Gerät zur Verfügung.

2.2.4.5 Sequenziervorgang mit dem MiSeq

Zum Starten der Hochdurchsatz-Sequenzierung wurde das *Manifest*, worin die zu analysierenden Target-Regionen aufgelistet sind, und das erstellte *Sample Sheet* (siehe Abschnitt 2.2.4.4) auf die lokale Festplatte des MiSeq-Gerätes geladen. Nachfolgend wurden die *Flow cell* sowie die befüllte Sequenzier-Kartusche (Details siehe Abschnitt 2.2.4.3) in das MiSeq-Gerät geladen. Die einzelnen Arbeitsschritte wurden nach dem *MiSeq User System Userguide* (Stand 2014) der Firma Illumina durchgeführt und sind nachfolgend kurz beschrieben:

Vor der Sequenzierung wurde auf dem Desktop des MiSeq Gerätes das Programm *Manage Instrument* → *Sequence* → *ohne Basespace* ausgewählt. Die *Flow cell* des *MiSeq Reagent Kits v2/v3* wurde mit dest. H₂O, Isopropanol erneut mit dest. H₂O nach Herstellerangaben gewaschen, um Salze der Pufferlösung zu entfernen, in der die *Flow cell* zuvor gelagert wurde. Nachfolgend wurden die *Flow cell*, der *PR2 Buffer* und die beladene Sequenzier-Kartusche in das MiSeq-Gerät geladen und der *Pre-Run* Sequenziervorgang gestartet. Dabei handelt es sich um einen Test vor der Sequenzierung, worin der verfügbare Speicher die *Flow cell* überprüft werden. Nach abgeschlossenem *Pre-Run* Sequenziervorgang wurde der Sequenziervorgang durch das Feld *Run* gestartet. Nach Beendigung der Sequenzierung wurde der Ordner mit den Rohdaten (Output) von der lokalen Festplatte des MiSeq manuell auf den internen Server des Instituts für Humangenetik kopiert (*MiSeq/Rohdaten/Lauf*) und die Daten vom MiSeq gelöscht, um Speicherkapazitäten für neue Sequenzierungen zu gewährleisten. Anschließend wurde ein *Post-Run-Wash* mit einer Lösung von 10% Tween 20 nach Herstellerangaben (Illumina) durchgeführt, wobei das MiSeq-Gerät gereinigt wurde und mögliche Reste der sequenzierten Library aus dem MiSeq-Gerät herausgewaschen wurden.

2.2.5 Bioinformatische Auswertung der Hochdurchsatz-Sequenzierungs-Daten

Im Rahmen dieser Arbeit wurde die Hochdurchsatz-Sequenzierung von Target-spezifischen Regionen mittels MiSeq-Gerät der Firma Illumina durchgeführt. Einen Überblick zur Auswertung der Hochdurchsatz-Sequenzierungs-Daten in Bezug auf die erstellten Datei-Formate und verwendeten Software-Programme/Pipelines sind in der Abbildung 20 zusammengefasst. Die verwendeten Fachbegriffe in dieser Arbeit sind zum Nachschlagen separat in einem Glossar am Anfang der Arbeit zusammengetragen.

Für die Auswertung der Hochdurchsatz-Sequenzierungen wurde für jeden Panel/Assay ein spezifisches *Manifest* von den Regionen von Interesse (engl. region of interests, ROI) und für jede Sequenzierung ein *SampleSheet* erstellt (siehe Abschnitt 2.2.4.4). Diese beiden Dateien waren die Voraussetzung und der Startpunkt für den MiSeq-Lauf sowie die anschließende bioinformatische Analyse.

Nach dem eigentlichen Sequenziervorgang wurde die erste Datenanalyse von dem MiSeq-Gerät durch den MiSeq-Reporter automatisiert vorgenommen. Hierbei wurden drei unterschiedliche Typen von Dateien für jede Probe erstellt.

- Fastq-Dateien, welche die Rohdaten darstellten und Informationen zur Sequenz und deren Qualität enthalten. Diese Dateien wurden zur Auswertung der mittels *SureSelectXT Panel* (Lymphoma-Panel) generierten Hochdurchsatz-Sequenzierungs-Daten verwendet.
- BAM-Dateien, worin die Konsensussequenz aus Vorwärts- und Rückwärtsstrang dargestellt wurde. Diese Dateien wurden zur Auswertung der mittels *Two-Step-PCR-Assay (ID3-Assay)* und *TruSight Rapid Capture Panel* (Cancer-Panel) generierten Hochdurchsatz-Sequenzierungs-Daten verwendet.
- VCF-Dateien (engl. variant call file), die eine Auflistung der detektierten Veränderungen im Vergleich zum Referenzgenom hg19 aufzeigen. Diese Dateien wurden zur Auswertung der mittels *TruSight Amplicon Panel* (Myeloid-Panel) generierten Hochdurchsatz-Sequenzierungs-Daten verwendet.

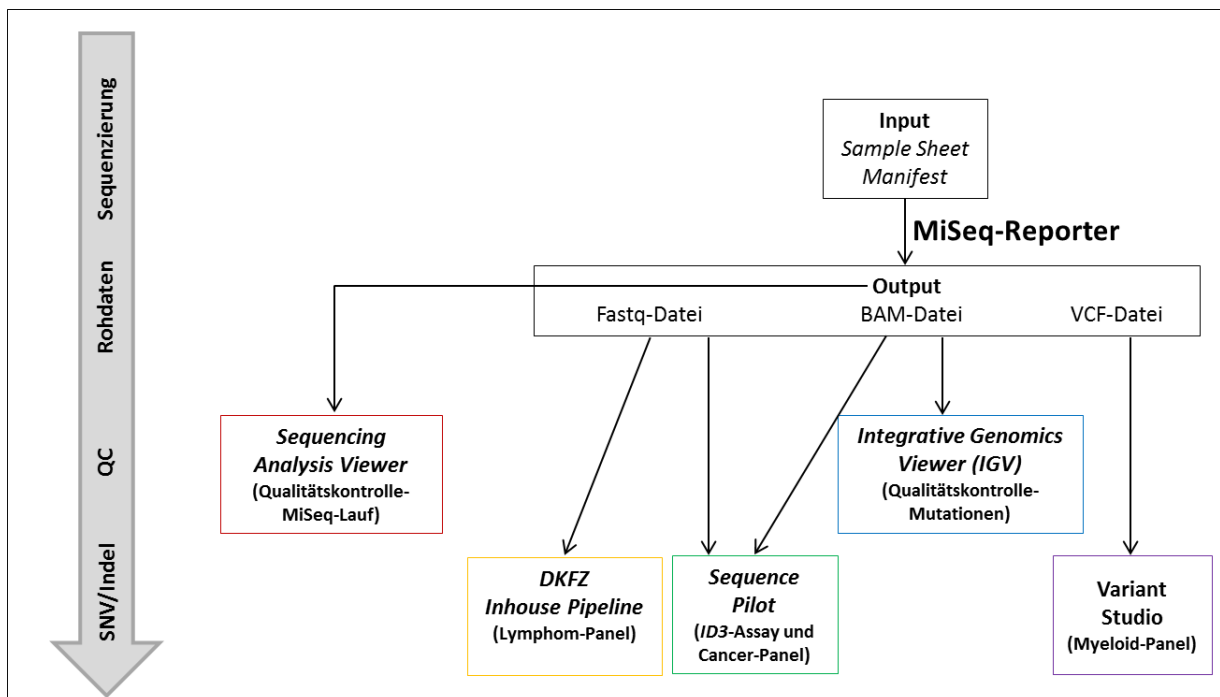


Abbildung 20: Schematische Darstellung des Ablaufprozesses der Datenauswertung mit dem MiSeq

Als Input für die Hochdurchsatz-Sequenzierung wurde das *SampleSheet* und *Manifest* in die Software des MiSeq-Geräts geladen. Mit Hilfe des MiSeq-Reporters wurden aus den Sequenzdaten Fastq-, BAM, und VCF-Dateien erzeugt. Diese Dateien wurden anschließend mit Hilfe von unterschiedlicher Software ausgewertet. Der *Sequencing Analysis Viewer* wurde zur Qualitätskontrolle des MiSeq-Laufs und der *Integrative Genomics Viewer (IGV)* zur Qualitätskontrolle einzelner detektierte Mutationen genutzt. Die Identifizierung von Mutationen (SNV bzw. Indel) erfolgte für das Lymphom-Panel mit einer *inhouse Pipeline* am DKFZ. Das Mutationsprofil der mittels *ID3-Assay* und *Cancer-Panel* analysierten Proben wurde mit der Software *Sequence Pilot* und der mittels *Myeloid-Panel* analysierten Proben mit dem *Variant Studio* (Illumina) bestimmt. Abkürzungen: SNV: engl. single nucleotide variant, Punktmutation. Indel: Insertion oder Deletion. QC: Qualitätskontrolle. DKFZ: Deutsches Krebsforschungszentrum in Heidelberg.

Zur ersten Beurteilung der Qualität der Hochdurchsatz-Sequenzierung wurden die Rohdaten des MiSeq-Laufs in die Software *Sequencing Analysis Viewer* geladen und einzelne Parameter des sequenzierten Laufs betrachtet. Als wichtige Faktoren zur Qualitätskontrolle wurden vier Qualitäts-Parameter (QC-Parameter) angesehen und für alle Hochdurchsatz-Sequenzierungen betrachtet:

- *Cluster Density*: Dieser Parameter beschreibt die Dichte der klonalen Cluster auf der *Flow cell* und liegt optimal bei 800 k/mm² (*MiSeq Reagent Kit v2*) bzw. 1200 k/mm² (*MiSeq Reagent Kit v3*).
- *QC30-Wert*: Dieser Parameter beschreibt den prozentualen Anteil an Reads, die weniger als einen Sequenzier-Fehler auf 1000 Basen zeigten.
- *Cluster Passing Filter*: Dieser Parameter beschreibt den Prozentsatz an Reads, die den Illumina-internen Filter passieren und in der Mutationsanalyse ausgewertet werden.
- *Fehlerrate*: Dieser Parameter beschreibt den Prozentsatz an Sequenzier-Fehlern im MiSeq-Lauf. Der Parameter wird anhand der sequenzierten PhiX-Kontrolle ermittelt, die als interne Illumina-Kontrolle mit einer bekannten Nukleotidsequenzabfolge mitgeführt wurde.

Zwischen diesen vier QC-Parametern besteht ein direkter Zusammenhang d.h. je dichter die *Cluster Density* am Optimum ist desto mehr Reads werden sequenziert und passieren den *Cluster Passing Filter*. Ist die *Cluster Density* jedoch zu hoch, kann die Fluoreszenz keiner Base eindeutig zugeordnet werden, was mit einer höheren *Fehlerrate* und einem geringeren *QC30-Wert* verbunden ist und zu einer erhöhten Anzahl an Reads führt, die nicht durch den *Cluster Passing Filter* gelangen. Die optimalen Qualitätswerte für die einzelnen Parameter sind abhängig von den eingesetzten Sequenzier-Reagenzien sowie der sequenzierten Readlänge (Tabelle 16). Insgesamt ist bekannt, dass die Qualität der Reads mit der Sequenzierlänge abnimmt und v3-Sequenzier-Reagenzien eine höhere *Cluster Density* auf der Flow cell ermöglichen, was zu mehr sequenzierten Reads und einem höheren Output führt.

Tabelle 16: Übersicht der optimalen Qualitätswerte der Sequenzier-Reagenzien des MiSeq Geräts

| Sequenzier-Kit | Flow cell | Reagenz | Zyklen | Cluster Density | Total Reads | Output | QC30 |
|----------------------------------------|-----------|---------|-----------------|---------------------------|-------------|--------|------|
| MiSeq Reagent Nano Kit v2 (500 Zyklen) | Nano | v2 | 500 (2x 250 bp) | 800 (k/mm ²) | 1 Mio. | 0,5 Gb | >75% |
| MiSeq Reagent Kit v2 (300 Zyklen) | Standard | v2 | 300 (2x 150 bp) | 800 (k/mm ²) | 9 Mio. | 5,1 Gb | >80% |
| MiSeq Reagent Kit v2 (500 Zyklen) | Standard | v2 | 500 (2x 250 bp) | 800 (k/mm ²) | 15 Mio. | 7,5 Gb | >75% |
| MiSeq Reagent Kit v3 (300 Zyklen) | Standard | v3 | 300 (2x 150 bp) | 1200 (k/mm ²) | 12,5 Mio. | 7,5 Gb | >75% |

QC30: QC30-Wert. v2/v3: Version der Sequenzier-Reagenzien. Die angegebenen Daten basieren auf den Angaben der Firma Illumina siehe www.illumina.com.

Es ist zu beachten, dass jedes Sequenzier-Kit aus drei Komponenten besteht: Sequenzier-Kartusche, Flow cell und Puffer. Der Puffer ist für jede Hochdurchsatz-Sequenzierung und in jedem Sequenzier-Kit gleich, wohingegen die Kartusche und Flow cell in den entsprechenden Sequenzier-Kits variieren. Die Auswahl des Sequenzier-Kits wird entsprechend den zu sequenzierenden Panel/Assay vorgenommen. Die Kartusche wird anhand der Sequenzlänge (300 bzw. 500 Zyklen) und dem gewünschten Output (v2 bzw. v3) der Library ausgewählt. Bei der Flow cell wird im Normalfall eine Standard-Flow cell genutzt (Output max. 7,5 Gb). Jedoch kann für kleine Panel und Assay auch ein Sequenzier-Kit mit einer kostengünstigeren Nano-Flow cell eingesetzt werden (Output max. 0,5 Gb).

Die vom MiSeq-Reporter gefilterten Reads, d.h. die Reads, die den *Cluster Passing Filter* passieren, wurden anschließend zur Mutationsanalyse in die Auswertungssoftware (*Sequence Pilot* bzw. *Illumina Variant Studio 2.2.*) geladen, worin die erzielte *Coverage* für jede einzelne Target-Region ermittelt wurde. Die *Coverage* ist für alle Datensätze ein entscheidender Parameter und beschreibt wie oft eine spezifische Region vollständig sequenziert wurde. Daher wurden zu Beginn der bioinformatischen Analysen für alle analysierten Assays/ Panels folgende zwei wichtige Parameter bestimmt, die sich aus der *Coverage* jeder Target-Region ermitteln lassen:

- *On-Target-Rate* (OTR): Prozentualer Anteil aller im Panel/Assay-Design enthaltenen Target-Regionen gegenüber den tatsächlich für jede Probe *gcovered* bzw. sequenzierten Target-Regionen.
- *On-Target-Coverage* (OTC): Beschreibt wie oft jede Target-Region für die einzelnen Proben tatsächlich sequenziert wurde.

Zur Identifizierung des Mutationsprofils der einzelnen Proben aus den Hochdurchsatz-Sequenzierungs-Daten ist es wichtig, dass jede analysierte Target-Region mit einer bestimmten *Coverage* abgedeckt wurde, um möglichst alle Mutationen zu detektieren. Bei der Analyse von Tumor-Proben sollte vorzugsweise die *Coverage* möglichst hoch sein, da der Tumorzellgehalt der Proben häufig ungewiss ist und auch Zellen vom Normalgewebe in der Probe sequenziert werden. Bei einem geringen Tumorzellgehalt ist es nur mit einer hohen *Coverage* (Tiefensequenzierung) der Target-Regionen möglich die Mutationen zu detektieren.

Je nach Zielsetzung der Projekte sowie eingesetzten Assays und Panel wurden unterschiedliche Schwellenwerte der *Coverage* für die einzelnen Hochdurchsatz-Sequenzierungen festgelegt. Die Grenzwerte der *Coverage* für die Teilprojekte, sowie die gemessenen Qualitätswerte für die durchgeführten MiSeq-Läufe innerhalb der vorliegenden Arbeit, sind in den Ergebnissen der einzelnen Assays bzw. Panel separat aufgelistet und beschrieben. Das generelle Vorgehen für die einzelnen Software-Programme und die Arbeitsschritte insbesondere zur Auswertung der Hochdurchsatz-Sequenzierungs-Daten im Rahmen der altersassoziierten Analysen von genetischen Aberrationen werden in den nachfolgenden Abschnitten genauer beschrieben.

2.2.5.1 Qualitätskontrolle des MiSeq-Laufs mittels der Software *Sequencing Analysis Viewer* von Illumina

Die Software *Sequencing Analysis Viewer* (ISAV) der Firma Illumina ermöglicht es eine Übersicht verschiedener Parameter des MiSeq-Laufs zu bekommen. Unter Anwendung des Manuals der Firma Illumina wurde der Output-Ordner des MiSeq-Laufs in die Software geladen und die vier zuvor beschriebenen QC-Parameter zur Qualitätskontrolle betrachtet (siehe Abschnitt 2.2.5). Da es von der Firma Illumina keine Schwellenwerte für gute und schlechte QC-Parameter gibt, wurden (*inhouse*)-QC-Werte festgelegt, die bei der Hochdurchsatz-Sequenzierung erzielt werden mussten. Die gewählten QC-Werte basierten auf den optimalen QC-Werten der MiSeq-Reagenzien die Illumina angibt sowie den eigenen Erfahrungen der Hochdurchsatz-Sequenzierungen. Die erzielten QC-Werte für die Hochdurchsatz-Sequenzierungen dieser Arbeit sind detailliert in den Ergebnissen unter Abschnitt 3.1. nachzulesen.

Für die weitere Mutationsanalyse wurden folgende minimale QC-Werte gewählt:

- 1.) *Q30-Wert* ist größer als 75%.
- 2.) *Cluster Passing Filter* beträgt mind. 70%.
- 3.) *Fehlerrate* ist max. 1%.

MiSeq-Läufe, die diese (*inhouse*)-QC-Werte nicht erfüllten, wurden erneut sequenziert und ggf. die *Library* Präparation wiederholt. Alle durchgeführten MiSeq-Läufe die diese (*inhouse*)-QC-Werte erfüllten, gingen in die weitere Auswertung ein. Dazu wurden Mutationsanalysen der sequenzierten Proben in verschiedenen Software-Programmen durchgeführt. Hierbei ist zu beachten, dass die Auswertung mit der Software *Variant Studio* und *Sequence Pilot* keine parallele Analyse eines Datensatzes von mehreren Proben ermöglichte und jede zu analysierende Probe separat in der Software betrachtet und ausgewertet wurde. Die Arbeitsschritte zur bioinformatischen Auswertung der einzelnen Panel/Assays werden in den nachfolgenden Kapiteln genauer beschrieben.

2.2.5.2 Qualitätskontrolle der mittels Hochdurchsatz-Sequenzierung identifizierten Mutationen im *Integrativen Genome Browser (IGV)* des Broad Instituts

Zur Qualitätskontrolle der identifizierten Mutationen wurden einzelne Mutationen im *Integrativen Genome Browser (IGV)* des Broad Instituts visualisiert^{223,224}, um technische Artefakte bestmöglich auszuschließen. Die ausgewählten Mutationen die mittels *IGV* detailliert betrachtet wurden, umfassten vorzugsweise Mutationen, die mit einer geringen Mutationsallelfrequenz (MAF) von unter 10% detektiert wurden. Weiterhin wurden genomische Regionen überprüft, worin viele Mutationen nacheinander detektiert wurden (mehr als 3 Mutationen in 50 bp). Zur Überprüfung der Mutationen wurde die webbasierte Version des *IGV* geöffnet und die zuvor vom MiSeq-Reporter erstellten BAM-Dateien eingeladen. Es wurden parallel die BAM-Dateien mehrerer Tumor-Proben, verschiedener Tumor-Entitäten sowie von gesunden Individuen (Kontroll-Proben) aus verschiedenen MiSeq-Läufen eingeladen und begutachtet. Mutationen, die mit einer geringen MAF von unter 10% in den Sequenzen verschiedener Tumor-Entitäten und gleichermaßen bei Tumor- und Kontroll-Proben identifiziert wurden, wurden als Artefakte angesehen und aus den weiteren Analysen ausgeschlossen.

2.2.5.3 Analyse mittels der *Software Variant Studio* von Illumina

Zur Analyse des *TruSight Amplicon Panels* (Myeloid-Panel) wurde die herstellereigene Software *Illumina Variant Studio (IVS)* verwendet. Hierbei handelt es sich um eine lokal-installierte Software zur Auswertung von VCF-Dateien, worin die zuvor mittels MiSeq-Reporter detektierten Mutationen, selektiert wurden. Die Selektion erfolgte automatisiert durch die Einstellung verschiedener Parameter in den Filtern der Software. Unter Anwendung des Manuals der Firma Illumina wurden zuerst in der Kategorie *Variant* das Feld *Pass Filter* bestätigt und *Quality >80* gesetzt. Bei diesen Parametern handelt es sich jeweils um Software-spezifische Filter der Firma Illumina, wodurch Reads mit einer schlechten Qualität aus der Mutationsanalyse ausgeschlossen wurden. Nachfolgend konnten unter der Verwendung weiterer Filter (z.B. Coverage, Gene) zusätzliche Sequenz-Daten aus den Mutationsanalysen ausgeschlossen werden. Im Rahmen der vorliegenden Arbeit wurden folgende Filter zur Analyse des *TruSight Amplicon Panels* (Myeloid-Panels) verwendet:

- *Coverage* $\geq 50x$: Die analysierte Target-Region wurde mind. 50-fach in der analysierten Probe sequenziert.
- Mutationsallelfrequenz (MAF) $\geq 15\%$: Die detektierte (MAF) betrug mind. 15% in der analysierten Probe
- Nur proteinverändernde Mutationen: Es wurden alle potentiell proteinverändernde Mutationen betrachtet inkl. single nucleotide polymorphism (SNP).

Alle identifizierten potentiell proteinverändernden Mutationen (inkl. Annotationen) wurden für jede analysierte Probe als Tabelle in Form einer .tsv-Datei abgespeichert. Anschließend wurde eine Datenbank-Recherche mittels dbSNP132, 1000 Genomes (1000 K), Exome Variant Server (ESP6500), Exome Aggregation Consortium (ExAC) und Catalogue of Somatic Mutations in Cancer (COSMIC) durchgeführt, wonach alle detektierten Mutationen mit einer MAF von über 1% in der Normalbevölkerung aus den weiteren Analysen ausgeschlossen wurden.

2.2.5.4 Analyse mittels der Software *Sequence Pilot* von JSI medical systems

Zur Analyse der Hochdurchsatz-Sequenziers-Daten des *Two-Step-PCR-Assays* (*ID3-Assay*) und dem *TruSight Rapid Capture Panel* (Cancer-Panel) wurde die kommerziell erhältliche Software *Sequence Pilot* mit dem Modul *SeqNext* verwendet (*JSI medical systems*). Diese Software ermöglicht es sowohl BAM- als auch Fastq-Dateien zu analysieren. In dieser Arbeit wurden ausschließlich BAM-Dateien verwendet und die Mutationsanalyse nach dem Manual des Herstellers durchgeführt. In der Auswertung mit dem Modul *SeqNext* wurden folgende Parameter angewandt:

- Alignment und Mapping mit dem Modul *SeqNext*
- Coverage $\geq 10x$
- Sequenzierinformationen vom Vorwärts- und/oder Rückwärtsstrang lagen vor

Die identifizierten Mutationen wurden in einer .pdf-Datei als Report abgespeichert. Anschließend wurden nur Splicing- und exonische Mutationen betrachtet und alle potenziell proteinverändernde Mutationen gegen verschiedene Datenbanken (dbSNP, 1000 K, ESP6500) abgeglichen. Dabei wurden bereits bekannte Sequenzveränderungen (SNPs), die in den Datenbank mit einer MAF von über 1% in der Normalbevölkerung aufgeführt waren, aus den weiteren Analysen ausgeschlossen. Abschließend wurden alle potentiell proteinverändernden Mutationen in einer .xlsx-Tabelle zusammengefasst.

2.2.5.5 Analyse mittels der *inhouse Pipeline* des Deutschen Krebsforschungszentrums

Die Auswertung der mittels *SureSelectXT Panel* (Lymphom-Panel) generierten Hochdurchsatz-Sequenzierungs-Daten wurde am Deutschen Krebsforschungszentrum (DKFZ) in Heidelberg in der Arbeitsgruppe zur Theoretischen Bioinformatik von Herrn Dr. Matthias Schlesner durch Herrn Dr. Bingding Huang durchgeführt. Im Rahmen der altersassoziierten genetischen Analysen wurden in der vorliegenden Arbeit fünf verschiedene Kohorte am DKFZ ausgewertet: *Jack* (siehe Abschnitt 2.1.2.2), *MMML-Senil* (siehe Abschnitt 2.1.2.1), *Vergleichsstudien* (siehe Abschnitt 2.1.2.1), *Kontrollen-Sehr alt* und *Kontrollen-Adult* (siehe Abschnitt 2.1.2.5). Hierbei ist zu beachten, dass die Sequenzierungs-Daten von *Jack*, *MMML-Senil* und den *Kontrollen-Adult* bzw. *-Sehr alt* mittels Target-Resequenzierung mit dem *SureSelectXT Panel* (Lymphom-Panel) in dieser Arbeit generiert wurden. Die Sequenzierungs-Daten der *Vergleichsstudien* kommen aus dem ICGC MMML-Seq Projekt, wobei es sich um *Whole Genome Sequencing* (WGS)-Daten der Patienten handelt. Zusätzlich ist beachten, dass in der Mutationsanalyse der vorliegenden Arbeit die Kohorte *Kontrollen-Sehr alt* nicht als Normalkontrollen verwendet wurden, da es sich bei diesen Proben um Gewebe mit einer entzündlichen Veränderung (Lymphadenitis) handelte und diese Lymphadenitis-Proben ggf. Lymphom-ähnliche genetische Veränderungen aufweisen. Die Auswertung der *Kontrollen-Sehr alt* erfolgte analog zu den Tumor-Proben aus *Jack* und *MMML-Senil*. Für die altersassoziierten Analysen der genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden daher nur die *Kontrollen-Adult* als Normalkontrollen verwendet und zu Vergleichszwecken eingesetzt. Eine Übersicht der Vorgehensweise bei der bioinformatischen Auswertung der Hochdurchsatz-Sequenzierungs-Daten der verschiedenen Kohorten soll das Flussdiagramm in Abbildung 21 darstellen. Zur Vereinfachung wurde in dieser Grafik die Kohorte *Kontrollen-Sehr alt* nicht separat aufgeführt.

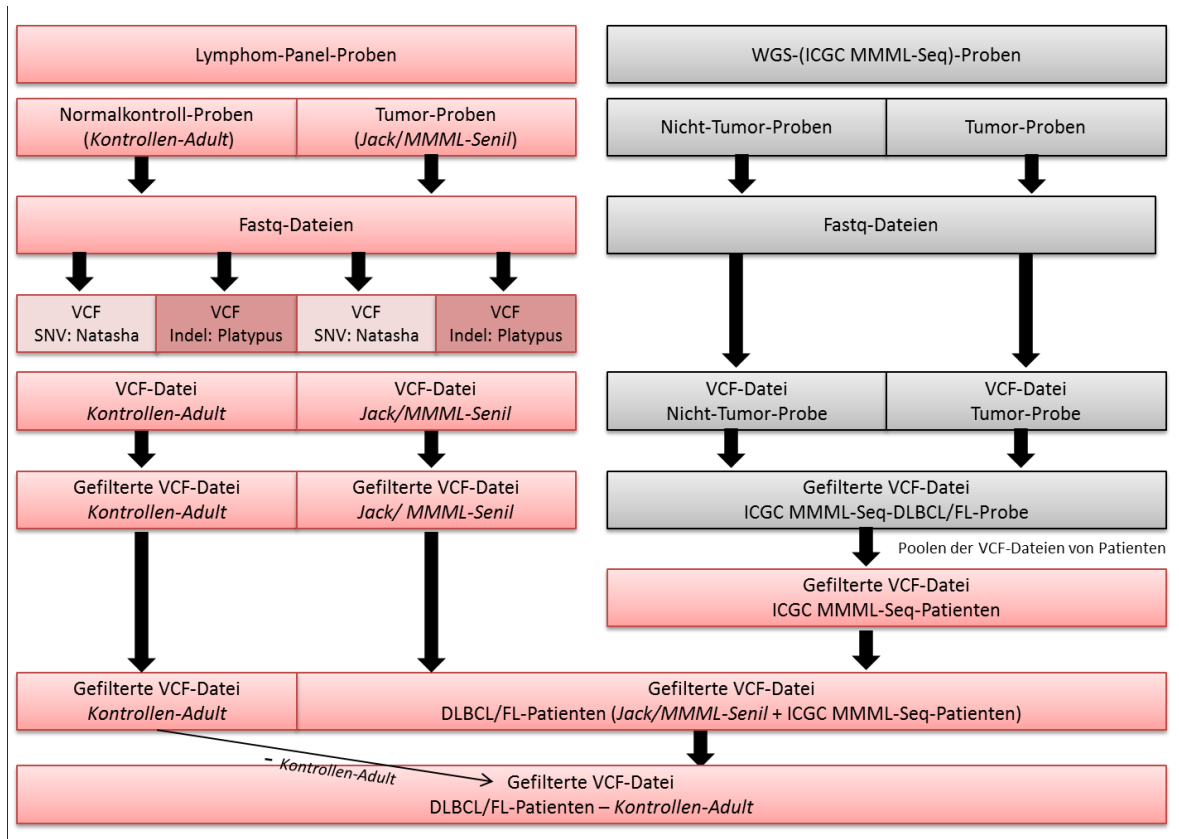


Abbildung 21: Schematische Darstellung der bioinformatischen Auswertung mittels der DKFZ inhouse Pipeline

Zur Analyse der altersassoziierten genetischen Aberrationen bei Patienten mit einem Keimzentrums-B-Zell-Lymphom wurden vier verschiedene Kohorte betrachtet: 1.) *Kontrollen-Adult*, 2.) *Jack* 3.) *MMML-Senil* und 4.) *Vergleichsstudien* (DLBCL- und FL-Patienten des ICGC MMML-Seq Projektes). In rot markiert ist Auswertung der mittels *SureSelectXT Panel* (Lymphom-Panel) sequenzierten Proben dargestellt. Bei dieser Auswertung handelt es sich um Kohorten von mehreren Patienten. Das Mutationsprofils der Kohorten *Kontrollen-Adult*, *Jack* und *MMML-Senil* wurde basierend auf den Rohdaten der Hochdurchsatz-Sequenzierungs-Daten (Fastq-Dateien) der mittels Lymphom-Panel analysierten Patienten bestimmt und die identifizierten Aberrationen in VCF-Dateien gespeichert. Die Auswertung der VCF-Dateien wurde mittels einer *DKFZ-inhouse-Pipeline* durchgeführt, wobei SNVs mit Hilfe des Programms *Natasha* (*DKFZ-inhouse*) und Indels unter Verwendung des Programms *Platypus*²²⁵ (Trust Centre for Human Genetics der Universität von Oxford) detektiert wurden. In grau ist die Auswertungen der *Whole Genome Sequencing* (WGS)-Daten der ICGC MMML-Seq-Proben dargestellt. Hierbei wurden in der primären Analyse jeweils die Mutationen von Tumor- und Nicht-Tumor-Proben eines Patienten identifiziert. Für die altersassoziierte Mutationsanalyse wurde der Pool der gefilterten VCF-Dateien (somatische Mutationen eines Patienten) von allen analysierten DLBCL- und FL-Patienten der ICGC MMML-Seq-Kohorte verwendet und auf die Target-Regionen des Lymphom-Panels reduziert. Die Mutationstabellen der ICGC-Kohorte und der mittels Lymphom-Panel analysierten Tumor-Proben (*Jack*, *MMML-Senil*) wurden in einer Datei zusammengefasst (DLBCL/FL-Patienten) und gegen alle detektierten SNVs und Indels der *Kontrollen-Adult* gefiltert. Abschließend erfolgte eine altersassoziierte Mutationsanalyse der Keimzentrums-B-Zell-Lymphome. Abkürzungen: SNV: engl. single nucleotide variant, Punktmutation. Indel: Insertion oder Deletion. VCF: engl. variant call file. FL: Follikuläres Lymphom. DLBCL: Diffus großzelliges B-Zell-Lymphom. WGS: engl. whole genome sequencing, Komplet-Genom-Sequenzierung. ICGC MMML-Seq: Verbundprojekt des internationalen Krebsgenomkonsortiums.

1.) Qualitätskontrolle des Lymphom-Panels

Zu Beginn der Analyse wurde eine Qualitätskontrolle der mittels Lymphom-Panel generierten Hochdurchsatz-Sequenzierungs-Daten durchgeführt und verschiedene QC-Parameter für jede analysierte Probe bestimmt (Total Reads, Mapping-Rate, Total-Mapped-Reads, Duplikat-Rate, Nicht-Duplikat-Reads, On-Target-Reads, Ø-OTR, Ø-OTC, Ø-Readlänge). Basierend auf dem Schwellenwert der WGS-Daten aus dem ICGC von einer *Coverage* von mind. 30x für Tumor-Proben wurde als Einschlusskriterium für die Mutationsanalysen eine min. Ø-OTC von 30x über alle Gene des Lymphom-Panels festgelegt. Alle Patienten mit einer geringeren Ø-OTC wurden aus den weiteren Analysen ausgeschlossen.

Zur Validierung der Ergebnisse wurden einzelne Mutationen mittels Sanger-Sequenzierung analysiert (N= 20). Die Selektionskriterien für die Auswahl der Proben sowie die durchgeführten Arbeitsschritte zur Validierung der Mutationen mittels Sanger-Sequenzierung wurden bereits im Abschnitt 2.2.3 detailliert beschrieben.

2.) Lymphom-Panel-Proben

Die *Kontrollen-Adult* wurden als Pool zusammengefasst und werden nachfolgend als Normalkontrollen bezeichnet. Die Tumor-Proben aus den Kohorten *Jack* und *MMML-Senil* werden im Folgenden als Lymphom-Panel-Tumor-Proben bezeichnet. Die Gesamtkohorte der mittels Lymphom-Panel analysierten Normalkontroll- und Tumor-Proben wird im Folgenden als Lymphom-Panel-Proben bezeichnet.

3.) Mutationsanalyse und Selektionskriterien der Lymphom-Panel-Proben

Die Identifizierung der SNVs von den Lymphom-Panel-Proben wurde mit der *inhouse*-Pipeline *Natasha* des DKFZ durchgeführt^{230,233,234}. Die Indels wurden mit Hilfe von *Platypus* identifiziert²³⁵. Anschließend wurden alle detektierten Mutationen gegen verschiedene Datenbanken (dbSNP, 1000K, ESP6500) für häufige Allelfrequenzen (SNPs) gefiltert und SNPs mit einer Häufigkeit von mind. 1% in der Normalbevölkerung aus den weiteren Analysen ausgeschlossen.

4.) Mutationsanalyse und Selektionskriterien der Vergleichskohorte

Als Vergleichskohorte für die Mutationsanalysen der Keimzentrums-B-Zell-Lymphome von senilen Patienten dienten die *Whole Genome Sequencing* (WGS)-Daten aus dem ICGC MMML-Seq Projekt. Dieser Datensatz bestand größtenteils aus adulten DLBCL- und FL-Patienten, wodurch das Altersspektrum der Patienten mit einem Keimzentrums-B-Zell-Lymphom und einer Erstdiagnose bis zu 80 Jahren abgedeckt wurde. Im Rahmen des Verbundprojektes ICGC MMML-Seq wurden von jedem Patienten die Tumor-Probe sowie die korrespondierende nicht- Tumor-Probe (überwiegend DNA aus dem peripheren Blut des Patienten) sequenziert, um die Keimbahnmutationen in den weiteren Mutationsanalysen auszuschließen. Aus dem ICGC MMML-Seq Projekt lag eine Mutationstabelle vor, die alle detektierten somatischen Mutationen der analysierten Patienten enthielt (Stand von Juni 2015). Diese Mutationstabelle wurde im ersten Auswertungsschritt auf die DLBCL- und FL-Patienten und die mittels Lymphom-Panel analysierten Target-Regionen reduziert. Die reduzierte Mutationstabelle der ICGC MMML-Seq-Proben diente als Basis für die altersassoziierten Analysen und wird im Folgenden als Mutationstabelle der Vergleichskohorte bezeichnet.

5.) Altersassoziierte Analysen von genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen

Für die altersassoziierte Auswertung der genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden die zuvor beschriebenen Mutationstabellen der Lymphom-Panel-Proben und der Vergleichskohorte zu einer Gesamtmutationstabelle zusammengefasst. Für die weiteren altersassoziierten Analysen wurde eine Gruppierung der analysierten Patienten mit einem DLBCL bzw. FL hinsichtlich des Ersterkrankungsalters vorgenommen (unabhängig von der zugehörigen Studie). Hierzu wurden die Altersgruppen: Gruppe 1 (0-39 Jahre), Gruppe 2 (40-59 Jahre), Gruppe 3 (60-79 Jahre) sowie Gruppe 4 (80-97 Jahre) eingeführt. Daraufhin erfolgten altersassoziierte Analysen der genetischen Aberrationen die in der Gesamtmutationstabelle zusammengetragen wurden (Stand vom 11. Juni 2015). Alle Gene, die in mind. fünf DLBCL bzw. FL-Patienten rekurrent mutiert vorlagen, wurden grafisch in zwei unterschiedlichen Formen dargestellt:

- **Oncoprints:** Darin wurden die verschiedenen Mutationstypen (Frameshift Deletion, Frameshift Insertion, Nicht-Frameshift Deletion, Nicht-Synonym, Splicing, Stopp-Mutation) in den rekurrent mutierten Genen für die analysierten Patienten dargestellt. Zusätzlich wurden die Keimzentrums-B-Zell-Lymphome den o.g. Altersgruppen zugeordnet und verschiedene Patienteninformationen in die Darstellung integriert und visualisiert (z.B. Erstdiagnosealter, FISH-Daten und die Gesamtanzahl an identifizierten Mutationen).
- **Boxplots:** Hierzu wurde das Ersterkrankungsalter des Patienten in Bezug auf das rekurrent mutierte Gen betrachtet und die Ergebnisse grafisch in Form von Boxplots dargestellt.

2.2.6 Analyse der DNA-Methylierung mittels Methylierungsarrays

Die DNA-Methylierung wurde unter Anwendung der Infinium HumanMethylation450 BeadChips, die im Folgenden als 450K bezeichnet werden, durchgeführt. Im Rahmen der vorliegenden Arbeit wurde das Frischmaterial von 28 Tumor-Proben der Kohorte *MMML-Senil* (siehe Abschnitt 2.1.2.1) und von 8 Normalkontroll-Proben der Kohorte *Kontrollen-Sehr alt* (siehe Abschnitt 2.1.2.5) mittels 450K analysiert. Die technische Durchführung der Methylierungsarrays wurde freundlicherweise von Frau Lorena Valles durchgeführt. In den nachfolgenden Abschnitten werden die durchgeführten Arbeitsschritte kurz beschrieben.

2.2.6.1 Natriumbisulfit-Konvertierung von DNA-Proben

Die Bisulfit-Konvertierung wurde mit dem EZ DNA Methylation Kit (Zymo Research) nach Herstellerangaben durchgeführt. Eine detaillierte Beschreibung des Verfahrens der Bisulfit-Konvertierung wurde zuvor in der Einleitung gegeben und ist dort unter Abschnitt 1.7.3 nachzulesen. Als Input wurde 1 µg gDNA in 45 µl H₂O eingesetzt und mit 5 µl M-Dilution Puffer versetzt. Die DNA-Probe wurde für 15 min bei 37°C inkubiert und anschließend 100 µl Konversions-Reagenz hinzugegeben. Die Probe wurde in den Cyclor gegeben und das Konvertierungsprogramm gestartet.

| Schritt | Zeit | Temperatur | Zyklen |
|---------------|--------|------------|--------|
| Denaturierung | 30 sec | 95°C | 16x |
| Konvertierung | 60 min | 50°C | |
| | ∞ | 4°C | |

Danach wurden 400 µl Binde-Puffer zur DNA-Lösung gegeben, dieser Mix auf die Zymo-Spin-Säule geladen und das Tube mit der Säule für 30 sec zentrifugiert. Der Durchfluss wurde verworfen und die Säule mit 100 µl M-Wasch Puffer gewaschen. 200 µl M-Desulphonierungs Puffer wurden hinzugegeben und für 20 min bei RT inkubiert. Die Säule wurde erneut zentrifugiert und der Waschschrift wiederholt. Abschließend wurde die konvertierte DNA in 10 µl H₂O eluiert und für die Analyse mittels Infinium HumanMethylation450 BeadChip verwendet.

2.2.6.2 Infinium HumanMethylation450 BeadChip (450K)

Für die Analyse der genomweiten DNA-Methylierung wurde der Infinium HumanMethylation450 BeadChip (450K) der Firma Illumina, der zur Technologie der Microarrays zählt, angewandt. Dieser Bead Chip wird im Folgenden als 450K bezeichnet. Ein 450K ist aus einer Matrix aufgebaut, auf der viele einzelsträngige Oligonukleotide immobilisiert wurden. Mit Hilfe des 450K ist es möglich eine gleichzeitige Analyse des Methylierungsprofils von 485.577 differentiellen CpGs in 21.231 Genen durchzuführen. Das Prinzip des 450K wurde zuvor in der Einleitung genauer beschrieben und ist dort dem Abschnitt 1.7.3 nachzulesen.

Die Durchführung des 450K erfolgte mit dem Protokoll „*Illumina Infinium HD Methylation Assay, Stand 2011*“ nach Herstellerangaben. Eine kurze schematische Beschreibung der Arbeitsabläufe der Methylierungsanalyse mittels 450K ist in dem nachfolgenden Abschnitt zusammengestellt. Sofern keine Herstellerangaben angegeben wurden, stammten die verwendeten Reagenzien aus dem *Infinium HumanMethylation450 BeadChip Kit* (Illumina).

Zur Bestimmung der DNA-Methylierung wurden die Natriumbisulfit-konvertierten DNA-Proben enzymatisch fragmentiert und mittels absolutem Isopropanol (Sigma Aldrich) präzipitiert. Das Präzipitat wurde mit einem Hybridisierungspuffer resuspendiert und die DNA-Hybridisierungslösung auf die 450K pipettiert. Es folgte eine Inkubation der 450K für mind. 16 h bzw. max. 22 h bei 48 °C im Hybridisierungssofen für 450K (Illumina). Während dieser Zeit hybridisierten die fragmentierten Natriumbisulfit-konvertierten DNA-Moleküle an die komplementäre Oligonukleotid-Sequenz auf dem 450K. Nicht hybridisierte DNA-Moleküle wurden anschließend durch Waschschrte entfernt. Nachfolgend wurden verschiedene mit Fluoreszenzfarbstoff-markierte dNTPs auf die 450K gegeben, wobei diese entsprechend der Basenabfolge komplementär in die DNA-Fragmente eingebaut und somit die Oligonukleotid-Sequenz verlängert wurde. Im Anschluss wurden Fluoreszenz-markierte Antikörper auf die 450K pipettiert, wobei diese an die Biotin-Markierung der eingebauten Nukleinbasen binden. Abschließend wurde die Emission der Fluorophore auf dem 450K durch ein iScan Gerät (Illumina, Grundeinstellung des iScans entsprechend der zu scannenden 450K) detektiert und in Form von .idat-Dateien gespeichert.

2.2.7 Bioinformatische Auswertung der DNA-Methylierungsdaten

Im Rahmen der altersassoziierten epigenetischen Analysen wurden in der vorliegenden Arbeit vier verschiedene Kohorte ausgewertet: *MMML-Senil* (siehe Abschnitt 2.1.2.1), *Vergleichskohorte* von Tumor-Proben für die DNA-Methylierungsanalysen (siehe Abschnitt 2.1.2.1), *Kontrollen-Sehr alt* und *Kontrollen-Jung* (siehe Abschnitt 2.1.2.5). Hierbei ist zu erwähnen, dass die 450K-Daten von *MMML-Senil* und den *Kontrollen-Sehr alt* in dieser Arbeit generiert wurden und die 450K-Daten der *Vergleichskohorte* und *Kontrollen-Jung* aus dem MMML- und/oder ICGC MMML-Seq Projekt stammen. Für die 450K-Analysen der Normalkontrollen ist zu beachten, dass sich die Kohorte *Kontrollen-Jung* (N= 8) aus jeweils zwei FACS-sortierten Zell-Populationen von vier verschiedenen gesunden Individuen zusammensetzt. Hingegen besteht die Kohorte *Kontrollen-Sehr alt* aus acht Patienten mit einer Lymphadenitis (8 unterschiedliche Individuen). Für die altersassoziierten Analysen der epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurde eine Pipeline unter Anwendung verschiedener Software-Programme angewandt. Ein Überblick zum Vorgehen bei der 450K-Analyse soll das Flusschema in Abbildung 22 geben.

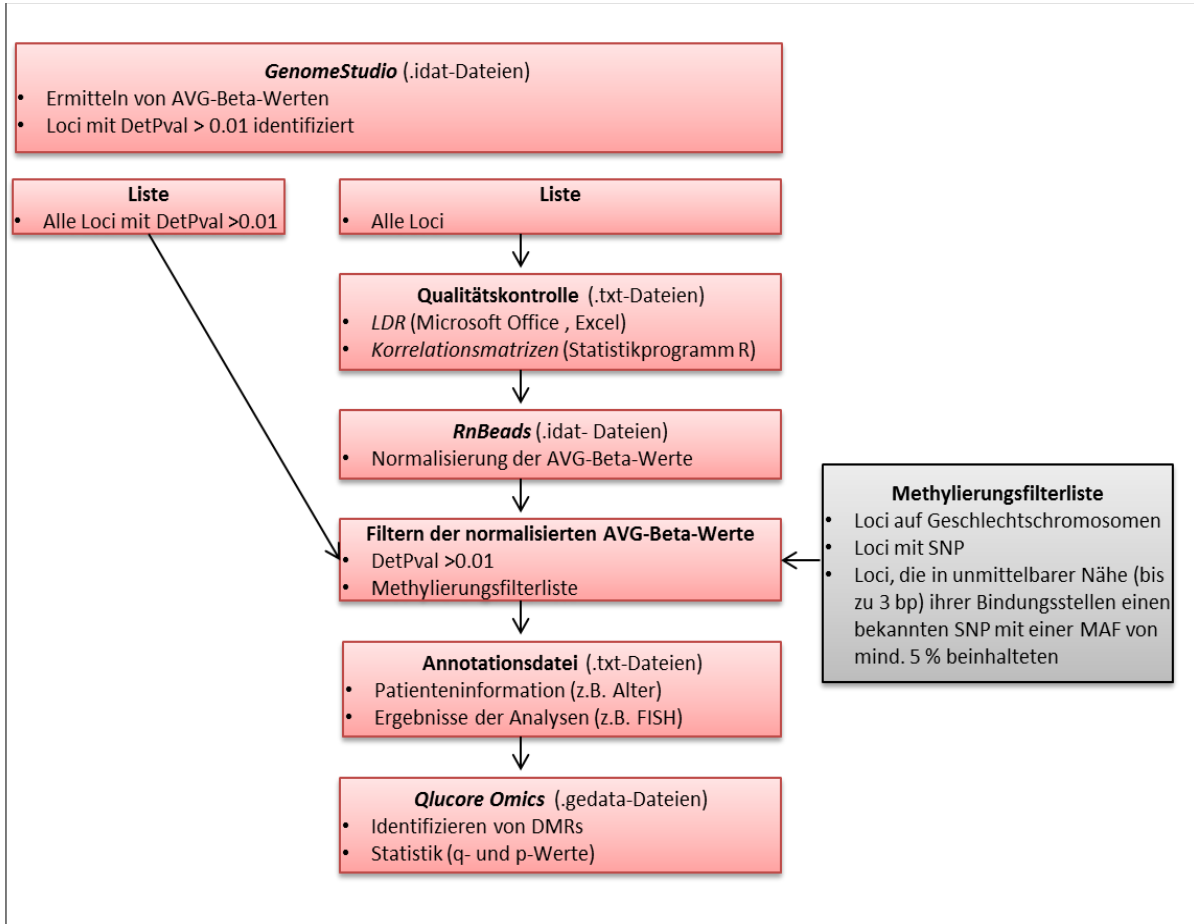


Abbildung 22: Schematische Darstellung der Auswertung der 450K-Methylierungsdaten

Die Rohdaten (.idat-Dateien) des zu analysierenden Datensatzes wurden in die *GenomeStudio* Software von Illumina eingeladen. Mit Hilfe der Software wurden die Methylierungswerte (AVG-Beta-Werte) für alle analysierten Loci ermittelt und eine Liste von allen Loci mit einem *Detection P-value* (*DetPval*) $\leq 0,01$ erstellt. Zur Qualitätskontrolle wurden *Korrelationsmatrizen* erstellt und die *Loci detection rate* (*LDR*) berechnet. Alle Proben des Datensatzes die in der *Korrelationsmatrix* unauffällig waren und alle Loci die eine *LDR* >99% hatten, wurden mittels *RnBeads* normalisiert. Anschließend wurden alle Loci mit *DetPval* $\leq 0,01$, die zuvor in *GenomeStudio* ermittelt wurden, herausgefiltert. Zusätzlich wurden alle Loci einer Methylierungs-Filter-Liste (interne Filterliste des Instituts für Humangenetik) aus der weiteren Analyse ausgeschlossen. Nachfolgend wurde eine Annotationsdatei erstellt und für alle Patienten des Datensatzes zusätzliche Informationen und Analyseergebnisse (z.B. Alter, FISH-Daten) eingetragen. Die Annotationsdatei wurde als .gedata-Datei umgespeichert und in die Software *Qlucore Omics* eingeladen. Abschließend wurde eine altersassoziierte Analyse des DNA-Methylierungsprofils vorgenommen und die dazugehörigen statistischen Parameter (p- und q-Wert) bestimmt. Abkürzungen: DMR: Differenziell methylierte Region. FISH: Fluoreszenz-insitu-Hybridisierung.

Zur Auswertung der DNA-Methylierung wurden zuerst die Rohdaten (.idat-Dateien) in das *Methylation Modul* der *GenomeStudio* Software (Illumina) geladen und die Methylierungswerte (AVG-Beta-Werte) für alle analysierten Loci generiert. Anschließend wurde die Qualität der Methylierungsdaten durch das Ermitteln der *Loci detection rate* (LDR) und der Erstellung von *Korrelationsmatrizen* kontrolliert. Es folgte die Hintergrundnormalisierung und Korrektur der Signalintensitäten mit dem Statistikprogramm R unter Verwendung des Pakets *RnBeads*. Die normalisierten AVG-Beta-Werte wurden abschließend mit der Software *Qlucore Omics Explorer* in einer logistischen Regression hinsichtlich altersassoziierter Veränderungen der DNA-Methylierung analysiert. Zusätzlich wurde das epigenetische Alter der Patienten, basierend auf dem DNA-Methylierungsmuster bei Keimzentrums-B-Zell-Lymphomen nach dem Modell von Steve Horvath, bestimmt. Die genauen Selektionskriterien, Darstellungsformen und das Vorgehen innerhalb der einzelnen Auswertungsschritte sind in den nachfolgenden Abschnitten dargestellt.

2.2.7.1 Qualitätskontrolle und Analyse der Rohdaten

1.) Initial-Analyse mittels *GenomeStudio* v2011.1 (*Methylation Modul* 1.9.0)

Die Rohdaten der 450K wurden in Form von .idat-Daten in das *Methylation Modul* der Software *GenomeStudio* (Illumina) geladen und der Analysetyp *Methylierungsanalyse* gewählt. Die Normalisierung mittels *GenomeStudio* erfolgte nur gegen interne Arraykontrollen, die ebenfalls auf dem 450K implementiert sind. Eine Hintergrundnormalisierung der AVG-Beta-Werte wurde durch *GenomeStudio* nicht durchgeführt, sondern erfolgte zu einem späteren Zeitpunkt mit dem Statistikprogramm R unter Verwendung des Pakets *RnBeads*. Die *GenomeStudio* Software generierte folgende Parameter, die in der weiteren 450K-Analyse verwendet wurden:

- AVG-Beta-Werte: Methylierungswerte, die zur Qualitätskontrolle mittels *Korrelationsmatrix* und zur weiteren Fragestellungs-bezogenen Methylierungsanalyse von Keimzentrums-B-Zell-Lymphomen verwendet werden.
- Detection P-value (DetPval): Wert zum Vergleich der gemessenen Fluoreszenzintensität der analysierten Loci und internen Arraykontrollen der Firma Illumina auf dem 450K. Aus dem ermittelten DetPval wurde eine Liste von allen analysierten Loci erstellt bei denen die gemessene Fluoreszenzintensität höher als der Hintergrund ($\text{DetPval} \leq 0,01$) war. Dadurch wurden alle Loci mit einer Falsch-Positiv-Rate von über 1% aus der weiteren 450K-Analyse ausgeschlossen. Auf diese Liste von CpG-Loci wurde u.a. zur Ermittlung der *Loci detection rate* (LDR) zurückgegriffen.

Zu allen Loci wurden zusätzlich folgende Informationen exportiert: *TargetID*, *Chromosom*, *Mapinfo*, *UCSC Refgenname*, *UCSC Genegroup* und *Relation to UCSC CpG Island*. Diese Angaben gehören zu einem Informationsset, das Illumina in der Software für jeden analysierten CpG-Loci mitlieferte und sich auf das Referenzgenom hg19 bezieht.

2.) Qualitätskontrolle mittels *Loci detection rate (LDR)*

In der vorliegenden Arbeit wurde zur Qualitätskontrolle der hybridisierten DNA-Proben u.a. die Bestimmung der *Loci detection rate (LDR)* angewandt. Die Bestimmung der *LDR* basierte auf allen analysierten Loci mit $\text{DetPval} \leq 0,01$ und ist als das Verhältnis der tatsächlich analysierten Loci zu der Gesamtanzahl der auf dem Array abgedeckten Loci ($N=485577$). Als Einschlusskriterium für die weitere 450K-Methylierungsanalysen wurde ein willkürlicher Grenzwert mit einer minimalen *LDR* von 99% gewählt.

3.) Qualitätskontrolle mittels *Korrelationsmatrizen*

Neben der *LDR* wurden *Korrelationsmatrizen* zur Qualitätskontrolle der hybridisierten DNA-Proben erstellt. Mit Hilfe der *Korrelationsmatrizen* wurde analysiert, ob sich die AVG-Beta-Werte einer Proben-Hybridisierung von den Hybridisierungen des gesamten zu analysierenden Datensatzes unterscheiden und ob ein linearer Zusammenhang zwischen den einzelnen Proben-Hybridisierungen vorliegt. Innerhalb der Qualitätskontrolle wurden zwei verschiedene *Korrelationsmatrixanalysen* durchgeführt, die in den nachfolgenden kurz beschrieben sind.

- *SNP-Korrelationsmatrix*

Zuerst wurden die Rohdaten (.idat-Dateien) aller mit "rs" bezeichneten Loci (Detail siehe Abschnitt 1.8.3 in der Einleitung) und die dazugehörigen AVG-Beta-Werte selektiert. Die rs-Loci wurden anschließend zur Erstellung der *SNP-Korrelationsmatrix* verwendet, um potenzielle Probenverwechslungen im analysierten Datensatz aufzudecken. Die Korrelationsanalyse der rs-Loci wurde in dem Statistikprogramm R nach Pearson mittels dem Befehl cor^{236} durchgeführt. Hierbei wurden die AVG-Beta-Werte der rs-Loci aller analysierten Proben miteinander verglichen. Die ermittelten Korrelationskoeffizienten der Hybridisierungen eines Datensatzes wurden zur grafischen Darstellung in Excel eingeladen und in einem asymmetrischen Farbcode bedingt formatiert (0=blau, ab 0,95=gelb, 1=rot). Diese Darstellungsform ermöglichte eine schnelle Unterscheidung zwischen Proben mit einem hohen Korrelationskoeffizienten (Nahe dem Wert 1, rot) und geringen Korrelationskoeffizienten (deutlich kleiner als der Wert 1, gelb bis blau).

Für die analysierten Tumor-Proben sollten die berechneten Korrelationskoeffizienten für die rs-Loci geringe Werte aufweisen (deutlich kleiner als der Wert 1), da keine biologischen oder technischen Replikate analysiert wurden. Für die FACS-sortierten Zell-Populationen vom gleichen Individuum werden hohe Korrelationskoeffizienten (nahe dem Wert 1) erwartet. Hierbei sollten die Hybridisierungen, die vom selben Individuum stammen, hohe Korrelationswerte von $\geq 0,99$ zueinander aufweisen. Alle analysierten DNA-Proben eines Individuums mit Korrelationswerte $< 0,99$ zueinander wurden aus den weiteren Analysen ausgeschlossen. Wurden Korrelationswerte von $\geq 0,99$ bei den Tumor- oder Lymphadenitisproben ermittelt, wurden diese Proben ausgeschlossen, da eine Probenverwechslung nicht ausgeschlossen werden konnte.

- *Gesamt-Korrelationsmatrix*

Die durchgeführten Arbeitsschritte zur Erstellung der *Gesamt-Korrelationsmatrix* wurden analog zur beschriebenen Erstellung der *SNP-Korrelationsmatrix* im vorherigen Abschnitt vorgenommen. Einziger Unterschied war, dass für die *Gesamt-Korrelationsmatrix* alle nicht-rs-Loci mit den dazugehörigen AVG-Beta-Werten verwendet wurden. Normalkontrollen zeigen untereinander standardmäßig ein homogenes DNA-Methylierungsmuster, sodass hohe Korrelationskoeffizienten (nahe dem Wert 1) in der *Gesamt-Korrelationsmatrix* erwartet werden. Hingegen ist das DNA-Methylierungsmuster von Tumor-Proben im Normalfall sehr heterogen, weswegen für diese Gruppe ein geringer Korrelationskoeffizient zu erwarten ist (deutlich kleiner als der Wert 1).

4.) Hintergrundkorrektur und Normalisierung mittels *RnBeads*

Im Anschluss an die Initialanalyse und Qualitätskontrolle der Rohdaten wurden die 450K-Methylierungsdaten mit dem Statistikprogramm R und dem Paket *RnBeads* ausgewertet. Es wurde eine Normalisierung des Hintergrunds und der Signalintensitäten des gesamten zu analysierenden Datensatzes vorgenommen, um die technische Varianz sowie potenzielle Batch-Effekte zwischen den Hybridisierungen zu minimieren²³⁷. Die Rohdaten (.idat-Dateien) der zu analysierenden Proben wurden hierfür in das R-Programm importiert und zuerst eine Normalisierung der Hintergrundfluoreszenz des Arrays durch eine Hintergrundsubtraktion mittels NOOB-Methode (engl. Normal-exponential using out-of-band probes) durchgeführt²³⁸. Anschließend wurde die speziell zur Analyse von 450K-Methylierungsdaten entwickelte Normalisierungsmethode SWAN (engl. subset-quantile within array normalisation) angewandt, welche die beobachteten Unterschiede in der AVG-Beta-Verteilung zwischen Infinium I und II Assay reduziert^{206,239,240}. Die Infinium Assay-Typen I und II des 450K wurden bereits in der Einleitung genauer beschreiben und Details sind dort im Abschnitt 1.7.3 nachzulesen. Die normalisierten AVG-Beta-Werte wurden abschließend in Form einer .txt-Datei exportiert und für die weitere Analyse genutzt.

2.2.7.2 Methylierungsanalyse von Keimzentrums-B-Zell-Lymphomen

1.) Selektion von Loci für die DNA-Methylierungsanalysen

Zur Selektion der Loci wurden die mittels RnBeads normalisierten AVG-Beta-Werte verwendet und die zuvor generierte .txt-Datei (Details siehe Abschnitt 2.2.7.1 „Hintergrundkorrektur und Normalisierung mittels *RnBeads*“) in Excel eingeladen. Anschließend wurden die AVG-Beta-Werte aller Loci (siehe Abbildung 22 in Abschnitt 2.2.7) der aus dem zu analysierenden Datensatz entfernt. Die Methylierungs-Filter-Liste mit den zu selektierenden Loci wurde von Frau Dr. Jana Gutwein im Rahmen ihrer Doktorarbeit am Institut für Humangenetik zusammengefasst und mir freundlicherweise zur Verfügung gestellt. Diese Liste beinhaltete u.A. die auf Geschlechtschromosomen lokalisierten Loci (Chr. X: N=11.232 und Chr. Y: N= 416), um geschlechtsspezifische Unterschiede im Methylierungsmuster in den weiteren Analysen auszuschließen. Zusätzlich wurden durch die Methylierungs-Filter-Liste alle SNPs (N= 65) und Sonden (N= 17043), die in unmittelbarer Nähe (bis zu 3 bp) ihrer 5'- Sondensequenz einen bekannten SNP mit einer MAF von mind. 5% beinhalteten, entfernt. Weiterhin wurden nur Loci in die finale 450K-Analyse eingeschlossen, deren zuvor die einen $\text{DetPval} \leq 0,01$ aufwiesen (detailliert beschrieben in Abschnitt 2.2.7.1). Die gefilterte Loci-Liste inkl. der entsprechenden AVG-Beta-Werte wurde abschließend als .txt-Datei gespeichert und für die finale 450K-Analyse verwendet.

2.) Identifizierung differentiell methylierter Loci mittels der Software *Qlucore*

Die Identifizierung differentiell methylierter Loci in senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom gegenüber adulten Patienten mit einem Keimzentrums-B-Zell-Lymphom wurde mit der Software *Qlucore Omics Explorer* (Qlucore) durchgeführt. Hierzu wurde vorab im Programm Microsoft Office Excel eine Annotationsdatei von allen Loci der zuvor gefilterten Loci-Liste erstellt (siehe Abschnitt 2.2.7.2 „Selektion von Loci für die DNA-Methylierungsanalysen“). Diese Annotationsdatei enthielt wichtige Informationen zu jeder einzelnen Probe (z.B. Alter, Geschlecht Diagnose) sowie die zu jeder Probe zugehörigen normalisierten AVG-Beta-Werte und wurde anschließend als .gedata-Datei gespeichert und in den *Qlucore Omics Explorer* eingeladen. In der Software *Qlucore Omics* wurde der Datensatz unter Verwendung eines Multigruppenvergleichs (ANOVA) und der linearen Regression analysiert. Die Software *Qlucore Omics* erlaubt es die verschiedenen in der Annotationsdatei hinterlegten Patienten-spezifischen Informationen in einem Gruppenvergleich gegen die AVG-Beta-Werte der DNA-Proben zu vergleichen. Mittels fließender Einstellung wurden die *False discovery rate (FDR)* und der P-Wert innerhalb der Analysen verändert. Zur Beurteilung der Signifikanz der differentiell methylierten Loci wurde die eingestellte *FDR* betrachtet, da diese gleichzeitig für multiples Testen mit der Korrektur nach Benjamin-Hochberg korrigiert²⁴¹. In der vorliegenden Arbeit wurden differentiell methylierte Loci als signifikant erachtet wenn eine $\text{FDR} \leq 0,050$ vorlag.

Alle durchgeführten Hauptkomponentenanalysen (engl. Principal component analysis, PCA) und hierarchischen Clusteranalysen wurden mit der Einstellung *Mean* $\rightarrow 0$ erstellt. Zusätzlich wurde mittels der Software *Qlucore Omics* in allen Analysen der Hybridisierungstag ausgeschlossen, um mögliche technische Artefakte zwischen den einzelnen Proben auszuschließen.

3.) Identifizierung des epigenetischen Alters nach Horvath

Zur altersassoziierten Analyse des DNA-Methylierungsmusters wurde das entsprechende epigenetische Alter der einzelnen Keimzentrums-B-Zell-Lymphome und der Normalkontrollen nach einem von Steve Horvath entwickelten Modell verwendet. Dazu werden die AVG-Beta-Werte von 353 CpGs, die Steve Horvath aus den 450K-Methylierungsdaten als altersassoziiert differenziell methyliert identifiziert hat, in die Modell-Rechnung einbezogen²². In dieser Arbeit wurde die gefilterte Loci-Liste nach der *RnBeads*-Analyse (siehe Abschnitt 2.2.7.2 „Selektion von Loci für die DNA-Methylierungsanalysen“) genutzt, um das epigenetische Alter der Keimzentrums-B-Zell-Lymphome und der Normalkontrollen mittels der frei zugänglichen Webseite zu berechnen (<https://dnamage.genetics.ucla.edu/>). Anschließend wurde das so bestimmte epigenetische Alter der einzelnen Proben in einer Excel-Datei gespeichert. Aus dem biologischen und dem epigenetischem (Horvath)-Alter wurde daraufhin eine Korrelation im Statistikprogramm R ermittelt und grafisch in Form von Punkt-, Linien- und Säulendiagrammen dargestellt.

2.2.8 Vergleiche der Hochdurchsatz-Sequenzierungs- und DNA-Methylierungs-Daten

In einem Vergleich der Hochdurchsatz-Sequenzierungs- und DNA-Methylierungs-Daten sollten Gene identifiziert werden, die sowohl das Mutationsprofil als auch das DNA-Methylierungsmuster in sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom (insbesondere DLBCL) beeinflussen. Dazu wurden zum einen Gene identifiziert, die in Patienten mit einem hohen biologischen als auch epigenetischen Alter Aberrationen aufweisen (siehe Abschnitt 2.2.8.1) und zum anderen wurden Gene ermittelt, die ggf. das epigenetische Alter in Keimzentrums-B-Zell-Lymphomen beeinflussen können (siehe Abschnitt 2.2.8.2). Das Vorgehen zum Vergleichen der Hochdurchsatz-Sequenzierungs- und Methylierungs-Daten und die dabei durchgeführten Arbeitsschritte werden in den nachfolgenden Abschnitten genauer beschrieben.

2.2.8.1 Rekurrent mutierte Gene in Patienten mit einem hohen biologischen und epigenetischen Alter

Im Rahmen der altersassoziierten Analysen in Keimzentrums-B-Zell-Lymphomen sollten rekurrent mutierte Gene identifiziert werden, die in Patienten mit einem hohen biologischen als auch epigenetischen Alter verändert sind. Hierzu wurden die in senilen Patienten mit einem DLBCL zuvor als rekurrent mutiert identifizierten Gene verwendet (siehe Abschnitt 2.2.5.5 „Altersassoziierte Analysen von genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen“, Abbildung Boxplot) und den analysierten Patienten das entsprechende epigenetischen Alter nach Horvath zugeordnet (siehe Abschnitt 2.2.7.2 „Identifizierung des epigenetischen Alters nach Horvath“). Anschließend wurde eine grafische Darstellung in Form eines Boxplots erstellt, worin das epigenetische Alter der Patienten im Bezug auf das rekurrent mutierte Gen betrachtet wurde.

Zudem wurde ein Vergleich der jeweils 20 rekurrent mutierten Gene mit dem höchsten medianen biologischen und epigenetischen Alter vorgenommen und mittels der Online-Software (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) ein Venn-Diagramm erstellt. Anhand dessen wurden die Gene identifiziert, die Aberrationen in DLBCL-Patienten mit einem hohen biologischen als auch epigenetischen Alter aufwiesen. Diese Gene wurden in einer Konsensus-Liste zusammengefasst und ein funktionelles Netzwerk mittels der Online-Software String (<http://string-db.org>) ermittelt.

2.2.8.2 Gene die das epigenetische Alter von Keimzentrums-B-Zell-Lymphomen beeinflussen

Weiterhin wurden in den altersassoziierten Analysen in Keimzentrums-B-Zell-Lymphomen Gene identifiziert, die das epigenetische Alter von Keimzentrums-B-Zell-Lymphomen beeinflussen (sog. „epigenetic modifiers“). Hierzu wurde die Differenz aus dem biologischen Alter (siehe Abschnitt 2.2.5.5 „Altersassoziierte Analysen von genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen“, Abbildung Boxplot) und dem epigenetischen Alter (siehe im Abschnitt 2.2.8.1 „Gene mit einem hohen biologischen und epigenetischen Alter in Keimzentrums-B-Zell-Lymphomen“, Abbildung Boxplot) bestimmt und im Folgenden als Delta-Alter bezeichnet. Anschließend wurde eine grafische Darstellung in Form eines Boxplots erstellt, worin das Delta-Alter des Patienten in Bezug auf das rekurrent mutierte Gen betrachtet wurde.

Für jedes rekurrent mutierte Gen (N= 50) wurde aus dem Delta-Alter der betroffenen Patienten der Median bestimmt. Anschließend wurde aus dem Median der einzelnen Gene der Gesamtmedian des Delta-Alter der Patienten (Gesamtmedian= 19 Jahre) und die Standardabweichung berechnet (σ = etwa 10).

Danach wurden die rekurrent mutierten Gene hinsichtlich des Gesamtmedian +/- einer halben Standardabweichung (Median 19 Jahre +/- 5 Jahre Standardabweichung) in drei Gruppen eingeteilt:

- Gruppe 1: Gene mit einem hohen Delta-Alter von ≥ 24 Jahre
- Gruppe 2: Gene mit einem mittleren Delta-Alter von 23,5-14,5 Jahre
- Gruppe 3: Gene mit einem geringen Delta-Alter von ≤ 14 Jahre

Anhand dessen wurden Gen-Gruppen identifiziert, die das epigenetische Alter in DLBCL beeinflussen und entweder zu einem deutlich höheren epigenetischen als biologischen Alter (Gruppe 1) oder zu einem deutlich geringeren epigenetischen als biologischen Alter (Gruppe 3) führen. Für diese Gen-Gruppe 1 bzw. 3 wurde jeweils ein funktionelles Netzwerk mittels der Online-Software String (<http://string-db.org>) berechnet.

2.2.9 Statistische Auswertung

In der vorliegenden Arbeit wurden im Rahmen der altersassoziierten Analysen bei Keimzentrums-B-Zell-Lymphomen verschiedene statistische Verfahren eingesetzt, um die erhobenen und bereits existierenden molekulargenetischen Daten auszuwerten. Dabei wurden drei verschiedene Tests mittels des Statistikprogramms R Version 3.0.1 angewandt und der P-Wert bestimmt. Die einzelnen Verfahren sind im nachfolgenden Kapitel genauer beschrieben.

2.2.9.1 Logistische Regression

Bei der logistischen Regression wird die Korrelation von Abhängigkeiten zwischen zwei oder mehr Merkmalen eines Kollektivs anhand von Messwerten analysiert. Mit Hilfe der Regressionsanalyse wird dann ein Korrelationskoeffizient ermittelt, der die Form der Abhängigkeit zwischen diesen Merkmalen beurteilt²⁴²⁻²⁴⁴. Dieser funktionale Zusammenhang stellt im einfachsten Fall eine lineare Regression dar²⁴⁵. In der vorliegenden Arbeit wurde die logistische Regression zur altersassoziierten Analyse von Keimzentrums-B-Zell-Lymphomen nach Klapper et al., (2012) verwendet. Dabei wurde das Patientenalter bei Erstdiagnose gegenüber genetischer Aberrationen von Keimzentrums-B-Zell-Lymphomen untersucht und der entsprechende P-Wert berechnet.

Im Detail wurden in dieser Arbeit folgende Parameter im Zusammenhang mit dem Patientenalter betrachtet:

- Ergebnisse der IHC-Analysen: Genexpression in den analysierten Proteinen BCL2, BCL6, MUM1/IRF4, Ki67 und CD10
- Ergebnisse der FISH-Analysen: Chromosomale Imbalancen unter Beteiligung der Gene *MYC*, *BCL2*, *BCL6* und *IGH*
- Weitere analysierte Patienten-Charakteristika: Geschlecht, DLBCL-Subtyp, Hans- und GEP-Klassifikator
- Ergebnisse der Hochdurchsatz-Sequenzierung und das Kandidatengen CD58

2.2.9.2 Mann-Whitney U Test

Zur statistischen Bewertung der Hochdurchsatz-Sequenzierungs-Daten wurde u.a. der Mann-Whitney U Test durchgeführt. Hierbei handelt es sich um einen Rangsummentest (Zwei-Stichproben-Test) zwischen zwei unabhängigen Gruppen, wobei der Frage nachgegangen werden soll, ob sich diese zwei Gruppen in einem Merkmal unterscheiden²⁴⁶⁻²⁴⁸. Zur Anwendung des Mann-Whitney U Tests wurden die Patienten nach dem Ersterkrankungsalter in zwei Gruppen eingeteilt: <80 Jahre und ≥80 Jahre. Anschließend wurden die analysierten Charakteristika wie z.B. detektierte Mutationen in einen 0- (negativ/wt) und 1- (positiv/mut) Code umgeschrieben und den Patienten die jeweiligen Altersgruppe (<80 bzw. ≥80) zugeordnet. Es folgte die Berechnung des P-Wertes mittels Statistikprogramms R durch den Befehl `wilcox.test {Dateiname}`, wobei die 1- (positiv/mut) Patienten der Altersgruppe <80 Jahre mit den 1- (positiv/mut) Patienten der Altersgruppe ≥80 Jahre verglichen wurden, um Signifikanzen hinsichtlich altersassoziierter Aberrationen bei den analysierten Patienten zu identifizieren. Der Mann-Whitney U Test wurde in der vorliegenden Arbeit für folgende Analyse genutzt:

- Vergleich von Ergebnissen der Hochdurchsatz-Sequenzierung (Patienten mit einem Ersterkrankungsalter bis zu 80 Jahren und einer Mutation im rekurrent mutierten Gen vs. Patienten mit einem Ersterkrankungsalter von mind. 80 Jahren und einer Mutation im rekurrent mutierten Gen)
- Vergleich von Ergebnissen der hierarchischen Clusteranalyse der Assoziationen in DLBCL (Patienten aus dem Cluster 1 mit einer 1 (positiv/mut) im analysierten Charakteristika vs. Patienten aus dem Cluster 2 mit einer 1 (positiv/mut) im gleichen analysierten Charakteristika)

2.2.9.3 Kruskal-Wallis Test

Innerhalb der statistischen Auswertung wurde auch der Kruskal-Wallis Test, der eine Erweiterung des Wilcoxon Tests ist, angewandt. Bei dem Kruskal-Wallis Test handelt es sich um einen nichtparametrischen Test der die Signifikanz von mehr als zwei Stichproben bestimmt²⁴⁹⁻²⁵¹. Mit Hilfe des Kruskal-Wallis Tests war es möglich eine Signifikanzanalyse der gesamten Alterskohorte ohne Intervalle und Altersgrenzen durchzuführen. In den Analysen dieser Arbeit wurden alle Patienten mit dem gleichen Ersterkrankungsalter als eine Gruppe zusammengefasst und hinsichtlich verschiedener Marker verglichen. Der P-Wert wurde im Statistikprogramm R durch den Befehl `kruskal.test{Dateiname}` ermittelt und die Verteilung grafisch mittels Boxplot dargestellt. Der Kruskal-Wallis-Test wurde in der vorliegenden Arbeit für folgende Analysen genutzt:

- Vergleich von Ergebnissen der IHC- und FISH Analysen: Genexpression in den analysierten Proteinen BCL2, BCL6, MUM1/IRF4, Ki67 und CD10 sowie chromosomale Imbalancen unter Beteiligung der Gene *MYC*, *BCL2*, *BCL6* und *IGH* (erstens: Ersterkrankungsalter(-sgruppe) vs. 1 (positiv/mut) im analysierten IHC- bzw. FISH-Marker, zweitens: Ersterkrankungsalter(-sgruppe) vs. 0 (negativ/wt) im analysierten IHC- bzw. FISH-Marker)
- Weitere analysierte Patienten-Charakteristika: Geschlecht, DLBCL-Subtyp, Hans- und GEP-Klassifikator (erstens: Ersterkrankungsalter(-sgruppe) vs. 1 (positiv/mut) im analysierten Charakteristika, zweitens: Ersterkrankungsalter(-sgruppe) vs. 0 (negativ/wt) im analysierten Charakteristika)
- Kandidatengen CD58 (erstens: Ersterkrankungsalter(-sgruppe) vs. genetische Aberration im *CD58*-Gen, zweitens: Ersterkrankungsalter(-sgruppe) vs. keine genetische Aberration im *CD58*-Gen)

2.2.9.4 Clustering zur Assoziationsanalyse

Um eine Einschätzung über die Interaktionen der analysierten Marker bzw. Charakteristika in DLBCL zu bekommen, wurden die Ergebnisse der Basis-Charakterisierung, DLBCL-Subtypen (*Hans*- und *GEP-Klassifikator*), Geschlechter und das Mutationsprofil einzelner Gene (z.B. *MYD88*, *CD79B*) in einem Farbcode in Form einer Heatmap dargestellt. Hierzu wurden die Ergebnisse der analysierten Charakteristika in einen 0- (negativ/wt) und 1- (positiv/mut) Code umgeschrieben und eine hierarchische Clusteranalyse der DLBCL-Patienten mit dem Statistikprogramm R und dem Befehl `heatmap.2{Dateiname}` durchgeführt.

Um mögliche altersassoziierte Effekte zu erkennen, wurden die analysierten Patienten in die Altersgruppen: Gruppe 1 (0-39 Jahre), Gruppe 2 (40-59 Jahre), Gruppe 3 (60-79 Jahre) sowie Gruppe 4 (80-97 Jahre) eingeteilt und in einem Farbcode über der Heatmap dargestellt. Basierend auf der Clusteranalyse wurde die Gesamtkohorte in zwei Gruppen eingeteilt (Cluster 1 und 2) und für jedes analysierte Charakteristika der P-Wert mittels Mann-Whitney U Test (siehe Abschnitt 2.2.9.2) berechnet.

2.2.10 Sonstige Verfahren zur Visualisierung und Auswertung

In der vorliegenden Arbeit wurden verschiedene grafische Übersichten zur Visualisierung der analysierten Kohorte und Mutationen erstellt. Zusätzlich wurden Auswertungen zur Assoziation zwischen den einzelnen Markern bzw. Charakteristika und den identifizierten Mutationen in Keimzentrums-B-Zell-Lymphomen durchgeführt. Die grafischen Darstellungen sowie das Vorgehen bei den Assoziationsanalysen ist in den nachfolgenden Abschnitten kurz erläutert.

2.2.10.1 Scatterplot-Boxplot-Grafiken

Im Rahmen der statistischen Auswertung wurden verschiedene grafische Darstellungen der Kohorten mittels dem Statistikprogramm R erzeugt, um die Altersverteilung der Patienten für einzelne Marker besser zu visualisieren (z.B. Geschlecht, Hans- und GEP-Klassifikator). Hierzu wurden Kombinationen aus Boxplots und Scatterplots erstellt, wobei jeder analysierte Patient bzw. Marker/Charakteristika durch einen Punkt dargestellt wurde. Die Boxplots geben das untere (25%) und obere (75%) Quartil der Altersverteilung der untersuchten Kohorten an, wobei der Median durch eine horizontale schwarze Linie gekennzeichnet ist. Die Antennen stellen das Minimum und Maximum der Altersverteilung dar und Ausreißer sind in Form von Kreisen in der Grafik angezeigt. Ein Beispiel für eine derartige Grafik ist in Abbildung 26 im Abschnitt 3.2.1 zu sehen.

2.2.10.2 MutationMapper zur Darstellung von Mutationen im Gen-Kontext

Zur Betrachtung der mittels Hochdurchsatz-Sequenzierung identifizierten genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden die detektierten Mutationen der Gene *TP53*, *IRF4*, *PRDM1*, *KMT2D* und *CD58* genauer analysiert. Hierzu wurden die Mutationen in das webbasierte Tool MutationMapper (<http://www.cbioportal.org/mutation-mapper.jsp>) eingeladen und eine grafische Übersicht erstellt. In der Grafik war die genomische Position der Aberration im analysierten Gen ersichtlich, wodurch mögliche Hotspots im analysierten Gen identifiziert wurden.

Weiterhin war zu erkennen, ob die Mutationen in einer bekannten Domäne oder über die komplette kodierende Region verteilt vorlagen. Zusätzlich wurden die verschiedenen Mutationstypen in einem unterschiedlichen Farbcode dargestellt (rot: trunkierende Mutation, grün: nicht synonyme Mutation, schwarz: In-Frameshift Deletion, violett: mehrere Mutationstypen wurden an dieser genomischen Position detektiert).

2.2.10.3 Stringanalyse und Venn-Diagramme

In dieser Arbeit wurden Mutationsanalysen von Keimzentrums-B-Zell-Lymphome durchgeführt und verschiedenen Genlisten erstellt. Abschließend wurden diese Genlisten miteinander verglichen, um gemeinsame und spezifisch rekurrent mutierte Gene für die jeweils analysierten Charakteristika bzw. Marker zu identifizieren. Dazu wurde unter Verwendung des webbasierten Tools Venny 2.0.2 (<http://bioinfogp.cnb.csic.es/tools/venny/>) ein Venn-Diagramm erstellt. Um eine grobe Einschätzung über die Assoziationen der gemeinsamen oder gruppenspezifisch mutierten Gene zu bekommen, wurde von einzelnen Gen-Gruppen Pathway-Analysen mittels String 10 durchgeführt. Hierzu wurden die Genlisten unter <http://string-db.org/> hochgeladen und die Assoziationen zwischen den Genen als Bilddatei (.png) gespeichert. Die Auswertung mittels Venn-Diagramm und Pathway-Analyse wurde in der vorliegenden Arbeit zum Vergleich der rekurrent mutierten Gene in den Mutationsanalysen der DLBCL nach dem biologischen und epigenetischen Alter (siehe Abschnitt 3.2.2.4) und dem Delta-Alter (3.2.2.) verwendet. Zusätzlich wurden diese Arbeitsschritte zum abschließenden Vergleich der genetischen Aberrationen in DLBCL und FL genutzt (3.2.1.3).

3 Ergebnisse

Das Ziel dieser Arbeit war die Identifizierung altersassoziierter genetischer und epigenetischer Aberrationen bei reifzelligen Keimzentrums-B-Zell-Lymphomen. Hierzu wurde gezielt eine Kohorte von senilen Patienten (mind. 80 Jahre) mit einem diffus großzelligem B-Zell-Lymphom (DLBCL) oder follikulärem Lymphom (FL) zusammengestellt und diese mittels Fluoreszenz-in-situ-Hybridisierung (FISH), Immunhistochemie (IHC), Hochdurchsatz-Sequenzierung und Methylierungsarrays analysiert. Die erhobenen Daten wurden anschließend in einer altersassozierten Analyse mit bereits vorhandenen Daten aus klinischen und Forschungsstudien hinsichtlich genetischer und epigenetischer Veränderungen, die insbesondere im sehr hohen Alter auftreten, verglichen. Für die Analyse des Mutationsprofils der Lymphom-Proben mittels Hochdurchsatz-Sequenzierung war es notwendig diese Methode am Institut für Humangenetik zu optimieren, was ein zusätzliches Ziel und wesentlicher Bestandteil der vorliegenden Arbeit war.

Zur Optimierung der Hochdurchsatz-Sequenzierung wurde eine Austestung unterschiedlicher Anreicherungsverfahren von verschiedenen Panel/Assay durchgeführt. Diese Panel/Assay wurden an Tumor- und Normalkontroll-Proben zu verschiedenen Fragestellungen angewandt und entsprechend optimiert. Die Sequenzierung der Libraries erfolgte mit dem MiSeq Gerät der Firma Illumina.

Der nachfolgende Ergebnisteil gliedert sich somit in zwei Haupt-Komplexe:

- 1) **Methodik:** Ergebnisse der Austestung und Optimierung verschiedener methodischer Verfahren zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung (Abschnitt 3.1)
- 2) **Keimzentrums-B-Zell-Lymphome:** Identifizierung von altersassozierten genetischen und epigenetische Aberrationen bei sehr alten Patienten (Abschnitt 3.2)

3.1 Austestung und Optimierung verschiedener methodischer Verfahren zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung

Zur Optimierung der Hochdurchsatz-Sequenzierung wurden verschiedene Panel bzw. Assays zur Probenanreicherung getestet und die jeweiligen Vor- und Nachteile der einzelnen Panel bzw. Assays bestimmt. Insgesamt wurden im Rahmen dieser Arbeit vier unterschiedliche Panel/Assays verwendet, die sowohl die Verfahren der PCR- als auch Hybridisierungs-basierten Anreicherungen abdecken.

- PCR-basierte Verfahren
 - a) Myeloid-Panel (*TruSight Amplicon Panel*, Illumina)
 - b) *ID3*-Assay (*Two-Step-PCR-Assay*, *Custom Assay* am Institut für Humangenetik erstellt)
- Hybridisierungs-basierte Verfahren
 - c) Cancer-Panel (*TruSight Rapid Capture Panel*, Illumina)
 - d) Lymphom-Panel (*SureSelectXT Panel*, *Custom Panel* der Firma Agilent)

Als Erstes wurde ein kommerziell erhältliches PCR-basiertes Myeloid-Panel der Firma Illumina ausgewählt (2.2.4.1). Nachfolgend wurde der Fokus auf Custom Designs gerichtet und zunächst ein kleiner Assay zur Abdeckung von einem Gen aufgesetzt (*ID3*-Assay). Dazu wurde nach dem *Two-Step-Verfahren* ein Assay zur Tiefensequenzierung von Exon 1 des *ID3*-Gens erstellt und optimiert (2.2.4.1). Nachdem zwei PCR-basierte Verfahren zur Probenanreicherung optimiert waren, wurde nun die Hybridisierungs-basierte Anreicherung unter Verwendung eines kommerziell erhältlichen Cancer-Panels der Firma Illumina getestet (2.2.4.2). Anschließend wurde nach einem Verfahren zur Probenvorbereitung der Keimzentrums-B-Zell-Lymphome gesucht. Basierend auf der Grundlage, dass eine möglichst große Anzahl an ausgewählte Target-Regionen analysiert werden sollte, wurde ein Hybridisierungs-basiertes Custom Panel ausgewählt. Zum damaligen Zeitpunkt war das Design von Hybridisierungs-basierten Custom Panels nur mit der Firma Agilent möglich, weshalb die Auswahl auf Agilent und nicht Illumina fiel. Dieses Custom Panel enthält ausgewählte Target-Regionen, die in Lymphomen rekurrent mutiert sind und wird in der vorliegenden Arbeit für die altersassoziierten Analysen von genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen verwendet (Lymphom-Panel, 2.2.4.2). In den folgenden Abschnitten werden zuerst die Ergebnisse der Optimierung der neuen methodischen Ansätze erläutert und anschließend eine kurze Beschreibung für die einzelnen Panel/Assay vorgenommen.

Im Rahmen der Optimierung der Probenvorbereitung wurden verschiedene Parameter z.B. Input, Ausgangsmaterial oder Normalisierung variiert, um möglichst viele Sequenzen mit bestmöglicher Qualität zu generieren (siehe Tabelle 17). Zum besseren Verständnis werden nachfolgend die wesentlichen Parameter kurz erläutert, die zur Bewertung der einzelnen Panel/Assay genutzt wurden.

Als Input wird die eingesetzte Menge an gDNA zur Probenvorbereitung bezeichnet, wobei der Input für die einzelnen Panel/Assay variierte. Als Output wird in dieser Arbeit die Menge an generierten Sequenzdaten bezeichnet, wobei diese abhängig von den verwendeten Sequenzier-Kartuschen und der entsprechenden Coverage ist und zwischen 0,5 und 7,5 Gb variiert. Für die kommerziell erhältlichen Panel wurden 50 ng Input verwendet, wohingegen bei den Custom Panels 200 ng Input je Probe benötigt wurden. Zusätzlich wurde das Ausgangsmaterial für die Probenvorbereitung der einzelnen Panel/Assay getestet. Die extrahierte DNA aus Frischmaterial wurde mit allen vier analysierten Panel/Assay durchgeführt. Die Anreicherung von FFPE-Material wurde nur mit den Custom Panels vorgenommen, da die getesteten kommerziellen Panel nach Herstellerangaben nicht FFPE-geeignet waren (siehe Tabelle 17).

Innerhalb der Optimierung der einzelnen Panel/Assay wurden verschiedene Sequenzier-Kit-Komponenten verändert, die einen Einfluss auf die Parameter haben. Hierbei ist zu berücksichtigen, dass jedes Sequenzier-Kit aus drei Komponenten besteht (Flow cell, Kartusche, Puffer), deren verwendete Größe und Version abhängig von dem zu sequenzierenden Panel/Assay sind (Details zu den SequenzierKit-Komponenten und einzelnen Fachbegriffen siehe unter Abschnitt 2.2.5). Im Rahmen der vorliegenden Arbeit wurde der Output an Sequenz-Daten und die Cluster Density in der Sequenzierung u.a. durch die Veränderung der Sequenzier-Kit-Komponenten variiert, indem die Sequenzier-Reagenzien von v2 auf v3 bzw. die Flow cell von Standard auf Nano gewechselt wurden (siehe Tabelle 17). Zudem ist die Normalisierung der Library ein wichtiger Optimierungs-Parameter, worin die Library auf die Konzentration eingestellt wird, die anschließend auf die Flow cell zur Sequenzierung geladen wird. Innerhalb dieser Arbeit wurde das Standard-Protokoll von Illumina (min. Input: 2 nM) verwendet und zusätzlich ein *Small-Scale-Protokoll* etabliert, dass es ermöglichte auch bei Libraries mit einer Konzentration von unter 2 nM eine höhere Cluster Density und mehr Output in der Sequenzierung zu erzielen (siehe Tabelle 17).

Zur Beurteilung der Genauigkeit der Sequenz-Daten wurde für alle analysierten Panel/Assay die Parameter Sensitivität und Spezifität bestimmt. Sensitivität steht dabei für richtig Positiv-Rate (Mutationen wurden mittels Sanger- und Hochdurchsatz-Sequenzierung identifiziert). Im Gegensatz dazu steht die Spezifität für die richtig Negativ-Rate, d.h. mit beiden Techniken wurden übereinstimmend keine Mutationen identifiziert.

Weiterhin wurde der Parameter Reproduzierbarkeit betrachtet, der aus dem Vergleich der mittels Hochdurchsatz-Sequenzierung analysierten Duplikate berechnet wurde. Hierzu wurden einzelne Proben zweifach angereichert und sequenziert. Anschließend wurde anhand der identifizierten Mutationsallelfrequenzen (MAF) das Bestimmungsmaß (R^2) zu einander bestimmt und alle Mutationen mit R^2 größer als 0,9 wurden als reproduzierbar bezeichnet.

Insgesamt wurden zur Optimierung der Hochdurchsatz-Sequenzierung 206 Proben unter Anwendung von vier unterschiedlichen Panel bzw. Assays analysiert. In der Tabelle 17 sind die wichtigsten Parameter und Ergebnisse zusammengefasst, die innerhalb der Optimierung der Anreicherungsverfahren für die einzelnen Panel bzw. Assays erzielt wurden. Eine detaillierte Auflistung der einzelnen Panel bzw. Assays und den dazugehörigen Parametern ist im Anhang in der Tabelle A2 zusammengestellt.

Tabelle 17: Übersicht zu den wichtigsten variierten Parametern bei der Optimierung verschiedener Panel/ Assays für die Hochdurchsatz-Sequenzierung

| Anreicherungsverfahren Panel/Assay | PCR-basiert | | Hybridisierungs-basiert | |
|-------------------------------------------|-----------------------------|----------------------------------|------------------------------------------|-------------------------------------------|
| | Myeloid-Panel | ID3-Assay | Cancer-Panel | Lymphom-Panel |
| Kit zur Probenvorbereitung | TruSight Amplicon Panel | Two-Step-PCR-Assay | TruSight Rapid Capture | SureSelectXT Panel |
| Hersteller | kommerziell Illumina | Custom inhouse | kommerziell Illumina | Custom Agilent |
| Ausgangsmaterial | Frischmaterial | Frisch- und FFPE-Material | Frischmaterial | Frisch- und FFPE-Material |
| Input gDNA [ng] | 50 | 200 | 50 | 200 |
| Polymerase | Illumina (TruSight Kit) | Life Technologies (AccuPrime) | Illumina (TruSight Kit) | Agilent Technologies (Herkulase II) |
| Normalisierung | Standard | Standard | Standard (N= 5-12) Small-Scale (N≤ 4) | Small-Scale |
| Probenanzahl pro MiSeq-Lauf | max. 12 (v3) max. 4 (v2) | max. 96 | max. 12 max. 4 (Nano) | max. 16 |
| Veränderte Sequenzier-Kit- Komponenten | v2-(N≤4) | - | Nano (N≤ 4) | - |
| Ø-Cluster Density [K/mm ²] | 1093 (v3)/ 249 (v2) | 598 (Nano) | 874/ 647 (Nano) | 393/ 1260 (Small-Scale) |
| Ø-Output [Gb] | 5,5 (v3)/ 1,5 (v2) | 0,4 (Nano) | 4,4/ 1,0 (Nano) | 1,4/ 3,2 (Small-Scale) |
| Sensitivität | 100% (17/17) | 100% (8/8) | 100% (13/13) | 100% (11/11) |
| Spezifität | NA | 100% (18/18) | 100% (9/9) | NA |
| Reproduzierbarkeit | NA | 100% (17/17) | NA | NA |

Abkürzungen: gDNA: genomische DNA, Standard: Protokoll zur Normalisierung von Libraries mit einer Konzentration ≥ 2 nM (Protokoll siehe Abschnitt 2.2.4.3). Small-Scale: Protokoll zur Normalisierung von Libraries mit einer Konzentration < 2 nM ((Protokoll siehe Abschnitt 2.2.4.3). N: Anzahl an analysierten Proben. v2: Sequenzier-Reagenzien Version v2. v3: Sequenzier-Reagenzien Version v3. Nano: Nano Flow cell. NA: Wert konnte für dieses Panel nicht berechnet werden.- : Die Sequenzier-Reagenzien wurden nicht verändert. Sequenzierungen die nicht explizit mit "Nano" oder "v3" gekennzeichnet wurden, sind mit den Standard-SequenzierKit-Komponenten (Standard Flow cell und v2) durchgeführt worden.

3.1.1 Myeloid-Panel (*TruSight Amplicon Panel*)

Die Probenvorbereitung mittels Myeloid-Panel ist ein PCR-basiertes Verfahren der Firma Illumina zur Anreicherung von 573 Amplikons in 54 Genen, die in myeloiden Neoplasien rekurrent mutiert sind. Im Rahmen dieser Arbeit wurden insgesamt 39 T-PLL-Patienten analysiert. Zur Analyse wurden jeweils zwölf Patienten unter Verwendung von v3-Reagenzien oder vier Patienten unter Verwendung von v2-Reagenzien gepoolt. Eine Beschreibung zu den unterschiedlichen Sequenzier-Reagenzien und Komponenten wurde zuvor unter dem Abschnitt 2.2.5 gegeben und ist dort nachzulesen. Insgesamt wurden fünf Sequenzierläufe (Paired-End, 2x 150 bp) mit dem Myeloid-Panel durchgeführt, wobei Lauf 4 nicht auswertbar war (Tabelle 18, rot markiert). In diesem MiSeq-Lauf wurde eine sehr hohen Cluster Density von über 1500 K/mm² erzielt. Die berechneten QC-Parameter zeigten, dass der Q30-Wert bei etwa 50% lag und nur etwa 30% der Reads den Illumina-spezifischen Filter passierten. Es ist zu berücksichtigen, dass innerhalb der Probenvorbereitung mit dem Myeloid-Panel keine PhiX-Kontrolle verwendet wurde und somit der QC-Parameter Fehlerrate nicht vorlag. Da die ermittelten QC-Parameter des Laufs 4 nicht den erfordernten *inhouse*-QC-Werten (siehe Abschnitt 2.2.5.1) entsprachen, wurde die zuvor erstellte Library erneut normalisiert und sequenziert (siehe Abschnitt 2.2.4.1). In der Auswertung des wiederholten Laufs (siehe Tabelle 18, Lauf 5) wurde eine deutlich geringere Cluster Density von 635 K/mm² erzielt. Die berechneten QC-Parameter lagen bei einem Q30-Wert von 96% und 96% der Reads passierten den Illumina-spezifischen Filter. Damit erfüllten insgesamt vier von fünf MiSeq-Läufen die *inhouse*-QC-Werte (Tabelle 18), wobei die Daten dieser vier Läufe anschließend in die detaillierte Sequenz-Auswertung eingingen.

Tabelle 18: Qualitätswerte der Hochdurchsatz-Sequenzierungen mit dem Myeloid-Panel

| Lauf | Reagenz | CD [K/mm ²] | Total reads | PF reads | Cluster PF [%] | PF Reads Identified [%] | Q30 [%] | OTR [%] | OTC [x] | Output [Gb] |
|------|-----------------|-------------------------|-------------|----------|----------------|-------------------------|---------|---------|---------|-------------|
| 1 | v3 (600 Zyklen) | 1323 | 30283852 | 26129148 | 86 | 97 | 93,4 | 99,5 | 6904 | 8,2 |
| 2 | v3 (600 Zyklen) | 827 | 19743330 | 18280892 | 93 | 97 | 96,0 | 99,5 | 4883 | 5,7 |
| 3 | v2 (300 Zyklen) | 249 | 4918393 | 4728420 | 96 | 97 | 98,0 | 99,5 | 4358 | 1,5 |
| 4 | v3 (600 Zyklen) | 1587 | 14855589 | 6048511 | 32 | 94 | 54,0 | 97,4 | 791 | 3,6 |
| 5 | v3 (600 Zyklen) | 635 | 15057785 | 13804377 | 92 | 96 | 96,0 | 98,1 | 3605 | 4,3 |

Abkürzungen: Reagenz: Verwendete Sequenzier-Reagenz von Illumina. CD: (engl. Cluster Density). PF: (engl. Passing Filter) Beschreibt die Sequenzen, die den Illumina-spezifischen Qualitätsfilter passieren und in die weitere Auswertung eingehen. Q30: Qualitätsparameter von Illumina. OTR: Durchschnittliche On-Target-Rate. OTC: Durchschnittliche On-Target-Coverage. Die Sequenzierung des Myeloid-Panels wurde für 12 Patienten mit v3 (Lauf 1, 2, 4 und 5) und für 4 Patienten mit v2 Reagenzien (Lauf 3) durchgeführt. Zur besseren Übersicht wurde der qualitativ schlechte MiSeq-Lauf 4 rot umrandet und die schlechten QC-Parametern in rot eingefärbt. Die Library vom Lauf 4 wurde erneut normalisiert und sequenziert (Lauf 5). Die Qualitätswerte der Sequenzierparameter im Wiederholungslauf sind deutlich besser und entsprechen der internen Qualitätskontrolle.

Anschließend erfolgte die Auswertung der Sequenzierungs-Daten mit der Software *Variant Studio* von der Firma Illumina (siehe Abschnitt 2.2.5.3), wobei zuerst die durchschnittliche On-Target-Rate (\emptyset -OTR) berechnet wurde. Die \emptyset -OTR lag bei etwa 99%, wobei über 90% der analysierten Amplikons in allen Proben eine Coverage $\geq 50x$ zeigten. Nachfolgend wurde die durchschnittliche On-Target-Coverage (\emptyset -OTC) für jede Probe ermittelt. Für eine Probe lag die \emptyset -OTC bei 3x, d.h. größtenteils wurden für diese Probe keine Amplikons sequenziert. Diese geringe \emptyset -OTC ist vermutlich auf die schlechte DNA-Qualität dieser Probe zurückzuführen. Weiterhin wurden sieben Amplikons identifiziert, die in allen Proben eine \emptyset -OTC $< 10x$ aufwiesen (Dropouts: nicht abgedeckte Regionen), wobei für drei Amplikons die \emptyset -OTC bei 0x lag. Bei diesen identifizierten Dropouts handelt es sich um Amplikons in den Genen *CDKN2A*, *CUX1*, *DNMT3A*, *SMC3* und *TET2*, die in repetitiven Sequenzen, Homopolymerbereichen und GC-reichen Regionen lagen. In den auswertbaren Target-Regionen des Myeloid-Panels konnten in 38 auswertbaren T-PLL-Patienten insgesamt 17 potentiell proteinverändernde Mutationen identifiziert und mittels Sanger-Sequenzierung validiert werden. Damit lag die ermittelte Sensitivität des Panels bei 100%. Die Spezifität konnte für das Myeloid-Panel nicht berechnet werden, da keine Proben mit bereits bekannten Mutationen sequenziert wurden. Diese Ergebnisse sind Teil des Manuskripts Lopez et al., „Genes encoding members of the *JAK-STAT* pathway or epigenetic regulators are recurrently mutated in T-cell prolymphocytic leukaemia“, das 2016 beim *British Journal of Haematology* veröffentlicht wurde.

3.1.2 *ID3*-Assay (*Two-Step-PCR-Assay*)

Bei dem *ID3*-Assay handelt es sich um einen PCR-basierten Custom Assay zur Anreicherung von vier Amplikons, die es ermöglichen das Mutationsprofil von Exon 1 und die zugehörigen Splicing Regionen sowie zwei Kontroll-Regionen mittels Hochdurchsatz-Sequenzierung zu analysieren (siehe Abschnitt 2.2.4.1). Im Rahmen der vorliegenden Arbeit wurden insgesamt 41 Proben mit dem *ID3*-Assay analysiert. Der Assay sollte neben Frischmaterial auch für Analysen von FFPE-Material geeignet sein. Daraufhin wurde zuerst ein Polymerase-Test an drei Zelllinien (Frischmaterial) und drei FFPE-Testproben vorgenommen.

Im Rahmen der Optimierung des *ID3*-Assays wurden vier verschiedene Polymerase Systeme getestet: Q5 Polymerase (NEB), Phusion Polymerase (NEB), AccuPrime Polymerase (Invitrogen) sowie FastStart Polymerase (Roche). Die Ergebnisse der Qualitätskontrolle der PCR-Produkte mittels Bioanalyzer und Qubit zeigten, dass unter den angewandten PCR-Bedingungen (siehe Abschnitt 2.2.4.1) keine Amplifikation mit der FastStart Polymerase stattfand (0/6). Unter Verwendung der Q5 Polymerase konnten 3/6 Proben sowie mit der Phusion Polymerase 5/6 Proben amplifiziert werden. Nur mittels der AccuPrime Polymerase wurden alle sechs Proben erfolgreich amplifiziert (Tabelle 19, grün markiert unter der Kategorie PCR).

Tabelle 19: Schematische Übersicht der Ergebnisse der Probenvorbereitung des ID3-Assays mit unterschiedlichen Polymerase Systemen an extrahierter DNA aus Frisch- und FFPE-Material

| Probe | Amplikon | Phusion (New England Biolabs) | | | Q5 (New England Biolabs) | | | FastStart (Roche) | | | AccuPrime (Life Technologies) | | |
|------------------------|----------|----------------------------------|------|-------------|-----------------------------|------|-------------|----------------------|-----|------------|----------------------------------|------|-------------|
| | | PCR | NGS | OTR [%] | PCR | NGS | OTR [%] | PCR | NGS | OTR [%] | PCR | NGS | OTR [%] |
| FFPE-1* | KR1 | grün | grün | 50 | grün | grün | 25 | rot | | 0 | grün | rot | 75 |
| | KR2 | grün | grün | | grün | grün | | rot | | | | | |
| | Ex1a | grün | grün | | grün | grün | | | | | | | |
| | Ex1b | grün | rot | | grün | rot | | | | | | | |
| FFPE-2* | KR1 | rot | | 50 | rot | | 25 | rot | | 0 | grün | grün | 75 |
| | KR2 | rot | | | rot | | | | | | | | |
| | Ex1a | rot | | | rot | | | | | | | | |
| | Ex1b | rot | | | rot | | | | | | | | |
| FFPE-3* | KR1 | grün | grün | 50 | grün | rot | 25 | rot | | 0 | grün | grün | 75 |
| | KR2 | grün | grün | | grün | rot | | | | | | | |
| | Ex1a | grün | grün | | grün | rot | | | | | | | |
| | Ex1b | grün | rot | | grün | rot | | | | | | | |
| Namalwa [#] | KR1 | grün | grün | 75 | grün | grün | 25 | rot | | 0 | grün | grün | 100 |
| | KR2 | grün | grün | | grün | grün | | rot | | | | | |
| | Ex1a | grün | grün | | grün | grün | | | | | | | |
| | Ex1b | grün | rot | | grün | rot | | | | | | | |
| BL-70 [#] | KR1 | grün | grün | 75 | rot | | 25 | rot | | 0 | grün | grün | 100 |
| | KR2 | grün | grün | | rot | | | | | | | | |
| | Ex1a | grün | grün | | grün | grün | | | | | | | |
| | Ex1b | grün | rot | | grün | rot | | | | | | | |
| Blue-1 [#] | KR1 | grün | grün | 75 | rot | | 25 | rot | | 0 | grün | grün | 100 |
| | KR2 | grün | grün | | rot | | | | | | | | |
| | Ex1a | grün | grün | | grün | grün | | | | | | | |
| | Ex1b | grün | rot | | grün | rot | | | | | | | |
| Gesamt: OTR [%] | | | | 62,5 | | | 25,0 | | | 0,0 | | | 87,5 |

Abkürzungen und Markierungen: PCR: (engl. polymerase chain reaction) Polymerasekettenreaktion. NGS: (engl. Next Generation Sequencing) Hochdurchsatz-Sequenzierung. KR: Kontrollregion (Region von der bekannt ist, dass keine ID3-Mutationen vorliegt). Ex: Exon. OTR: Durchschnittliche On-Target-Rate. rot markiert: es wurde kein PCR-Produkt mittels Qubit oder Bioanalyzer gemessen bzw. es konnte kein Amplifikat für das Amplikon sequenziert werden (Coverage=0x). grün markiert: es wurde ein PCR-Produkt mittels Qubit oder Bioanalyzer gemessen bzw. es konnte ein Amplifikat für das Amplikon sequenziert werden, d.h. die Coverage lag bei mind. 1x. *: FFPE-Material. #: Frischmaterial.

Nachfolgend wurden 18 B-Zell-Lymphom-Zelllinien mit einem bereits bekannten Mutationsstatus aus Richter et al. mittels ID3-Assay und unter Verwendung der AccuPrime Polymerase angereichert (17/18 in Duplikaten). Abschließend wurden sowohl die Proben der Amplifikate aus dem Polymerase-Test (N= 6) als auch die Amplifikate der B-Zell-Lymphom-Zelllinien (N= 35) in einem MiSeq-Lauf mit einer Nano Flow cell und v2-Reagenzien mit dem MiSeq-Gerät (Paired-End, 2x250 bp) sequenziert.

Im Gegensatz zum Myeloid-Panel wurde für den *ID3*-Assay in der Probenvorbereitung die PhiX-Kontrolle mitgeführt und eine darauf basierende Fehlerrate von 1% für den MiSeq-Lauf ermittelt. Zudem lag der Cluster Passing Filter bei 85% und der Q30-Wert bei 90%. Somit wurden insgesamt alle *inhouse*-QC-Werte (siehe Abschnitt 2.2.5.1) in der Hochdurchsatz-Sequenzierung des *ID3*-Assays erfüllt (siehe Tabelle 20) und alle Sequenz-Daten dieses MiSeq-Laufs in der weiteren Auswertung genauer betrachtet.

Tabelle 20: Qualitätswerte der Hochdurchsatz-Sequenzierung mit dem *ID3*-Assay

| Lauf | CD [K/mm ²] | Total reads | PF reads | Cluster PF [%] | PF Reads Identified [%] | Q30 [%] | Fehlerrate [%] | OTR [%] | OTC [x] | Output [Gb] |
|------|----------------------------|----------------|-------------|-------------------|----------------------------|------------|-------------------|------------|------------|----------------|
| 1 | 598 | 980178 | 835128 | 85 | 77 | 90 | 1 | 88 | 4179 | 0,4 |

Abkürzungen: CD: (engl. Cluster Density). PF: (engl. Passing Filter) Beschreibt die Sequenzen, die den Illumina-spezifischen Qualitätsfilter passieren und in die weitere Auswertung eingehen. Q30: Qualitätsparameter von Illumina. Fehlerrate: Fehlerrate des MiSeq-Lauf basierend auf der mitgeführten PhiX-Kontrolle. OTR: Durchschnittliche On-Target-Rate. OTC: Durchschnittliche On-Target-Coverage.

Die Auswertung der Sequenzierungs-Daten erfolgte mit der Software *Sequence Pilot* im Modul *SeqNext* nach den im Methodenteil beschriebenen Auswertungsparametern (siehe Abschnitt 2.2.5.4). Um eine abschließende Beurteilung der getesteten Polymerase Systeme unter den angewandten Bedingungen zu geben, wurde zuerst für jede Library, die im Rahmen des Polymerase-Tests analysiert wurde, die Ø-OTR ermittelt (siehe Tabelle 19, N= 6). Die berechnete Ø-OTR betrug für die AccuPrime Polymerase: 87,5%, Phusion Polymerase: 62,5%, Q5 Polymerase: 25% und Roche Polymerase: 0%. Die Ergebnisse der Sequenzierungs-Daten zeigten, dass unter Verwendung der Q5 und Phusion Polymerase von keiner Probe alle Target-Regionen vollständig sequenziert wurden (0/6). Hingegen wurden unter Einsatz der AccuPrime Polymerase bei 5/6 Proben alle Target-Regionen vollständig sequenziert (Tabelle 19, grün markiert unter der Kategorie NGS). Die Ergebnisse der Ø-OTR zeigten, dass in der Probenvorbereitung mittels AccuPrime Polymerase die meisten Amplifikate sequenziert wurden. Die Berechnung der weiteren Parameter des Assays (z.B. Sensitivität) erfolgte anhand der Sequenz-Daten der B-Zell-Lymphom-Zelllinien mit bereits bekanntem Mutationsstatus²²⁰ (mutiert in Exon 1: N= 8; mutiert in Exon 2: N= 1; unmutiert: N=9).

17/18 Zelllinien wurden als Duplikate präpariert und in einem Lauf sequenziert, um die Reproduzierbarkeit des *ID3*-Assay zu beurteilen. Da sich die bekannte Mutation der Zelllinie Namalwa teilweise über Exon 2 streckt und demnach nicht vollständig durch den *ID3*-Assay abgedeckt wird, wurde die Namalwa-Probe als technische Kontrolle mitgeführt, aber keine Duplikate präpariert.

Die Ergebnisse der Hochdurchsatz-Sequenzierung zeigten, dass die Sensitivität des Assays bei 100% lag d.h. alle Mutationen der analysierten Zelllinien im Exon 1 des *ID3*-Gens konnten verifiziert werden (8/8). Zusätzliche potenziell proteinverändernde Mutationen wurden unter den angewandten Auswertungsparametern nicht identifiziert (18/18, Spezifität: 100%). Weiterhin zeigten die analysierten Duplikate ähnlich detektierte MAFs, was die Reproduzierbarkeit der Ergebnisse bestätigte (R^2 : 0,996, Abbildung 23).

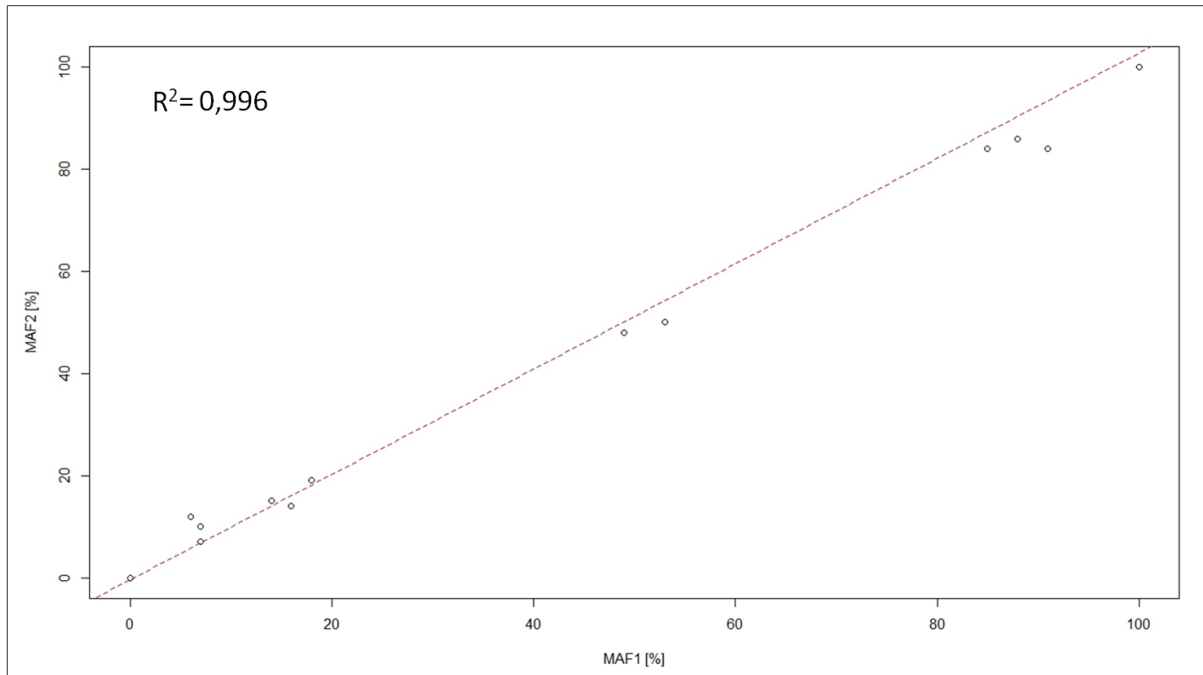


Abbildung 23: Grafische Darstellung zur Überprüfung der Reproduzierbarkeit des *ID3*-Assays

Die Abbildung zeigt die Korrelation der MAFs für die identifizierten Mutationen der analysierten Zelllinien mittels *ID3*-Assay (N= 17). Das ermittelte Bestimmungsmaß von $R^2 = 0,996$ zeigte, dass die Ergebnisse des *ID3*-Assays in der Hochdurchsatz-Sequenzierung reproduzierbar waren.

Die Ergebnisse der Mutationsanalyse des *ID3*-Assays der B-Zell-Lymphom-Zelllinien aus der Hochdurchsatz-Sequenzierung im Vergleich zu den Ergebnissen der Sanger-Sequenzierung sind in der nachfolgenden Tabelle 21 zusammengestellt.

Tabelle 21: Ergebnisse der Verifizierung der Hochdurchsatz-Sequenzierungs-Daten mit dem *ID3*-Assay im Vergleich zu den bekannten Mutationen in den B-Zell-Lymphom-Zelllinien aus den Sanger-Sequenzierungs-Daten von Richter et al., 2012.

| Zelllinie | Entität | <i>ID3</i> -Mutation mittels Sanger-Sequenzierung aus Richter et al., 2012 | <i>ID3</i> -Two-Step-PCR-Assay | | | | | |
|------------|----------------|----------------------------------------------------------------------------|--------------------------------|-----|----|------------|-----|----|
| | | | Duplikat 1 | | | Duplikat 2 | | |
| | | | C | MAF | V | C | MAF | V |
| BL-2 | BL | c.300G>A | 2592 | 7 | ja | 3636 | 7 | ja |
| | | c.300+1G>C | | | | | | |
| BL-41 | BL | c.[202G>C]; [202G>C]; p.[Glu68fs*];[Glu68fs*] | 3575 | 100 | ja | 2706 | 100 | ja |
| BL-70 | BL | c.138-264del; p.Cys47Profs*32 | 8975 | 91 | ja | 653 | 84 | ja |
| | | c.300+1G>A | 8975 | 6 | ja | 653 | 12 | ja |
| Blue-1 | BL | c.236-240delACCTG; p.Asp79Alafs*13 | 2250 | 18 | ja | 2182 | 19 | ja |
| CA-46 | BL | c.160C>G; p.Leu54Val | 2610 | 14 | ja | 3033 | 15 | ja |
| | | c.190C>T; p.Leu64Phe | 2610 | 85 | ja | 3033 | 84 | ja |
| Daudi | BL | c.160C>G; p.Leu54Val | 1677 | 53 | ja | 6329 | 50 | ja |
| | | c.241C>T; p.Gln*81 | 1677 | 49 | ja | 6329 | 48 | ja |
| DG-75 | BL | wt | 2363 | wt | ja | 2092 | wt | ja |
| EB-1 | BL | wt | 2689 | wt | ja | 1503 | wt | ja |
| Namalwa | BL | c.220-360+66del; p.Ile74Valfs*26 | 1727 | - | - | | | |
| Raji | BL | wt | 2053 | wt | ja | 1713 | wt | ja |
| Ramos | BL | wt | 1662 | wt | ja | 2918 | wt | ja |
| U-698-M | BL | c.166C>T; p.Pro56Ser | 2944 | 88 | ja | 3097 | 86 | ja |
| | | c.233T>C; p.Leu78Pro | 2944 | 7 | ja | 3097 | 10 | ja |
| MC-116 | B-Zell-Lymphom | c.300G>A | 2427 | 16 | ja | 3575 | 14 | ja |
| | | c.300+1G>A | 2427 | 16 | ja | 3575 | 14 | ja |
| HT | DLBCL-GCB | wt | 1297 | wt | ja | 2934 | wt | ja |
| Karpas-422 | DLBCL | wt | 3213 | wt | ja | 3315 | wt | ja |
| SU-DHL-10 | DLBCL-GCB | wt | 3384 | wt | ja | 4501 | wt | ja |
| SU-DHL-5 | DLBCL-GCB | wt | 9776 | wt | ja | 1824 | wt | ja |
| SU-DHL-6 | DLBCL-GCB | wt | 5204 | wt | ja | 3435 | wt | ja |

Abkürzungen und Verweise: C: Coverage. MAF: Mutationsallelfrequenz in%. V: Verifizierung. *ID3*-Mutation: Nomenklatur nach HGVS unter Verwendung des humanen Referenzgenoms hg19. BL: Burkitt-Lymphom. DLBCL-GCB: Diffus großzelliges B-Zell-Lymphom von Keimzentrumsähnlichen B-Zellen. wt: Wildtyp (unmutierte Probe). Verifizierung: Bekannte Mutation aus Richter et al., 2012 konnte in der Lymphom-Zelllinie mittels *ID3*-Assay im Rahmen dieser Arbeit verifiziert werden. Duplikat 1/2: Für die gleiche Zelllinie wurden zwei Libraries unabhängig voneinander präpariert. -: Keine Angabe möglich, da keine Mutation detektiert wurde. Da die bereits bekannte Mutation der Zelllinie Namalwa nicht vollständig vom *ID3*-Assay abgedeckt wird, wurde die extrahiert DNA von Namalwa als technische Kontrolle mitgeführt, aber keine Duplikate generiert.

Zusammenfassend konnte in der vorliegenden Arbeit ein Hochdurchsatz-Assay zur Tiefensequenzierung von Exon 1 und zwei Kontroll-Regionen des *ID3*-Gens von extrahierter DNA aus Frisch- und FFPE-Material mit dem *Two-Step-PCR*-Verfahren unter Verwendung des AccuPrime Polymerase Systems optimiert werden.

3.1.3 Cancer-Panel (*TruSight Rapid Capture Panel*)

Die Probenvorbereitung mittels Cancer-Panel ist ein Hybridisierung-basiertes Verfahren der Firma Illumina zur Anreicherung von ~etwa 4000 Target-Regionen in 94 Genen, die in malignen Tumoren bekanntermaßen rekurrent mutiert sind. Im Rahmen dieser Arbeit wurden insgesamt 28 Patienten mit diesem Panel analysiert (siehe Abschnitte 2.1.1.4). Zur Sequenzanalyse wurden jeweils 12 Patienten unter Verwendung des Standard-Protokolls bzw. 4 Patienten unter Verwendung des im Rahmen dieser Arbeit optimierten *Small-Scale-Protokolls* (Durchführung siehe Abschnitt 2.2.4.3) gepoolt. Dieses *Small-Scale-Protokoll* ermöglicht die Sequenzierung von Libraries mit einer Konzentration unter 2 nM zu sequenzieren. Die mittels *Small-Scale-Protokoll* normalisierten Proben wurden zusätzlich mit einer Nano Flow cell sequenziert (Details zu den Sequenzier-Kit-Komponenten siehe Abschnitt 2.2.5).

Insgesamt wurden vier MiSeq-Läufe, zwei Läufe mit jeweils zwölf und zwei Läufe mit jeweils vier Patienten durchgeführt. Alle MiSeq-Läufe wurden unter der Verwendung von v2-Reagenzien durchgeführt (Paired-End, 2x150 bp). In der Probenvorbereitung mittels Cancer-Panel wurde die PhiX-Kontrolle mitgeführt und eine darauf basierende Fehlerrate von unter 0,5% für alle MiSeq-Läufe ermittelt. Zudem lagen der Cluster Passing Filter bei über 80% und der Q30-Wert bei über 90% für alle MiSeq-Läufe. Somit wurden insgesamt in allen vier MiSeq-Läufen die *inhouse*-QC-Werte erfüllt (Tabelle 22) und alle Sequenz-Daten dieser vier Läufe wurden in der weiteren Auswertung genauer betrachtet.

Tabelle 22: Qualitätswerte der Hochdurchsatz-Sequenzierungen mit dem Cancer-Panel

| Lauf | CD [K/mm ²] | Total reads | PF reads | Cluster PF [%] | PF Reads Identified [%] | Q30 [%] | Fehlerrate [%] | OTR [%] | OTC [x] | Output [Gb] |
|------|----------------------------|----------------|--------------|-------------------|----------------------------|------------|-------------------|------------|------------|----------------|
| 1 | 997 | 18308 356 | 1499048 9 | 82 | 96 | 92 | 0,45 | 100, 0 | 697 | 4,7 |
| 2 | 751 | 14104 455 | 1291706 4 | 92 | 95 | 94 | 0,37 | 99,9 | 391 | 4,1 |
| 3 | 886 | 47895 96 | 3875029 | 81 | 95 | 93 | 0,35 | 99,8 | 355 | 1,2 |
| 4 | 407 | 23103 54 | 2071952 | 90 | 95 | 95 | 0,29 | 99,8 | 188 | 0,7 |

Abkürzungen: CD: (engl. Cluster Density). PF: (engl. Passing Filter) Beschreibt die Sequenzen, die den Illumina-spezifischen Qualitätsfilter passieren und in die weitere Auswertung eingehen. Q30: Qualitätsparameter von Illumina. Fehlerrate: Fehlerrate des MiSeq-Lauf basierend auf der mitgeführten PhiX-Kontrolle. OTR: Durchschnittliche On-Target-Rate. OTC: Durchschnittliche On-Target-Coverage. Grau markiert: Zur besseren Übersicht und Abgrenzung zwischen den verwendeten Normalisierungs-Protokollen und den Sequenzier-Kit-Komponenten wurden alle MiSeq-Läufe unter Verwendung des *Small-Scale-Protokolls* und der Nano Flow cell grau eingefärbt. Die MiSeq-Läufe ohne Farbe (weiß) wurden mit dem Standard-Protokoll der Firma Illumina normalisiert und eine Standard Flow cell verwendet.

Nachfolgend wurde die Auswertung der Sequenzierungs-Daten mit der Software *Sequence Pilot* im Modul SeqNext durchgeführt (siehe Abschnitt 2.2.5.4). Die ermittelte Ø-OTR lag für alle vier MiSeq-Läufe bei über 99,8%. Die berechnete Ø-OTC lag für die mittels Standard-Protokoll normalisierten Proben bei 544x und für alle mittels *Small-Scale-Protokoll* normalisierten Proben bei 272x.

In den ausgewerteten Target-Regionen konnten bei 13 der 31 analysierten Proben jeweils eine potenziell proteinverändernde Mutationen (N=13) identifiziert werden. Bei 9/13 Mutationen handelte es sich um bereits bekannte Mutationen aus vorangegangenen Sanger-Sequenzierungen. 4/13 Mutationen wurden bei Indexpatienten identifiziert und anschließend mittels Sanger-Sequenzierung validiert. Somit wurden insgesamt 13/13 Mutationen der mittels Cancer-Panel analysierten Proben erfolgreich validiert und daraufhin als richtig positiv bewertet (Sensitivität: 100%). In den 9 Proben mit bereits bekannten Mutationen aus der Sanger-Sequenzierung wurden in den Hochdurchsatz-Sequenzierungs-Daten keine zusätzlichen potenziell proteinverändernde Mutationen unter den angewandten Auswertungsparametern identifiziert (9/9, Spezifität: 100%).

Zusammenfassend konnten in der vorliegenden Arbeit zwei Protokolle (Standard- und *Small-Scale-Protokoll*) zur Probenvorbereitung und anschließenden Hochdurchsatz-Sequenzierung mit dem Cancer-Panel der Firma Illumina etabliert werden.

3.1.4 Lymphom-Panel (*SureSelectXT Panel*)

Die Probenvorbereitung zur Anreicherung von über 13.000 Target-Regionen in 118 Genen, die in Lymphomen häufig mutierte vorliegen erfolgte mit einem Hybridisierung-basierten SureSelect XT Custom Panel der Firma Agilent Technologies (Details zum Design und der Durchführung siehe Abschnitt 2.2.4.2). Im Rahmen der vorliegenden Arbeit wurden 80 Patienten-Proben (19 Proben als Duplikate) mit dem Lymphom-Panel analysiert. Für einen MiSeq-Lauf wurden 12 bis 16 Proben zu einer Library gepoolt und insgesamt zehn MiSeq-Läufe bei einer Readlänge von 2x 100 bp unter der Verwendung von v2-Reagenzien durchgeführt.

Beim Lymphom-Panel erfolgte das Custom Design als *in-siliko*-Verfahren, sodass ein etabliertes Herstellerprotokoll für die Normalisierung und Sequenzierung der Library nicht vorlag. Daraufhin wurden die Libraries in den ersten Sequenzierungen nach dem Standard-Verfahren der Firma Illumina normalisiert und jeweils 8 pM Library sowie 5% PhiX-Library auf die Sequenzier-Kartusche geladen (Tabelle 23, Lauf 1-4). Nachdem der ermittelte Output für alle Sequenzierungen bei unter 2 Gb lag, wurde der Input auf 10 pM Library erhöht. Der berechnete Output lag daraufhin in der Sequenzierung bei etwa 2,6 Gb, was immer noch deutlich geringer war als der max. zu erzielende Wert (5,1 Gb). Da die Normalisierung der Library nach dem Standard-Protokoll von Illumina nicht mehr verändert werden konnte (10 pM max. Konzentration), wurde das zuvor optimierte *Small-Scale-Protokoll* zur Normalisierung der Lymphom-Panel-Proben getestet (optimiert für kleine Probenansätze im Rahmen des *TruSight Rapid Capture Panels*, siehe Abschnitt 2.2.4.2). Der Output konnte damit von ~2,6 Gb auf ~4,1 Gb erhöht werden. Daraufhin wurden alle weiteren Library-Normalisierungen mit dem *Small-Scale-Protokoll* durchgeführt (Lauf 6-10). Die Sequenzierungs-Ergebnisse zeigten, dass mit dem *Small-Scale-Protokoll* eine deutlich höhere Cluster Density (Standard: 390 K/mm², *Small-Scale*: 1260 K/mm²) sowie ein höherer Output an Sequenz-Daten (Standard: 1,6, *Small-Scale*: 3,5 Gb) erzielt wurden. Die PhiX-Kontrolle basierende Fehlerrate lag bei <0,5% für alle MiSeq-Läufe. Zudem lag für alle MiSeq-Läufe der Cluster Passing Filter bei >75% und der Q30-Wert bei > 90%. Insgesamt wurden in allen 10 MiSeq-Läufen die *inhouse-QC*-Werte erfüllt (Tabelle 23), sodass die Sequenz-Daten aller zehn Läufe in der weiteren Auswertung genauer betrachtet wurden.

Tabelle 23: Qualitätswerte der Hochdurchsatz-Sequenzierungen mit dem Lymphom-Panel

| Lauf | CD [K/mm ²] | Total reads | PF reads | Cluster PF [%] | PF Reads Identified [%] | Q30 [%] | Fehler- rate [%] | Input [pM] | PhiX [%] | Output [Gb] |
|------|----------------------------|-------------|----------|-------------------|----------------------------|------------|---------------------|---------------|-------------|----------------|
| 1 | 266 | 5490878 | 4950548 | 90 | 81 | 95 | 0,25 | 8 | 5 | 1,0 |
| 2 | 321 | 6411265 | 6061588 | 95 | 84 | 94 | 0,22 | 8 | 5 | 1,2 |
| 3 | 153 | 3527479 | 3056869 | 87 | 89 | 95 | 0,30 | 8 | 5 | 0,6 |
| 4 | 466 | 9170849 | 8627671 | 94 | 86 | 91 | 0,16 | 8 | 5 | 1,8 |
| 5 | 740 | 14227608 | 13038849 | 92 | 90 | 96 | 0,25 | 10 | 5 | 2,6 |
| 6 | 1374 | 24770202 | 20416526 | 82 | 90 | 92 | 0,30 | 25 | 1 | 4,1 |
| 7 | 984 | 18543052 | 16345373 | 88 | 88 | 93 | 0,27 | 25 | 1 | 3,3 |
| 8 | 1174 | 21304696 | 17683262 | 83 | 76 | 93 | 0,28 | 25 | 1 | 3,6 |
| 9 | 1170 | 21065604 | 17614046 | 84 | 90 | 93 | 0,28 | 25 | 1 | 3,6 |
| 10 | 1594 | 28677660 | 22249654 | 78 | 86 | 91 | 0,32 | 25 | 1 | 4,5 |

Abkürzungen: CD: (engl. Cluster Density). PF: (engl. Passing Filter) Beschreibt die Sequenzen, die den Illumina-spezifischen Qualitätsfilter passieren und in die weitere Auswertung eingehen. Q30: Qualitätsparameter von Illumina. Fehlerrate: Fehlerrate des MiSeq-Lauf basierend auf der mitgeführten PhiX-Kontrolle. PhiX: Kommerziell erhältliche Library des PhiX-Virus der Firma Illumina. Zur besseren Übersicht und Abgrenzung zwischen den verwendeten Normalisierungsprotokollen der DNA-Libraries wurden die MiSeq-Läufe unter Verwendung des Small-Scale-Protokolls grau eingefärbt. Die Libraries ohne Farbe (weiß) wurden mit dem Standard-Protokoll der Firma Illumina durchgeführt.

Nachfolgend wurden alle Hochdurchsatz-Sequenzierungs-Daten zur weiteren bioinformatischen Analyse an das Deutsche Krebsforschungszentrum übermittelt (Arbeitsgruppe Theoretische Bioinformatik von Herrn Dr. Matthias Schlesner). Für die Auswertung der Lymphom-Panel-Proben ist zu beachten, dass die Duplikate einer Probe am DKFZ zusammengeführt und im weiteren als eine Probe ausgewertet wurden. Zuerst wurde eine Qualitätskontrolle der Sequenzierungs-Daten mittels einer *DKFZ-inhouse-Pipeline* durchgeführt und verschiedene Qualitätsparameter wie z.B. Ø-OTR, Ø-OTC und Duplikat-Rate bestimmt (siehe detaillierte Tabelle A3 im Anhang). Die Ø-OTR für alle analysierten Proben mit dem Lymphom-Panel lag bei 44% (min: 23%; max: 68%). Anschließend wurde die Ø-OTC für jede analysierte Probe pro Gen berechnet und grafisch dargestellt (Abbildung 24). Die Ø-OTC über alle analysierten Proben ergab einen Wert von 35x (min: 5x, max: 106x). Basierend auf der min. Ø-OTC der WGS-Daten aus der Vergleichskohorte des ICGC MMML-Seq Projektes mit einer Coverage von mind. 30x wurde als Einschlusskriterium für die Mutationsanalysen jedes Patienten eine min. Ø-OTC von 30x über alle Target-Regionen des Lymphom-Panels gewählt. Insgesamt erfüllten die Sequenzierungs-Daten von 50 Patienten die Ø-OTC-Bedingungen. Für 30 Patienten wurden die Ø-OTC-Bedingungen nicht erzielt, weshalb diese Patienten aus der weiteren Analyse ausgeschlossen wurden.

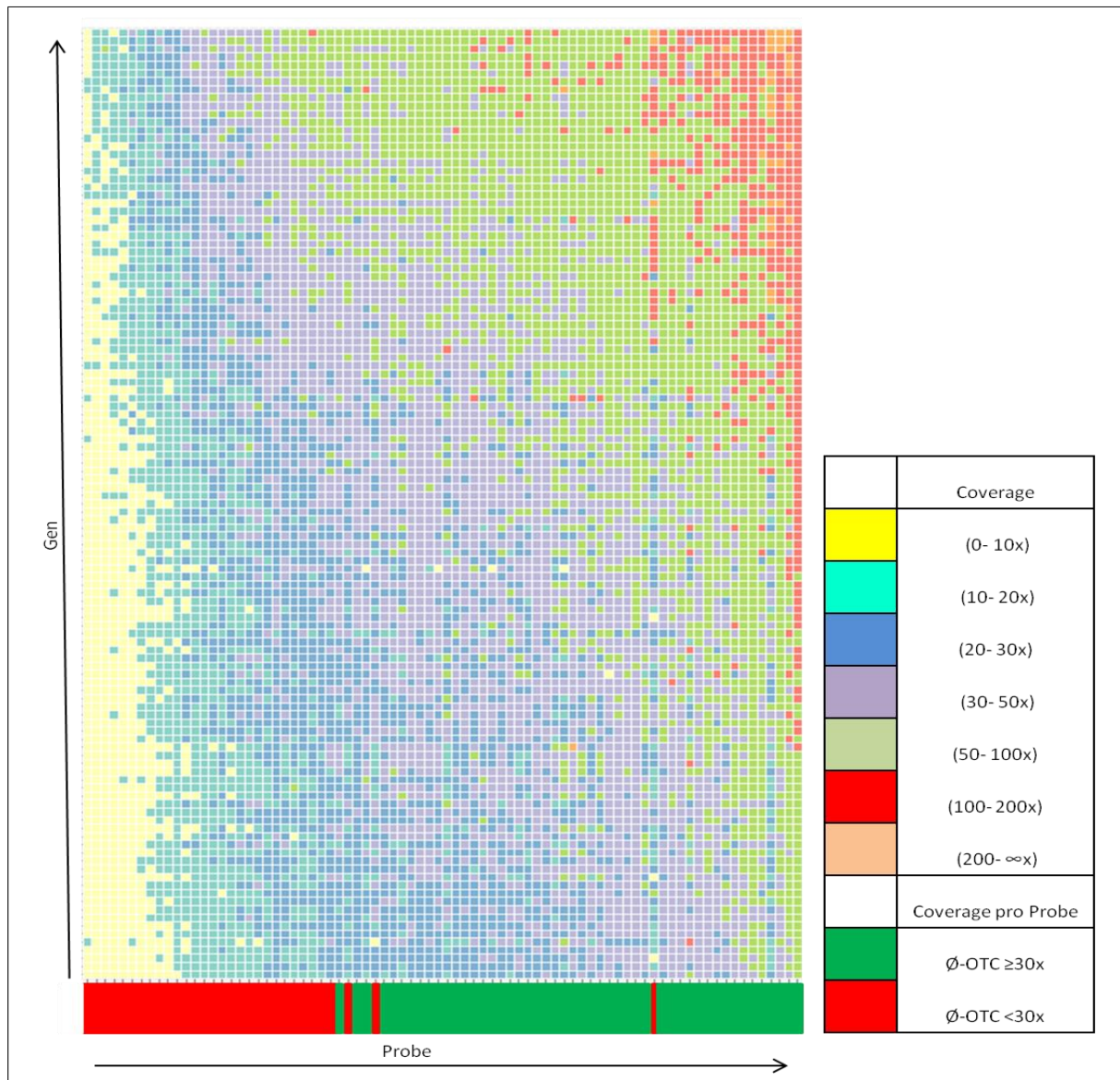


Abbildung 24: Schematische Darstellung der durchschnittliche On-Target-Coverage pro Gen

Die Coverage der 80 Lymphom-Panel-Proben wurde am Deutschen Krebsforschungszentrum in Heidelberg berechnet und grafisch in Form von einer Matrix dargestellt. In der Grafik stellt jede Spalte eine analysierte Probe und jede Zeile ein analysiertes Gen dar. Jedes farbige Rechteck innerhalb der Matrix spiegelt die $\bar{\varnothing}$ -OTC einer Probe über ein Gen wider. Als Einschlusskriterium für die weiteren Mutationsanalysen wurde eine durchschnittliche On-Target-Rate ($\bar{\varnothing}$ -OTC) von 30x gewählt. Insgesamt wurde für 50/80 Proben eine $\bar{\varnothing}$ -OTC $\geq 30x$ erzielt (grün markiert). Für die weiteren 30/80 analysierten Proben wurde eine $\bar{\varnothing}$ -OTC $< 30x$ ermittelt, weshalb diese Proben aus den weiteren Analysen ausgeschlossen wurden (rot markiert).

Zur Überprüfung der Sensitivität des Lymphom-Panels wurden 20 potentiell proteinverändernde Mutationen mit unterschiedlicher MAF (von etwa 0,15 bis 0,90) mittels Sanger-Sequenzierung analysiert (Tabelle 24). Für 9/20 Mutationen konnten keine spezifischen PCR-Produkte generiert werden, was am Ehesten mit einer schlechten DNA-Qualität der FFPE-Proben zu erklären ist. Alle weiteren 11 Mutationen konnten erfolgreich mittels Sanger-Sequenzierung validiert werden, woraufhin die ermittelte Sensitivität des Panels bei 100% lag. Die Spezifität konnte für das Lymphom-Panel nicht berechnet werden, da keine Proben mit bereits bekannten Mutationen sequenziert wurden.

Tabelle 24: Übersicht der Sanger-Sequenzierungs-Ergebnisse zur Validierung der mittels Lymphom-Panel detektierten Mutationen

| Patient | Gen | Transkript | Mutation | Funktion | Typ | MAF | Produktgröße [bp] | Validierung |
|---------|-----------------|-----------------|--------------------------------|---------------------|-------|------|-------------------|------------------|
| Jack-3 | <i>CREBBP</i> | NM-004380 | c.T4504A p.Trp1502Arg | nicht-synonym | SNV | 0,30 | 300 | ja |
| Jack-27 | <i>CREBBP</i> | NM-004380 | c.C4471A p.Gln1491Lys | nicht-synonym | SNV | 0,19 | 300 | ja |
| Jack-4 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,50 | 336 | ja |
| Jack-10 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,63 | 336 | ja |
| Jack-27 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,27 | 336 | ja |
| Jack-28 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,89 | 336 | ja |
| Jack-32 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,27 | 336 | ja |
| Jack-36 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,60 | 336 | ja |
| Jack-43 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,43 | 336 | ja |
| Jack-25 | <i>TNFRSF14</i> | ENST00000451778 | c.G200A p.Cys67Tyr | nicht-synonym | SNV | 0,59 | 250 | ja |
| Jack-54 | <i>TNFRSF14</i> | ENST00000451778 | c.C245T p.Thr82Ile | nicht-synonym | SNV | 0,77 | 250 | ja |
| Jack-28 | <i>CREBBP</i> | NM-004380 | c.G4232A p.Gly1411Glu | nicht-synonym | SNV | 0,25 | 306 | kein PCR-Produkt |
| Jack-34 | <i>CREBBP</i> | NM-004380 | c.3172-3179del p.Glu1058fs* | Frameshift Deletion | INDEL | NA | 329 | kein PCR-Produkt |
| Jack-51 | <i>CREBBP</i> | NM-004380 | c.4548-4549del p.His1517fs* | Frameshift Deletion | INDEL | NA | 300 | kein PCR-Produkt |
| Jack-57 | <i>CREBBP</i> | NM-004380 | c.G4337A p.Arg1446His | nicht-synonym | SNV | 0,31 | 315 | kein PCR-Produkt |
| Jack-19 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,76 | 336 | kein PCR-Produkt |
| Jack-37 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,81 | 336 | kein PCR-Produkt |
| Jack-48 | <i>MYD88</i> | NM-002468 | c.T794C p.Leu265Pro | Stopp-Mutation | SNV | 0,50 | 336 | kein PCR-Produkt |
| Jack-24 | <i>TNFRSF14</i> | ENST00000451778 | c.G178A p.Gly60Ser | nicht-synonym | SNV | 0,16 | 202 | kein PCR-Produkt |
| Jack-39 | <i>TNFRSF14</i> | ENST00000451778 | c.305delC p.Ala102fs* | Frameshift Deletion | INDEL | NA | 228 | kein PCR-Produkt |

Abkürzungen: NA: nicht verfügbar. Mutation: Nomenklatur nach HGVS unter Verwendung des humanen Referenzgenoms hg19. Funktion: Gibt den Mutationstyp an. SNV: (engl. single nucleotide variant). Punktmutation. Indel: Insertion oder Deletion. MAF: Mutationsfrequenz in der Hochdurchsatz-Sequenzierung. Produktgröße: Fragmentlänge des PCR-Produktes

Zusammenfassend wurde in dieser Arbeit an Frisch- und FFPE-Material ein Lymphom-Panel zur Hochdurchsatz-Sequenzierung von ausgewählten Lymphom-spezifischen Target-Regionen mit dem *SureSelectXT*-Verfahren der Firma Agilent Technologies erstellt und optimiert. Die Hochdurchsatz-Sequenzierungs-Daten der mittels Lymphom-Panel analysierten Tumor- (N= 41) und Kontroll-Proben (N= 9), die eine min. Coverage von 30x aufwiesen, wurden im Rahmen der vorliegenden Arbeit zur altersassoziierten Mutationsanalyse von Keimzentrums-B-Zell-Lymphomen verwendet. Im Zusammenhang des Ersterkrankungsalters der betroffenen Patienten sind die Ergebnisse der Mutationsanalyse der Lymphom-Panel-Proben in Kombination mit den WGS-Daten des ICGC MMML-Seq Projektes in dem nachfolgenden Kapitel 3.2. detaillierter beschrieben.

3.2 Altersassoziierte genetische und epigenetische Aberrationen bei reifzelligem Keimzentrums-B-Zell-Lymphomen

In der vorliegenden Arbeit sollten altersassoziierte Analysen von genetischen und epigenetischen Aberrationen bei Patienten mit einem Keimzentrums-B-Zell-Lymphom durchgeführt werden. Hierzu wurden Patienten mit einer Diagnose von diffus großzelligem B-Zell-Lymphom (DLBCL) oder follikulärem Lymphom (FL) nach WHO-Klassifikation 2008 mittels verschiedener molekulargenetischer Methoden analysiert. Die Betrachtung der Altersverteilung der Patienten mit DLBCL oder FL aus den klinischen Studien RICOVER und NHL-B sowie den Forschungsstudien MMML und ICGC MMML-Seq zeigte, dass für die Analysen der vorliegenden Arbeit molekulargenetische Daten von Patienten aus einem breiten Altersspektrum vorlagen, diese Daten aber überwiegend die Altersgruppe der Kinder und Erwachsenen bis zu einem Alter von 80 Jahren umfassten (Abbildung 25). Da der Fokus dieser Arbeit in der Analyse von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom (≥ 80 Jahre) lag, wurde gezielt eine neue Kohorte von sehr alten (senilen) Patienten mit einem Keimzentrums-B-Zell-Lymphom erstellt, die nachfolgend als *Jack* bezeichnet wird. Zudem wurden 28 Patienten aus dem Verbundprojekt MMML mit einem Ersterkrankungsalter von mind. 80 Jahren zu *MMML-Senil* zusammengefasst und ergänzende molekulargenetische Analysen zu den bereits vorhandenen Untersuchungen vorgenommen. Zudem wurden bereits existierende Daten von größtenteils jüngeren Patienten mit einem Keimzentrums-B-Zell-Lymphom (< 80 Jahre) aus verschiedenen *Vergleichsstudien* zu Vergleichszwecken verwendet. Im Rahmen dieser Arbeit wurden verschiedene altersassoziierte genetischen und epigenetischen Analysen an einer Gesamtkohorte von 1892 Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem Ersterkrankungsalter von 3 bis 97 Jahren durchgeführt. Dazu wurden alle senilen Patienten als Kohorte *DLBCL/FL-Senil* (N= 161) zusammengefasst und alle Patienten mit einem Ersterkrankungsalter unter 80 Jahren als Kohorte *DLBCL/FL-Adult* bezeichnet (N= 1731). Eine Übersicht der analysierten Patienten-Kohorten in Bezug auf die einzelnen Studien und das Ersterkrankungsalter ist in der nachfolgenden Abbildung 25 dargestellt.

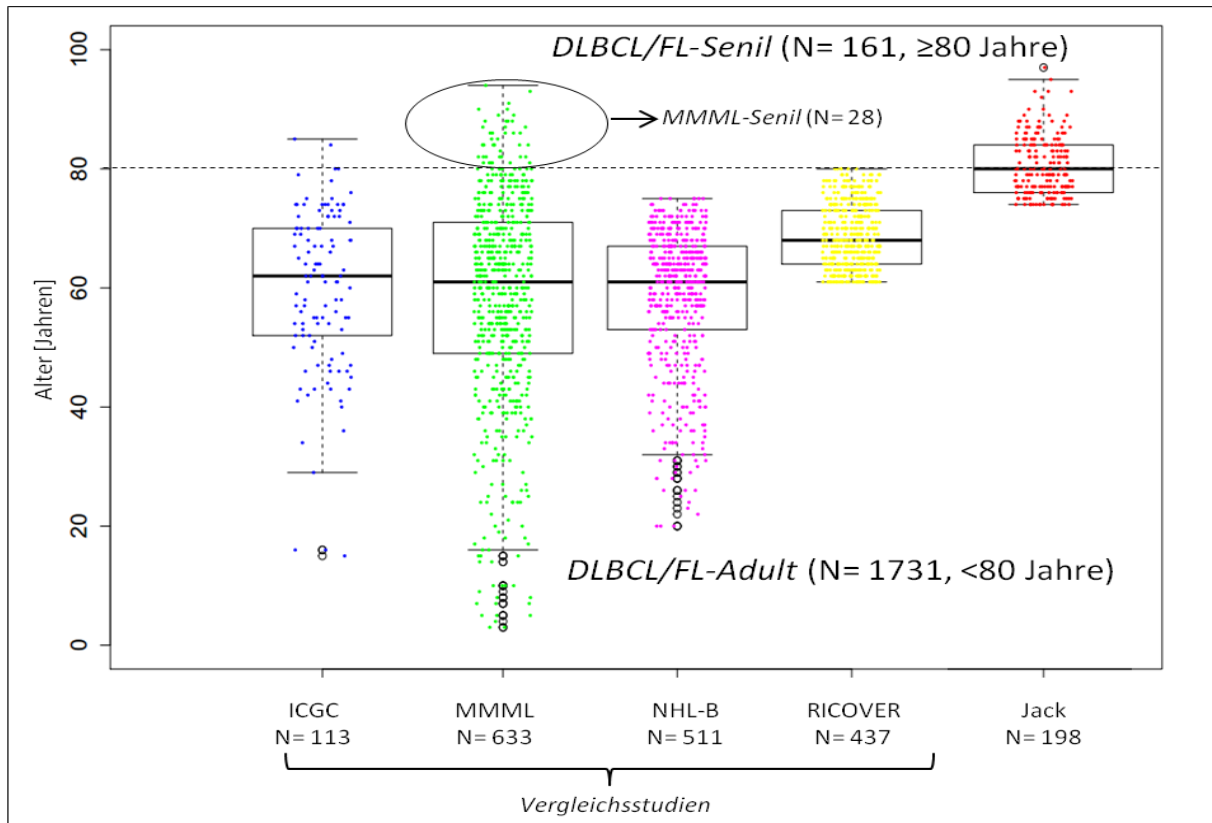


Abbildung 25: Schematische Darstellung der analysierten Patienten-Kohorten in Bezug auf die einzelnen Studien und das Ersterkrankungsalter. In den grafischen Darstellungen stellt jeder Punkt einen analysierten Patienten dar. Die Boxplots geben das untere (25%) und obere (75%) Quartil der Altersverteilung der untersuchten Kohorte an, wobei der Median durch eine horizontale, schwarze Linie in den Boxplots gekennzeichnet ist. Die Antennen stellen das Minimum und Maximum der Altersverteilung dar und Ausreißer sind in Form von Kreisen in der Grafik eingefügt. Im Rahmen der vorliegenden Arbeit wurden genetische und epigenetische Analysen von Keimzentrums-B-Zell-Lymphomen durchgeführt. Hierzu wurde eine Kohorte von sehr alten Patienten neu identifiziert und charakterisiert (*Jack*). Zusätzlich konnte die bereits extrahierte DNA von 28 Tumor-Proben von senilen Patienten aus dem MMML-Verbundprojekt analysiert werden (*MMML-Senil*). Zum Vergleich konnten bereits existierende molekulargenetische Daten von Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den klinischen (NHL-B, RICOVER) sowie Forschungsstudien (MMML, ICGC) genutzt werden. Das Gesamtkollektiv dieser vier Studien wird als *Vergleichsstudien* bezeichnet. Alle Patienten mit einem Ersterkrankungsalter von mind. 80 Jahren werden als Kohorte *DLBCL/FL-Senil* bezeichnet (N= 161) und in den altersassoziierten Analysen von genetischen und epigenetischen Aberrationen mit der Kohorte *DLBCL/FL-Adult* (<80 Jahre) verglichen. Abkürzungen: J: Jahre. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. RICOVER und NHL-B: Klinische Studie von Patienten mit einem Keimzentrums-B-Zell-Lymphom.

Insgesamt wurden alle 1892 Patienten im Rahmen dieser Arbeit oder innerhalb der verschiedenen Studien mittels Fluoreszenz-in-situ-Hybridisierung (FISH)- und Immunhistochemie (IHC) analysiert. Zusätzlich wurde von den Tumor-Proben einzelner Patienten das Mutations- und/oder das DNA-Methylierungsprofil analysiert oder die bereits existierenden Daten aus den Studien verwendet. Die neu erhobenen und bereits existierenden molekulargenetischen Daten aller Patienten wurden anschließend in einer altersassoziierten Analyse hinsichtlich genetischer und epigenetischer Aberrationen ausgewertet. Eine detaillierte Übersicht zu den durchgeführten Analysen in Bezug auf die analysierten Patienten-Kohorten ist im Anhang in der Tabelle A1 zu finden. Die Ergebnisse der einzelnen molekulargenetischen Analysen in Keimzentrums-B-Zell-Lymphomen sind in den nachfolgenden Kapiteln detaillierter beschrieben: genetische Aberrationen siehe unter Abschnitt 3.2.1 und epigenetischen Aberrationen siehe unter Abschnitt 3.2.2

3.2.1 Genetische Aberrationen bei reifzelligem Keimzentrums-B-Zell-Lymphomen

Im Rahmen dieser Arbeit zur Identifizierung von altersassoziierten genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden zuerst verschiedene Parameter wie das Geschlecht, Lymphom-Entität, Lymphom-Subtyp und Hans-/ GEP-Klassifikator (Details siehe unter Abschnitt 2.1.2.2) mit dem Ersterkrankungsalter von Patienten mit einem DLBCL oder FL verglichen. Anschließend wurden in Keimzentrums-B-Zell-Lymphomen die chromosomalen Imbalancen mittels Fluoreszenz-in-situ-Hybridisierung (FISH), die Genexpression mittels Immunhistochemie (IHC) und das Mutationsprofil von Lymphom-assoziierten Genen mittels Hochdurchsatz-Sequenzierung genauer betrachtet. Im Folgenden werden die Daten aus den FISH- und IHC-Analysen als Basis-Charakterisierung bezeichnet. Es erfolgte eine altersassoziierte Auswertung hinsichtlich genetischer Aberrationen in Keimzentrums-B-Zell-Lymphomen, worin alle o.g. Parameter und die Basis-Charakterisierungs- und Sequenz-Daten detailliert analysiert wurden. Insgesamt wurden 1892 Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem Ersterkrankungsalter von 3 bis 97 Jahren analysiert (Median= 65 Jahre). Zu Beginn wurde der komplette Datensatz hinsichtlich der Parameter Geschlecht und Lymphom-Entität ausgewertet, worin die Altersverteilungen der analysierten Patienten zeigten, dass mit dem Alter die Anzahl an weiblichen Patienten zunahm und überwiegend DLBCL diagnostiziert wurde (Abbildung 26). Aufgrund der unterschiedlichen Altersstrukturen der Patienten mit DLBCL und FL wurden die Entitäten in den weiteren Analysen separat betrachtet.

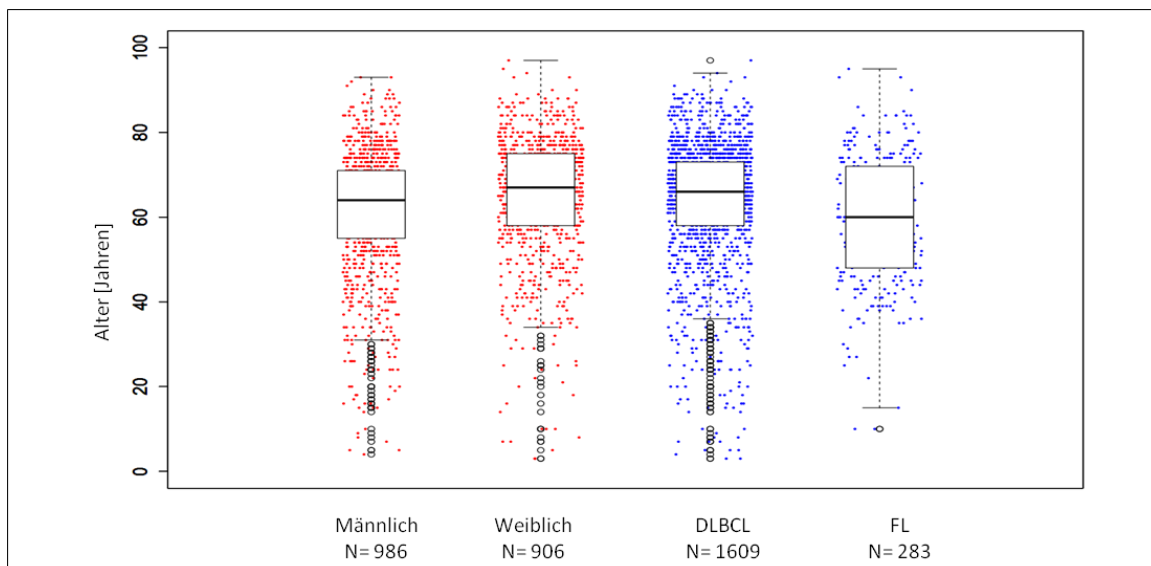


Abbildung 26: Altersverteilung der Geschlechter und Lymphom-Entitäten von Keimzentrums-B-Zell-Lymphomen

In den grafischen Darstellungen stellt jeder Punkt einen analysierten Patienten mit dem entsprechenden Charakteristika dar. Ersterkrankungsalter der Patienten in Bezug auf Geschlecht (rot) und Lymphom-Entitäten (blau). In der Analyse der Altersverteilung der Patienten hinsichtlich des Geschlechts waren mit zunehmendem Alter häufiger Frauen erkrankt. Zusätzlich wurden bei Patienten im hohen Alter häufiger DLBCL als FL diagnostiziert.

3.2.1.1 Charakterisierung genetischer Aberrationen bei folliculären Lymphomen

Im Rahmen der vorliegenden Arbeit wurden insgesamt 283 Patienten mit einer Diagnose folliculäres Lymphom (FL) und einem Ersterkrankungsalter von 10 bis 95 Jahren analysiert (Median: 60 Jahre). Dieses Patienten-Kollektiv setzte sich aus sehr alten Patienten der neu erstellten Kohorte *Jack* (N= 38) und 245 Patienten aus den Forschungsstudien MMML (N= 188) und ICGC MMML-Seq (N= 57) zusammen (Abbildung 27A). Die klinischen Studien wurden in diesen Analysen nicht berücksichtigt, da der Autorin keine Daten zu FL-Patienten aus NHL-B und RICOVER vorlagen.

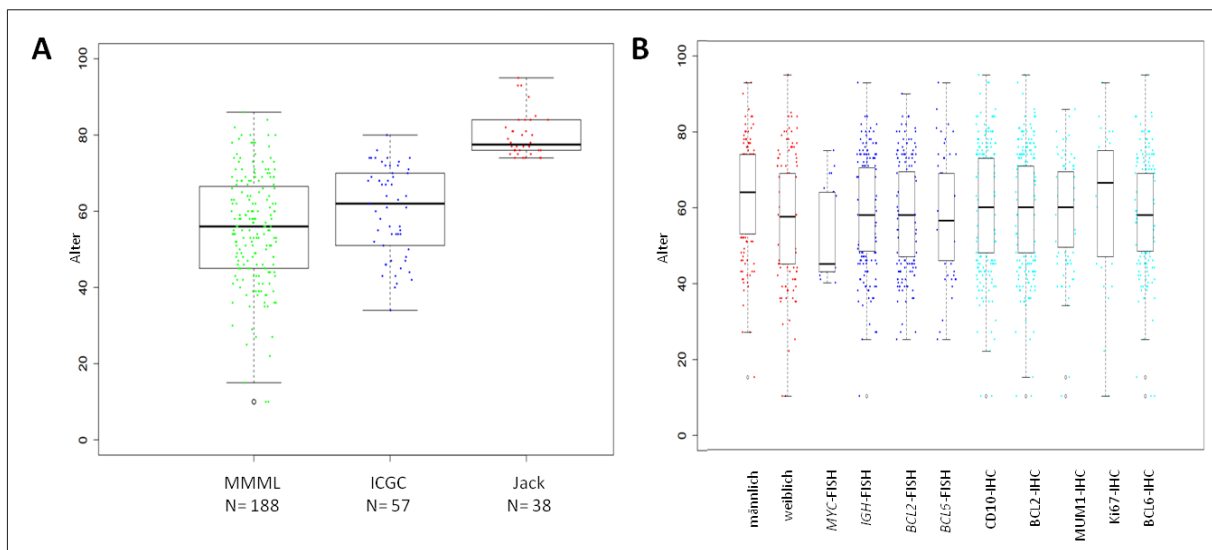


Abbildung 27: Patienten-Kohorte und Altersverteilung der analysierten Charakteristika von FL-Patienten

A: Altersverteilung von Patienten aus verschiedenen klinischen und Forschungsstudien. Die Betrachtung der analysierten Studien zeigte, dass eine Kohorte von FL-Patienten im breiten Altersspektrum von 10-93 Jahren analysiert wurden (N= 265). B: Altersverteilung der Patienten in Bezug auf Geschlecht (rot), Charakteristika der FISH- (dunkelblau) und IHC-Daten (hellblau). Der Vergleich der Altersverteilung der Patienten zeigte, dass die Patienten im hohen Alter häufiger Männer sind und chromosomale Imbalancen im *IGH*-Lokus aufweisen bzw. verstärkt das Protein Ki67 exprimieren. Abkürzungen: FISH: Daten aus den Fluoreszenz-in-situ-Hybridisierung-Analysen. IHC: Daten aus den Immunhistochemie-Analysen.

Für alle 283 Patienten lagen die Information zum Geschlecht und für viele Patienten die Daten zur Basis-Charakterisierung vor. Die Betrachtung der Geschlechter-Verteilung zeigte, dass annähernd bei genauso vielen Frauen (138/283, 49%) wie Männern (145/283, 51%) ein FL diagnostiziert wurde. Die Ergebnisse der FISH-Analysen zeigten, dass bei FL-Patienten chromosomale Imbalancen unter Beteiligung der Gene *BCL2* (188/263, 71%), *BCL6* (58/260, 22%), *MYC* (23/253, 9%) und *IGH* (199/262, 76%) vorlagen. Weiterhin wurden in fast allen Tumor-Proben die Proteine *BCL2* (261/273, 96%), *BCL6* (188/196, 96%) und *CD10* (238/262, 89%), sowie in etwa der Hälfte das Protein *MUM1/IRF4* (91/165, 55%) exprimiert. Hingegen war die Expression des Proteins *Ki67* in den Tumor-Proben deutlich geringer (32/263, 12%).

Anschließend wurde das Geschlecht und die Basis-Charakterisierungs-Daten der FL-Patienten im Zusammenhang mit dem Alter des Patienten bei Erstdiagnose betrachtet (Abbildung 27B). Die Ergebnisse spiegelten wider, dass mit zunehmendem Alter bei mehr Männern (Median: 64 Jahre) als Frauen (Median: 58 Jahre) ein FL diagnostiziert wurde. Zudem lagen Translokationen im *MYC*-Gen eher bei jüngeren Patienten vor. Assoziationen zwischen dem Ersterkrankungsalter der Patienten und chromosomalen Imbalancen unter Beteiligung der Gene *BCL6*, *BCL2* und *IGH* sowie die Expression der Proteine BCL6, CD10, BCL2, MUM1/IRF4 und Ki67 waren anhand der erstellten Grafiken nicht eindeutig erkennbar (Abbildung 27B).

Abschließend wurde die Verteilung der Geschlechter sowie die Basis-Charakterisierungs-Daten in den verschiedenen Altersgruppen genauer betrachtet (Gruppe 1: 0-39 Jahre, Gruppe 2: 40-59 Jahre, Gruppe 3: 60-79 Jahre und Gruppe 4: 80-97 Jahre). Die Ergebnisse zeigten, dass etwa 82% (232/283) der Patienten ein Ersterkrankungsalter zwischen 40 und 80 Jahren hatten. Nur etwa 10% (27/283) der Patienten waren sehr jung (Gruppe 1) bzw. 8% sehr alt (Gruppe 4). Eine detaillierte Übersicht zu den Ergebnissen der Basis-Charakterisierung in Bezug auf die vier Altersgruppen ist im Anhang in der Tabelle A4 zusammengestellt. Insbesondere spiegelten die FISH- und IHC-Daten wider, dass FL-Patienten mit zunehmendem Alter seltener chromosomale Imbalancen in den Genen *MYC* (Gruppe 2: 12%, Gruppe 4: 0%) und *BCL2* (Gruppe 2: 78%, Gruppe 4: 65,0%) aufwiesen und weniger *BCL6* exprimieren (Gruppe 1: 100%, Gruppe 4: 86,0%). Anschließend wurden die Ergebnisse der FISH- und IHC-Analysen mit den statistischen Verfahren der logistischen Regression (LR) und des Kruskal-Wallis Tests (KW) ausgewertet. Dabei wurde analysiert inwieweit ein statistischer Zusammenhang (P-Wert) zwischen den einzelnen analysierten Charakteristika in Bezug auf das Erstdiagnosealter des Patienten besteht. Alle Charakteristika mit einem ermittelten signifikanten P-Wert (P-Wert $\leq 0,050$) in der LR oder KW wurden anschließend grafisch dargestellt (Abbildung 28).

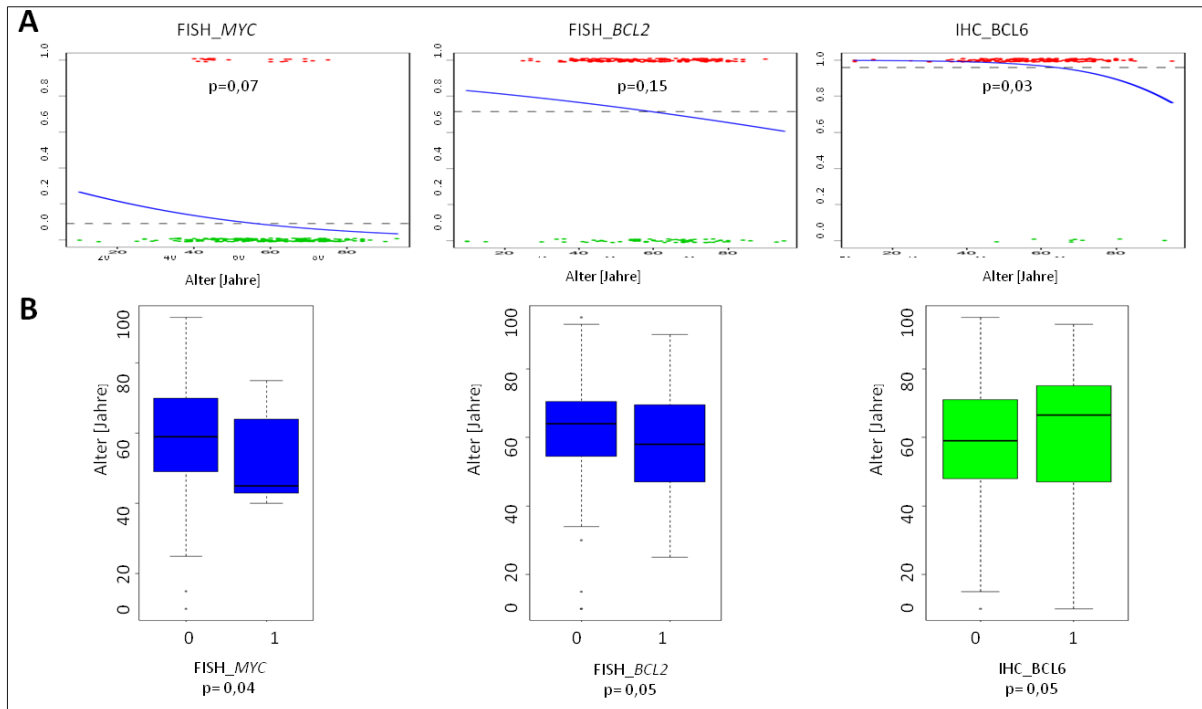


Abbildung 28: Statistik von altersassoziierten Charakteristika mit signifikantem P-Wert bei FL-Patienten

In dieser Grafik sind nur die Charakteristika dargestellt, die mittels logistischer Regression oder Kruskal-Wallis Test mit dem Ersterkrankungsalter der Patienten assoziiert waren. A: Logistische Regression der mittels FISH erhobenen Daten für die Gene *MYC* und *BCL2* und der mittels Immunhistochemie analysierten Expression des Proteins *BCL6*. In der Darstellung der logistischen Regression stellt jeder Punkt einen analysierten Patienten dar. Eine Aberration wurde in Form von roten Punkten in der Grafik eingefügt. Keine Aberration/ Wildtyp ist durch grüne Punkte gekennzeichnet. Die gestrichelte Linie stellt das mediane Alter des Patienten-Kollektivs dar. Auf der x-Achse ist das Ersterkrankungsalter der Patienten mit zunehmendem Alter aufgetragen. B: Kruskal-Wallis Test der mittels FISH erhobenen Daten für die Gene *MYC* und *BCL2* und der mittels Immunhistochemie analysierten Expression des Proteins *BCL6*. Auf der x-Achse sind die Werte 0 und 1 dargestellt, wobei 0 für negativ (Wildtyp) und 1 für positiv (Bruch, Imbalanz) steht.

Zusammenfassend konnten in den statistischen Auswertungen chromosomale Imbalanzen der Gene *MYC* und *BCL2* sowie die Expression des Proteins *BCL6* als signifikante altersassoziierte genetische Aberrationen in FL-Patienten identifiziert werden.

3.2.1.1.1 Mutationsprofil in folliculären Lymphomen

Zur Identifizierung von altersassoziierten Mutationen von FLs lagen insgesamt Hochdurchsatz-Sequenzierungs-Daten von 50 Patienten im Alter von 40 bis 84 Jahren vor (Median: 62,5 Jahre). Der Datensatz setzte sich aus den Daten der Target-Resequenzierung mittels Lymphom-Panel (N=4, Median: 79 Jahre) und den WGS-Daten aus dem ICGC MMML-Seq Projekt (N= 46, Median: 62 Jahre) zusammen. Bei der Auswertung der Hochdurchsatz-Sequenzierungs-Daten der Keimzentrums-B-Zell-Lymphome ist zu beachten, dass von jedem Patienten des ICGC MMML-Seq die Tumor-Probe und die korrespondierende nicht-Tumor-Probe sequenziert wurden. Anschließend wurden alle identifizierten Mutationen der nicht-Tumor-Probe aus den Mutationsdaten der Tumor-Probe entfernt, um mögliche Keimbahnmutationen in der weiteren Analyse auszuschließen. Für die Lymphom-Panel-Proben lagen nur Tumor-Proben vor, weshalb zusätzlich B-Zell-Populationen gesunder älterer Individuen als Kontrollen verwendet wurden.

Hierzu wurden FACS-sortierte B-Zellen aus dem peripheren Blut von zwei gesunden Spendern sequenziert und alle Sequenzdaten in der bioinformatischen Auswertung zu einem Kontrollpool zusammengefasst. Daraufhin wurden in der vorliegenden Arbeit alle detektierten Mutationen des Kontrollpools von den detektierten Mutationen der Tumor-Proben (Lymphom-Panel-Tumor-Proben und ICGC MMML-Seq Patienten) abgezogen. Zusätzlich wurden die Mutationsdaten der Tumor-Proben gegen Mutationen mit einer MAF von über 1% in der Normalbevölkerung gefiltert. Eine detaillierte Beschreibung des Vorgehens bei der bioinformatischen Auswertung ist im Abschnitt 2.2.5.5 nachzulesen. Die finale Mutationsanalyse der Keimzentrums-B-Zell-Lymphome erfolgte an dem gefilterten Datensatz und ausschließlich für Tumor-Proben die eine minimale On-Target-Coverage (OTC) von 30x in der Hochdurchsatz-Sequenzierung aufwiesen (Qualitätskontrolle siehe Abschnitt 3.1.4). Insgesamt wurden 241 somatische Mutationen in den 50 analysierten FL-Patienten identifiziert, die als Basis-Datensatz für die primäre altersassoziierte Mutationsanalyse in FL dienen. Die potentiell proteinverändernde Mutationen wurden in 61 Genen detektiert, wobei 14/241 (6%) Mutationen die Splicing-Region der analysierten Transkripte betrafen. 155/241 (64%) waren nicht-synonyme Mutationen, 35/241 (15%) Frameshift Deletionen oder Insertionen, 34/241 (14%) Stopp-Mutationen und 3/241 (1%) nicht-Frameshift Deletionen. In der weiteren Mutationsanalyse wurden zuerst alle Gene identifiziert, die in mindestens 5 Patienten eine potenziell proteinverändernde Aberration zeigten. Diese Gene werden im Folgenden als rekurrent mutiert bezeichnet. Zur Beurteilung der Häufigkeit von Mutationen in den rekurrent mutierten Genen wurde die Mutationsfrequenz (MF) ermittelt und grafisch dargestellt (Abbildung 29A). Innerhalb dieses Datensatzes konnten insgesamt neun Gene als rekurrent mutiert in FL identifiziert werden, wobei jedes dieser neun Gene eine MF von mind. 10% aufwies.

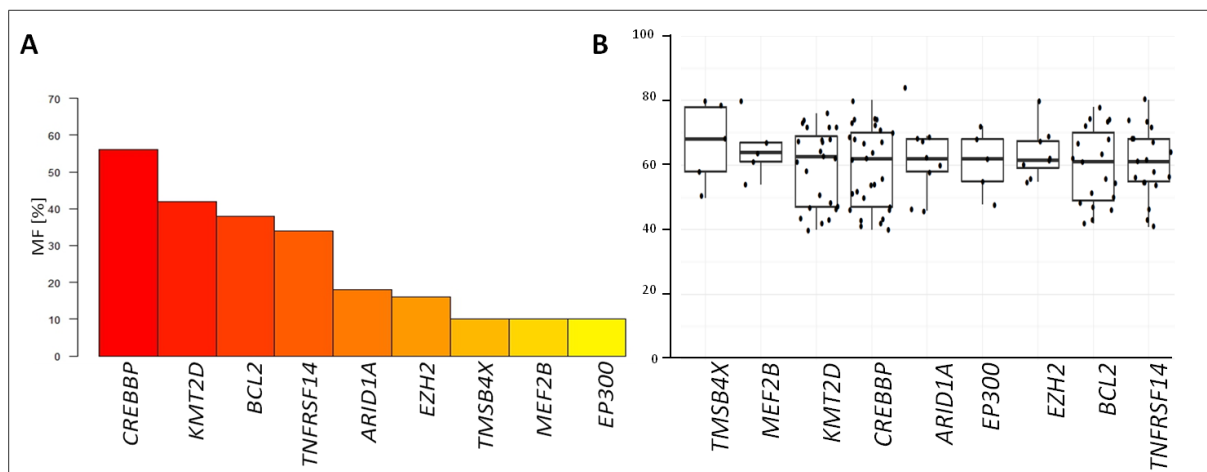


Abbildung 29: Mutationsfrequenz und Altersverteilung der rekurrent mutierten Gene in FL

Als rekurrent mutiert wurden alle Gene bezeichnet, die in mindestens 5 Patienten potentiell proteinverändernde Mutationen aufwiesen. A: Um die Unterschiede in der Mutationsfrequenz (MF) der einzelnen Gene besser darzustellen, wurde in dem Balkendiagramm die Skalierung der y-Achse von 100% auf 70% reduziert. Insgesamt wurden in 50 analysierten FL-Patienten 9 Gene als rekurrent mutiert identifiziert, die häufig in FL mutiert sind und eine MF von mind. 10% aufwiesen. B: Das höchste mediane Alter in der Gruppe von Patienten mit einer Mutation wurde mit 68 Jahren im *TMSB4X*-Gen ermittelt.

Zur Visualisierung der Altersverteilung der rekurrent mutierten Gene wurden diese Gene in Form von Boxplots dargestellt und nach dem medianen Alter der entsprechend mutierten Patienten geordnet (Abbildung 29B). Das durchschnittliche mediane Alter der mutierten Patienten über die neun Gene lag bei 63 Jahren. Für das *TMSB4X*-Gen wurde in der analysierten Patienten-Kohorte das höchste mediane Alter mit 68 Jahren für Mutationen in FL ermittelt (Tabelle 25).

Tabelle 25: Altersverteilung und Statistik der rekurrent mutierten Gene in FL

| Gen | min. Alter [Jahre] | max. Alter [Jahre] | Medianes Alter [Jahre] | LR P-Wert | MF [%] |
|-----------------|--------------------|--------------------|------------------------|-----------|--------|
| <i>TMSB4X</i> | 58 | 78 | 68 | 0,388 | 10 |
| <i>MEF2B</i> | 61 | 67 | 64 | 0,501 | 10 |
| <i>KMT2D</i> | 47 | 69 | 62 | 0,050 | 42 |
| <i>CREBBP</i> | 47 | 70 | 62 | 0,100 | 56 |
| <i>ARID1A</i> | 58 | 68 | 62 | 0,925 | 18 |
| <i>EP300</i> | 55 | 68 | 62 | 0,877 | 10 |
| <i>EZH2</i> | 59 | 68 | 62 | 0,933 | 16 |
| <i>BCL2</i> | 49 | 71 | 61 | 0,276 | 38 |
| <i>TNFRSF14</i> | 55 | 68 | 61 | 0,985 | 34 |

Nur für das *KMT2D*-Gen konnte mittels logistischer Regression ein signifikanter P-Wert von 0,05 ermittelt (grau markiert). Abkürzungen und Verweise: min. Alter: minimales Alter der entsprechend mutierten Patienten innerhalb des Boxplots (untere Quartilgrenze in Abbildung 30B). max. Alter: maximales Alter der entsprechend mutierten Patienten innerhalb des Boxplots (obere Quartilgrenze in Abbildung 30B). Medianes Alter: medianes Alter der entsprechend mutierten Patienten. LR: Logistische Regression. MF: Mutationsfrequenz innerhalb der analysierten Patienten-Kohorte.

Die Ergebnisse der Mutationsanalysen aus der Abbildung 29 wurden mit dem statistischen Verfahren der logistischen Regression (LR) ausgewertet und der P-Wert für jedes rekurrent mutierte Gen in Bezug auf das Ersterkrankungsalter des Patienten berechnet. Die statistische Auswertung konnte nur für das *KMT2D*-Gen einen signifikanten P-Wert von 0,050 ermitteln und die LR zeigte, dass mit zunehmendem Alter die Mutationsrate bei FL-Patienten in *KMT2D* signifikant abnahm (Abbildung 30). Alle weiteren acht Gene zeigten mittels LR keine signifikante Assoziation zwischen den Mutations-Ergebnissen und dem Alter der Patienten (Tabelle 25).

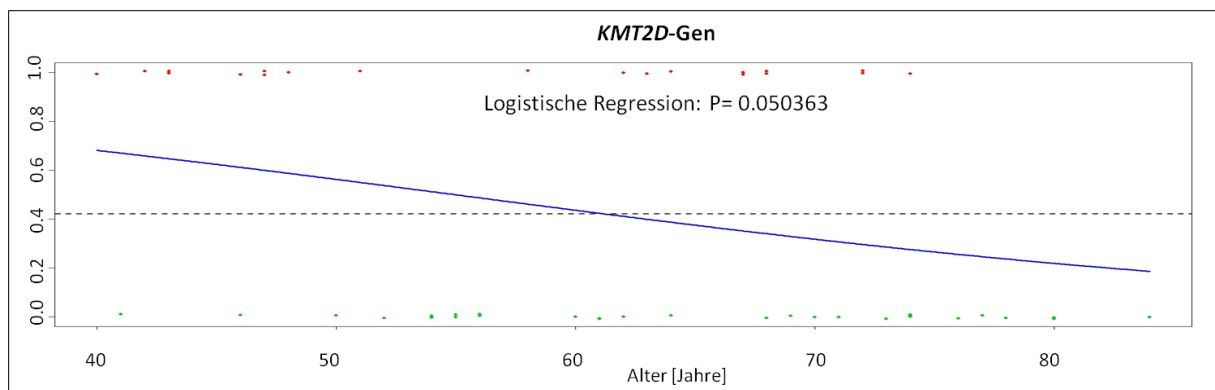


Abbildung 30: Statistische Auswertung von FL-Patienten mit einer Mutation im *KMT2D*-Gen

Um die Unterschiede des Mutationsstatus in *KMT2D* besser zu veranschaulichen, wurde in der Grafik die Skalierung der x-Achse an die Altersverteilung der analysierten Patienten-Kohorte angepasst und die x-Achse von 40 bis 85 Jahren dargestellt. In der Darstellung der logistischen Regression stellt jeder Punkt einen analysierten Patienten dar. Eine Aberration wurde in Form von roten Punkten in der Grafik eingefügt. Keine Aberration/ Wildtyp ist durch grüne Punkte gekennzeichnet. Die gestrichelte Linie stellt das mediane Alter des Patienten-Kollektivs dar. Auf der x-Achse ist das Ersterkrankungsalter der Patienten mit zunehmendem Alter aufgetragen. Die Grafik spiegelt die Ergebnisse der logistischen Regression der Mutationsanalyse von FL-Patienten im *KMT2D*-Gen wider. Die statistische Analyse zeigte, dass mit zunehmendem Alter signifikant weniger potentiell proteinverändernde Mutationen in FL-Patienten identifiziert wurden.

Zur Beurteilung der Aberrationen im *KMT2D*-Gen wurde eine grafische Übersicht erstellt, worin alle identifizierten Mutationen hinsichtlich der Mutationstypen und der Position im Protein eingetragen wurden (Abbildung 31). Das Mutationsprofil im *KMT2D*-Gen der FL-Patienten zeigte, dass die Aberrationen überwiegend trunkierende Mutationen (32/37) waren, die über die komplette kodierende Sequenz des Gens verteilt waren. Bei 12/24 Patienten wurde nur eine Mutation in *KMT2D* identifiziert. Zusätzlich wurden im *KMT2D*-Gen von 11/24 Patienten jeweils zwei und in 1/24 Patienten drei Mutationen nachgewiesen. Bei diesen Patienten war mind. eine der detektierten Mutationen trunkierend. Ob es sich bei den Mutationen um bi- oder monoallelische Aberrationen handelt, kann mit der angewandten Methode nicht nachvollzogen werden, da keine Angaben zu dem mutierten Allel existieren.

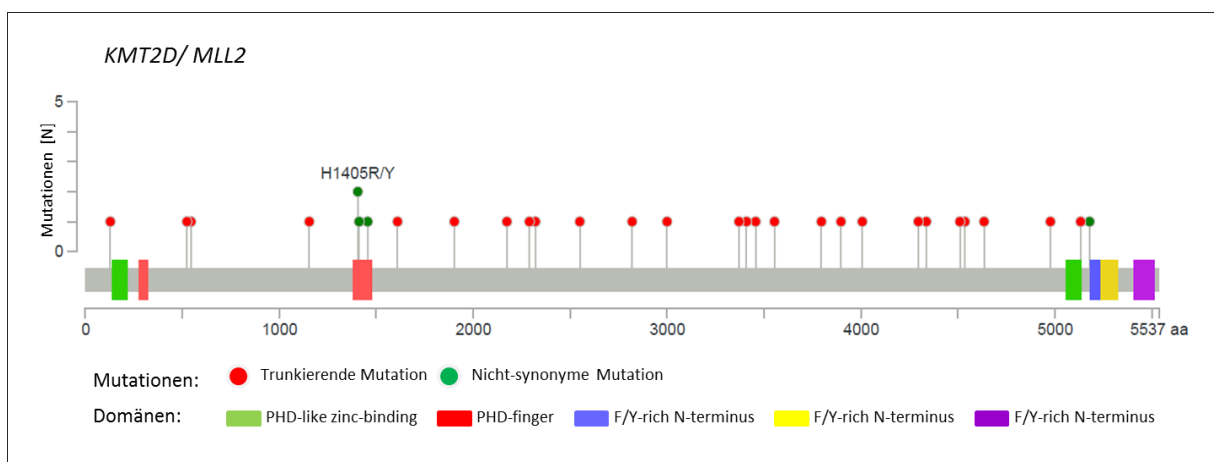


Abbildung 31: Schematische Darstellung der Mutationen im *KMT2D*-Gen in FL

Die Domänen-Namen wurden nicht in die deutsche Sprache übersetzt und sind in Englischer Sprache entsprechend den Angaben des webbasierten Tools MutationMapper übernommen. In der Grafik sind Positionen gekennzeichnet an denen zwei unterschiedliche Mutationen in der genomischen Sequenz identifiziert wurden. Die entsprechenden Mutationen sind darüber angegeben (H1405R/Y). D.h. an dieser Position wurde die Mutation H1405R und H1405Y in der analysierten Patienten-Kohorte detektiert. Insgesamt verteilten sich die identifizierten Mutationen in *KMT2D* über die kodierende Sequenz des Gens, wobei überwiegend trunkierende Mutationen identifiziert wurden (N= 32/37).

Zusammenfassend konnte in der statistischen Analyse der Hochdurchsatz-Sequenzierung das *KMT2D*-Gen (*MLL2*) als signifikanter Marker für das Ersterkrankungsalter bei FL-Patienten identifiziert werden (medianes Alter der betroffenen Patienten: 62 Jahre).

3.2.1.1.2 Zusammenfassung der altersassoziierten genetischen Marker in Patienten mit einem folliculären Lymphom

In der vorliegenden Arbeit wurden Patienten mit einem Follikulären Lymphom (FL) und einem Ersterkrankungsalter von 10 bis 93 Jahren mit verschiedenen molekulargenetischen Methoden analysiert. Mittels der Basis-Charakterisierung wurden verschiedene altersassoziierte Marker identifiziert, die sich mit dem Ersterkrankungsalter der FL-Patienten veränderten. Zusammenfassend wurde in den altersassoziierten Analysen dieser Arbeit gezeigt, dass mit zunehmendem Alter bei mehr Männern ein FL diagnostiziert wurde. Chromosomale Imbalancen unter Beteiligung des *IGH*-Lokus nahmen mit dem Alter signifikant zu und bei den Genen *BCL6* und *MYC* ab. In den Hochdurchsatz-Sequenzierungsanalysen wurden 9 Gene identifiziert, die bei FL-Patienten rekurrent mutiert vorlagen. Das Gen *KMT2D* wurde als signifikanter Marker in FL identifiziert, wobei die Mutationsfrequenz mit dem Ersterkrankungsalter der betroffenen Patienten abnahm (Median: 62 Jahre) und vorwiegend trunkierende Mutationen in der kodierenden Region vorlagen.

Da der Fokus der vorliegenden Arbeit auf der Analyse von altersassoziierten Aberrationen in Keimzentrums-B-Zell-Lymphomen mit einem sehr hohen Ersterkrankungsalter von vorwiegend über 80 Jahren lag und das mediane Alter der analysierten FL-Patienten bei 58 Jahren lag, wurden im Folgenden die genetische Aberrationen bei DLBCL-Patienten genauer betrachtet.

3.2.1.2 Charakterisierung genetischer Aberrationen bei diffus großzelligen B-Zell-Lymphomen

Im Rahmen der vorliegenden Arbeit wurden insgesamt 1609 Patienten mit einer Diagnose DLBCL und einem Ersterkrankungsalter von 3 bis 97 Jahren analysiert (Median: 66 Jahre). Dieses Kollektiv setzte sich aus Patienten der neu erstellten Kohorte Jack (N= 160, Median: 80 Jahre), 492 Patienten aus den Forschungsstudien MML (N= 445, Median: 64 Jahre) und ICGC MML-Seq (N= 56, Median: 62 Jahre) sowie 948 Patienten aus den klinischen Studien NHL-B (N= 511, Median: 61 Jahre) und RICOVER (N= 437, Median: 68 Jahre) zusammen.

Für alle insgesamt 1609 Patienten lagen die Daten zum Geschlecht, dem histopathologischen DLBCL-Subtypen und der Immunhistochemie vor. Zudem waren für einen Großteil der Patienten die FISH-Daten (N= 1561, Details siehe Tabelle 26 am Ende dieses Kapitels), die Daten zum DLBCL-Subtyp nach dem *Hans*- (N=1189) und *GEP-Klassifikator* (N= 398) vorhanden. Die Betrachtung der Geschlechter-Verteilung zeigte, dass in der analysierten Patienten-Kohorte etwa genauso häufig bei Männern (841/1609, 52%) wie Frauen (768/1609, 48%) ein DLBCL diagnostiziert wurde. Die Ergebnisse der histologischen DLBCL-Subtypen spiegelten wider, dass 1074/1609 (67%) centroblastisch, 139/1609 (9%) immunoblastisch, 36/1609 (2%) anaplastisch-großzellig, 30/1609 (2%) T-Zellreich waren und 311/1609 (19%) dem Subtyp NOS (engl. not otherwise specified; nicht anders klassifiziert) zugeordnet wurden.

Die Ergebnisse der FISH-Analysen zeigten, dass bei DLBCL-Patienten chromosomale Imbalancen unter Unterbeteiligung der Gene *IGH* (363/757, 48%), *BCL6* (273/971, 28%), *BCL2* (145/956, 15%) und *MYC* (106/988, 11%) vorlagen. Weiterhin wurden in vielen Tumor-Proben verstärkt die Proteine MUM1/IRF4 (996/1234, 81%), BCL6 (973/1194, 82%), BCL2 (1077/1336, 81%), und Ki67 (808/1325, 61%) exprimiert. Wohingegen die Expression von CD10 in der analysierten Patienten-Kohorte deutlich geringer war und das Protein in weniger als der Hälfte der Tumore exprimiert wurde (517/1343, 39%).

Anschließend wurde das Geschlecht, die histopathologischen DLBCL-Subtypen, die Basis-Charakterisierungs-Daten sowie der *Hans*- und *GEP-Klassifikator* der DLBCL-Patient im Zusammenhang mit dem Alter des Patienten bei Erstdiagnose betrachtet (einzelne Charakteristika sind in der Abbildung 32 dargestellt). Die Ergebnisse zeigten, dass mit zunehmendem Alter bei mehr Frauen (Median: 68 Jahre) als Männern (Median: 64 Jahre) ein DLBCL diagnostiziert wurde. In der Betrachtung der gesamten Patienten-Kohorte hinsichtlich der histologischen DLBCL-Subtypen und der Basis-Charakterisierungs-Daten war ohne statistische Verfahren keine direkte Assoziation der analysierten Charakteristika zum Alter zu erkennen (Abbildung 32).

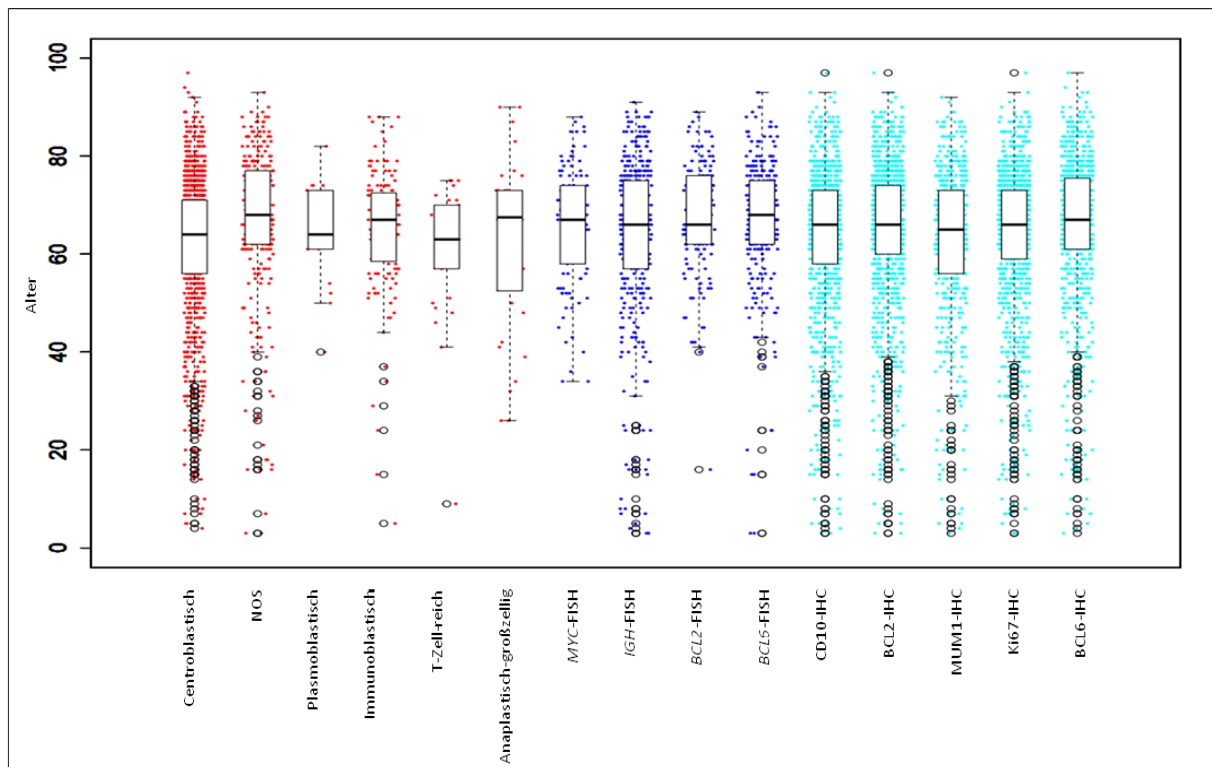


Abbildung 32: Altersverteilung von analysierten Charakteristika in DLBCL-Patienten

In dieser Grafik sind nur die Ergebnisse der histopathologischen DLBCL-Subtypen (rot), FISH- (dunkelblau) und IHC-Analysen (hellblau) von den 1503 analysierten DLBCL-Patienten in Assoziation zum Ersterkrankungsalter der Patienten dargestellt. Der Vergleich der Altersverteilung der Patienten zeigte ohne statistische Verfahren keinen direkten Zusammenhang zwischen den histopathologischen DLBCL-Subtypen und den Basis-Charakterisierungs-Daten. Abkürzungen: FISH: Daten aus den Fluoreszenz-in-situ-Hybridisierung-Analysen. IHC: Daten aus den Immunhistochemie-Analysen.

Die Altersverteilung in DLBCL hinsichtlich des *Hans-Klassifikators* verdeutlicht, dass mit zunehmendem Alter häufiger der nonGCB- (Median: 67 Jahre) als GCB-Subtyp (Median: 65 Jahre) identifiziert wurde und die Ergebnisse mittels *GEP-Klassifikator* zeigten, dass im höheren Alter häufiger der ABC- (Median: 67 Jahre) als GCB-Subtyp (59 Jahre) identifiziert wurde.

Abschließend wurde die Verteilung der Geschlechter, histopathologischen Subtypen, Basis-Charakterisierungs-Daten und den DLBCL-Subtypen in den verschiedenen Altersgruppen genauer betrachtet (Gruppe 1: 0-39 Jahre, Gruppe 2: 40-59 Jahre, Gruppe 3: 60-79 Jahre und Gruppe 4: 80-97 Jahre). Eine detaillierte Übersicht zu den Ergebnissen der analysierten Charakteristika in Bezug auf die vier Altersgruppen ist in der Tabelle 26 am Ende dieses Kapitels zusammengestellt.

Insgesamt zeigten die Ergebnisse, dass über 70% (1178/1609) der analysierten DLBCL-Patienten ein Ersterkrankungsalter zwischen 60 und 100 Jahren hatten. Nur etwa 20% (313/1609) der Patienten waren junge (Gruppe 2) bzw. 7% (118/1609) sehr junge Patienten (Gruppe 1). Der Vergleich der histopathologischen DLBCL-Subtypen verdeutlichte, dass im hohen Alter mehr DLBCL vom Subtyp NOS (Gruppe 3: 20%, Gruppe 4: 36%) und weniger centroblastische DLBCL (Gruppe 3: 65%, Gruppe 4: 55%) diagnostiziert wurden. Die Betrachtung der FISH-Daten zeigte keine Veränderungen von chromosomalen Imbalanzen der Gene *BCL2*, *BCL6*, *MYC* und *IGH*, die vorwiegend bei Patienten im hohen Alter auftraten.

Die FISH-Ergebnisse zeigten lediglich eine altersassoziierte Veränderung beim Extremgruppenvergleich, wobei Veränderungen der Häufigkeit der chromosomalen Imbalanzen eher zwischen den sehr jungen (Gruppe 1) und jungen Patienten (Gruppe 2) identifiziert wurden. Hierbei wurden Brüche in den Genen *MYC* (Gruppe 1: 6%, Gruppe 2: 16%), *BCL2* (Gruppe 1: 2%, Gruppe 2: 21%), *BCL6* (Gruppe 1: 15%, Gruppe 2: 30%) und *IGH* (Gruppe 1: 45%, Gruppe 2: 56%) deutlich häufiger ab einem Ersterkrankungsalter von 40 Jahren identifiziert. Brüche im *BCL2*- (Gruppe 4: 20%) und *IGH*-Gen (Gruppe 4: 48%) scheinen bei senilen Patienten wieder abzunehmen, wohingegen sich die Häufigkeit von Brüchen in *MYC* (Gruppe 4: 11%) und *BCL6* (Gruppe 4: 27%) auch bei senilen Patienten kaum veränderte. Die IHC-Daten zeigten, dass die Expression des Proteins CD10 (Gruppe 1: 48,0%, Gruppe 4: 35%) kontinuierlich mit Zunahme des Alters des Patienten abnahm und Ki67 (Gruppe 1: 52%, Gruppe 4: 79%) sowie *BCL2* (Gruppe 1: 72%, Gruppe 4: 84%) mit zunehmendem Alter verstärkt exprimiert wurden. Zudem spiegelten die IHC-Ergebnisse wider, dass die Expression des Proteins MUM1/IRF4 (Gruppe 1: 75%, Gruppe 3: 83%) bis zum Alter von 80 Jahren kontinuierliche zunahm und in den Tumoren der senilen Patienten deutlich weniger MUM1/IRF4 (Gruppe 4: 74%) vorlag. Die Expression des Proteins *BCL6* in den Tumoren veränderte sich kaum mit dem Ersterkrankungsalter der Patienten (Gruppe 1: 84%, Gruppe 4: 80%).

Die Ergebnisse des Hans-Klassifikators zeigten, dass ab 60 Jahren verstärkt DLBCL vom nonGCB-Subtyp diagnostiziert wurden (Gruppe 1 und 2: 144/327 44%, Gruppe 3 und 4: 447/862 52%). Unter Verwendung des GEP-Klassifikators wurden ab 60 Jahren deutlich häufiger DLBCL vom ABC-Subtyp nachgewiesen (Gruppe 1 und 2: 43/171 25%, Gruppe 3 und 4: 104/226 46%).

Anschließend wurden die Ergebnisse der Verteilung der Geschlechter, histopathologischen Subtypen, Basis-Charakterisierungs-Daten und den DLBCL-Subtypen mit den statistischen Verfahren der logistischen Regression (LR) und des Kruskal-Wallis Tests (KW) ausgewertet. Eine Übersicht einzelner signifikanter Charakteristika ist in der nachfolgenden Abbildung 33 dargestellt. Alle berechneten Ergebnisse sind in der Tabelle 26 aufgelistet.

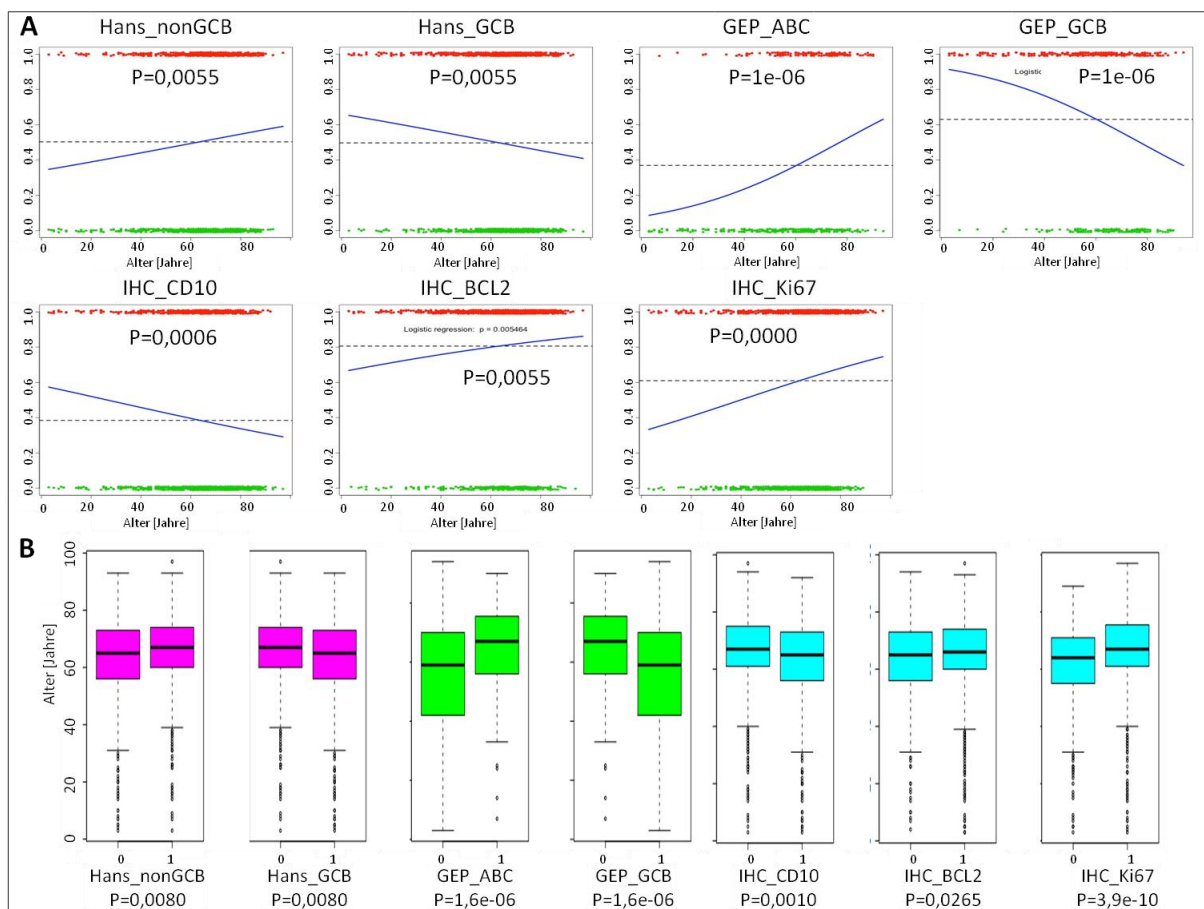


Abbildung 33: Statistik von altersassoziierten Charakteristika mit significantem P-Wert bei DLBCL-Patienten

A: Logistische Regression der Altersverteilung von Patienten hinsichtlich DLBCL-Subtypen nach dem *Hans*- und *GEP*-Klassifikator sowie Charakteristika der IHC-Daten. In der Darstellung der logistischen Regression stellt jeder Punkt einen analysierten Patienten dar. Eine Aberration wurde in Form von roten Punkten in der Grafik eingefügt. Keine Aberration/Wildtyp ist durch grüne Punkte gekennzeichnet. Die gestrichelte Linie stellt das mediane Alter des Patienten-Kollektivs dar. Auf der x-Achse ist das Ersterkrankungsalter der Patienten mit zunehmendem Alter aufgetragen. B: Kruskal-Wallis Test der Altersverteilung von Patienten hinsichtlich DLBCL-Subtypen nach dem *Hans*- und *GEP*-Klassifikator sowie Charakteristika der IHC-Daten. Auf der x-Achse sind die Werte 0 und 1 dargestellt, wobei 0 für negativ (Wildtyp) und 1 für positiv steht.

Die statistischen Analysen ergaben, dass im hohen Alter bei signifikant mehr Frauen ein DLBCL diagnostiziert wurde (LR: $P = <0,0010$; KW: $P = <0,0010$). Zudem wurde bei senilen Patienten signifikant häufiger ein DLBCL vom Subtyp NOS (LR: $P = <0,0010$; KW: $P = <0,0010$) diagnostiziert, wohingegen centroblastische DLBCL (LR: $P = <<0,0010$; KW: $P = <0,0010$) signifikant häufiger bei jüngeren Patienten identifiziert wurden. In der statistischen Auswertung der Basis-Charakterisierungs-Daten wurden signifikante P-Werte für die Proteine BCL2 (LR: $P = 0,006$; KW: $P = 0,023$), Ki67 (LR: $P = <0,0010$; KW: $P = <0,0010$) und CD10 (LR: $P = <0,0010$; KW: $P = <0,0010$) berechnet. Die Proteinexpression von BCL6 und MUM1/IRF4 sowie die analysierten Gene *MYC*, *BCL2*, *BCL6* und *IGH* wurden nicht als signifikante Marker in Bezug auf das Ersterkrankungsalters des Patienten identifiziert.

Die statistische Auswertung der *Hans*- und *GEP*-Klassifikatoren zeigte, dass bei jüngeren Patienten signifikant häufiger der GCB-Subtyp (*Hans*: LR: $P = 0,010$; KW: $P = 0,010$; *GEP*: LR: $P = <0,0010$; KW: $P = <0,0010$) nachgewiesen wurde. Mit zunehmendem Alter der Patienten waren die Tumore signifikant häufiger vom nonGCB- (*Hans*: LR: $P = 0,010$; KW: $P = 0,010$) bzw. ABC-Subtyp (*GEP*: LR: $P = <0,0010$; KW: $P = <0,0010$).

Zusammenfassend konnte in der statistischen Auswertung der Geschlechter, histopathologischen Subtypen, Basis-Charakterisierungs-Daten und DLBCL-Subtypen in Assoziation mit dem Alter der Patienten gezeigt werden, dass mit zunehmendem Alter häufiger bei Frauen ein DLBCL diagnostiziert wurde und dass der histopathologische DLBCL-Subtyp NOS im Alter zunahm. Weiterhin wurde gezeigt, dass DLBCL-Patienten im höheren Alter häufiger Tumore vom ABC- bzw. nonGCB-Subtyp hatten und die Tumore von Patienten im hohen Alter eher eine hohe Expression von BCL2, MUM1 und Ki67 aufwiesen, wohingegen CD10 vorwiegend bei jüngeren Patienten stark exprimiert wurde.

Tabelle 26: Charakteristika des Gesamtkollektives von analysierten Patienten mit einem DLBCL

| | Kinder/Junge | Erwachsene | | Senile | Total | LR P-Wert | KW P-Wert |
|-------------------------------------------|-----------------|------------------|-------------------|------------------|--------------------|--------------|--------------|
| | 0-39 Jahre | 40-59 Jahre | 60-79 Jahre | ≤80 Jahre | 0-100 Jahre | | |
| Median Alter (min. – max.) | 29 (3-39) | 53 (40-59) | 68 (60-79) | 84 (80-97) | 66 (3-97) | | |
| Patienten | 118/ 100% | 313/ 100% | 1041/ 100% | 137/ 100% | 1609/ 100% | | |
| Geschlecht | | | | | | | |
| Männlich | 85/118 72,0% | 183/313 58,5% | 522/1041 50,1% | 51/137 37,2% | 841/1609 52,3% | <0,0001 | <0,0001 |
| Weiblich | 33/118 28,0% | 130/313 41,5% | 519/1041 49,9% | 86/137 62,8% | 768/1609 47,7% | <0,0001 | <0,0001 |
| DLBCL Morphologie | | | | | | | |
| NOS | 20/118 16,9% | 37/313, 11,8% | 204/1041 19,5% | 49/137 35,8% | 311/1609 19,3% | 0,0002 | <0,0001 |
| Centroblatisch | 86/118 72,9% | 226/313 72,2% | 680/1041 65,3% | 75/137 54,7% | 1074/1609 66,7% | 0,0011 | <0,0001 |
| Plasmoblatisch | 0/118 0% | 4/313 1,3% | 12/1041 1,2% | 1/137 0,7% | 17/1609 1,1% | 0,8428 | 0,8491 |
| Immunoblastisch | 6/118 5,1% | 31/313 9,9% | 94/1041 9,0% | 8/137 5,8% | 139/1609 8,6% | 0,3259 | 0,6046 |
| T-Zell-reich | 1/118 0,8% | 8/313 2,6% | 21/1041 2,0% | 0/137 0% | 30/1609 1,9% | 0,2952 | 0,1444 |
| Anaplastisch-großzellig | 5/118 4,2% | 7/313 2,2% | 20/1041 2,0% | 4/137 2,9% | 36/1609 2,2% | 0,6370 | 0,9253 |
| Molekularer Subtyp | | | | | | | |
| <i>GEP-Klassifikator</i> | | | | | | | |
| ABC | 8/61 13,3% | 35/110 31,8% | 88/192 45,8% | 16/34 47,1% | 147/398 36,9% | <0,0001 | <0,0001 |
| GCB | 53/61 83,3% | 75/110 68,2% | 98/192 51,1% | 18/34 52,9% | 251/398 63,1% | <0,0001 | <0,0001 |
| NA | 57/118 48,3% | 203/313 64,9% | 849/1041 81,6% | 102/137 74,5% | 1211/1609 75,3% | | |
| <i>Hans-Klassifikator</i> | | | | | | | |
| nonGCB | 38/94 40,4% | 106/233 45,5% | 396/756 52,4% | 51/106 48,1% | 598/1189 50,3% | 0,0055 | 0,0080 |
| GCB | 56/94 59,6% | 127/233 54,5% | 345/756 45,6% | 55/106 51,9% | 591/1189 49,7% | 0,0055 | 0,0080 |
| NA | 24/118 20,3% | 80/313 25,6% | 285/1041 27,4% | 31/137 22,6% | 420/1609 26,1% | | |
| Fluoreszenz-in-situ-Hybridisierung | | | | | | | |
| MYC Bruch | 3/58 5,7% | 25/160 15,6% | 66/698 9,5% | 12/74 16,2% | 106/988 10,7% | 0,4345 | 0,6260 |
| BCL2 Bruch | 1/66 1,5% | 29/137 21,2% | 96/656 14,6% | 19/96 19,8% | 145/956 15,2% | 0,1617 | 0,9152 |
| BCL6 Bruch | 10/677 14,9% | 41/135 30,4% | 194/672 28,9% | 28/96 29,2% | 273/971 28,1% | 0,0810 | 0,3708 |
| IGH Bruch | 29/64 45,3% | 73/130 56,1% | 216/469 46,1% | 45/93 48,4% | 363/757 48,0% | 0,1789 | 0,1306 |
| Immunhistochemie | | | | | | | |
| CD10 pos | 47/99 47,5% | 113/253 44,7% | 312/863 36,2% | 45/128 35,2% | 517/1343 38,5% | 0,0006 | 0,0010 |
| BCL2 pos | 68/95 71,6% | 191/246 77,6% | 707/862 82,0% | 111/133 83,5% | 1077/1336 80,6% | 0,0055 | 0,0265 |
| BCL6 pos | 77/92 83,7% | 198/241 82,2% | 621/765 81,2% | 77/96 80,2% | 973/1194 81,5% | 0,3380 | 0,4134 |
| Ki67 pos | 52/100 52,0% | 124/254 48,8% | 533/845 63,1% | 99/126 78,6% | 808/1325 61,0% | <0,0001 | <0,0001 |
| MUM1 pos | 70/93 75,3% | 181/232 78,1% | 668/805 83,0% | 77/104 74,0% | 996/1234 80,7% | 0,2900 | 0,2708 |

Abkürzungen und Verweise: min: Minimum; max: Maximum; pos: positiv; NOS: (engl. not otherwise specified) nicht weiter spezifiziert. LR: Logistische Regression. KW: Kruskal Wallis Test. Die Ergebnisse für die analysierten Marker und Charakteristika sind nach den analysierten Altersgruppen und für die Gesamtkohorte zusammengestellt. Im MMML-Verbundprojekt wurden die FISH-Analysen mit einer Fusion-Sonde (engl. *dual color dual fusion probe*) hinsichtlich t(14;18) durchgeführt. Die Ergebnisse sind in dieser Tabelle in den Brüchen der Gene *BCL2* (Chromosom 18) und *IGH* (Chromosom 14) integriert. Aus der NHL-B Studie lagen ausschließlich FISH-Analysen für das *MYC*-Gen vor. Aus allen Studien waren IHC-Daten verfügbar, wonach mittels *Hans-Klassifikator* die Tumor-Proben in die DLBCL-Subtypen nonGCB und GCB eingeteilt wurden. Der *GEP-Klassifikator* war nur für Tumor-Proben der Forschungsstudien MMML und ICGC MMML-Seq verfügbar.

3.2.1.2.1 Mutationsprofil in diffus großzelligen B-Zell-Lymphomen

Im Folgenden wird ausschließlich auf die Auswertung der Daten der Tumor-Proben eingegangen, die eine minimale \emptyset -OTC von 30x in der Hochdurchsatz-Sequenzierung aufwiesen (Qualitätskontrolle siehe Abschnitt 3.1.4). Zur finalen Auswertung der Mutations-Daten wurden insgesamt 84 DLBCL-Patienten in einem Alter von 15 bis 97 Jahren (Median: 75 Jahre) betrachtet. Der Datensatz setzte sich aus den Lymphom-Panel-Proben (N= 37, Median: 83 Jahre) und den WGS-Daten aus dem ICGC MMML-Seq Projekt (N= 47, Median: 62 Jahre) zusammen. Die bioinformatische Auswertung erfolgte analog zu dem Vorgehen bei der Analyse des Mutationsprofils in folliculären Lymphomen (siehe unter Abschnitt 3.2.1.1). Insgesamt wurden 1036 somatische Mutationen in den 84 analysierten DLBCL-Patienten identifiziert, die als Basis-Datensatz für die primäre altersassoziierte Mutationsanalyse in DLBCL dienen. Die potentiell proteinverändernde Mutationen wurden in 95 Genen detektiert, wobei 53/1036 (6%) Mutationen die Splicing-Region der analysierten Transkripte betrafen. 811/1036 (64%) waren nicht-synonyme Mutationen, 75/1036 (15%) Frameshift Deletionen oder Insertionen, 78/1036 (14%) Stopp-Mutationen und 19/1036 (1%) nicht-Frameshift Deletionen oder Insertionen.

In der weiteren Mutationsanalyse wurden zuerst alle rekurrent mutierten Gene identifiziert. Zur Beurteilung der Häufigkeit von Mutationen in den rekurrent mutierten Genen wurde die Mutationsfrequenz (MF) ermittelt und grafisch dargestellt (Abbildung 34).

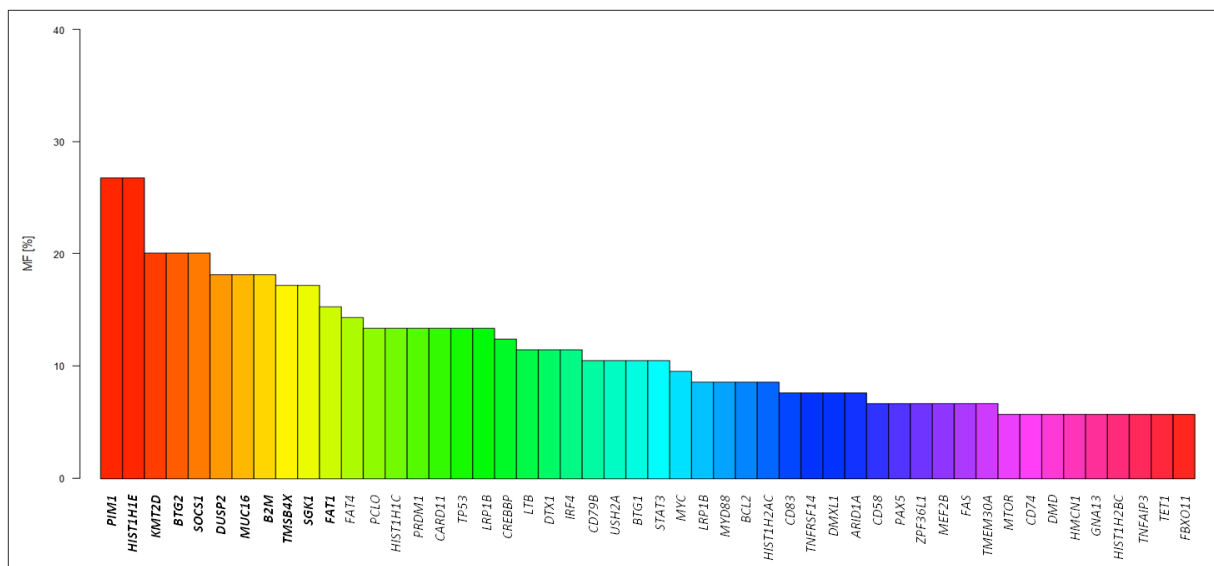


Abbildung 34: Mutationsfrequenz der rekurrent mutierten Gene in DLBCL

Um die Unterschiede in der Mutationsfrequenz (MF) der einzelnen Gene besser darzustellen, wurde in dem Balkendiagramm die Skalierung der y-Achse von 100% auf 40% reduziert. Als rekurrent mutiert wurden alle Gene bezeichnet, die in mindestens 5 Patienten potentiell proteinverändernde Mutationen aufwiesen. Insgesamt wurden in den 84 analysierten Patienten 50 Gene identifiziert, die häufig in DLBCL mutiert sind. Davon wurde für 11 Gene eine Mutationsfrequenz (MF) von über 15% ermittelt (fett markiert).

Innerhalb dieses Datensatzes konnten insgesamt 50 Gene als rekurrent mutiert in DLBCL identifiziert werden, wobei elf Gene eine MF von über 15% aufwies (*PIM1*, *HIST1H1E*, *KMT2D/MLL2*, *BTG2*, *SOCS1*, *DUSP2*, *MUC16*, *B2M*, *TMSB4X*, *SGK1*, *FAT1*). Diese zehn Gene werden im Folgenden als „Top 11 Gene“ bezeichnet. Die durchschnittliche MF für die Top 11 Gene in DLBCL lag bei 20%.

Anschließend wurden die Top 11 Gene in einer Pathway-Analyse mit dem webbasierten Tool String 10 analysiert und das Ergebnis in der Abbildung 35 dargestellt. Die Analyse zeigte, dass nur zwei der elf exprimierten Proteine miteinander interagierten (*PIM1* und *SOCS1*). Für alle weiteren neun Proteine wurden in der durchgeführten Analyse keine Interaktionen identifiziert.

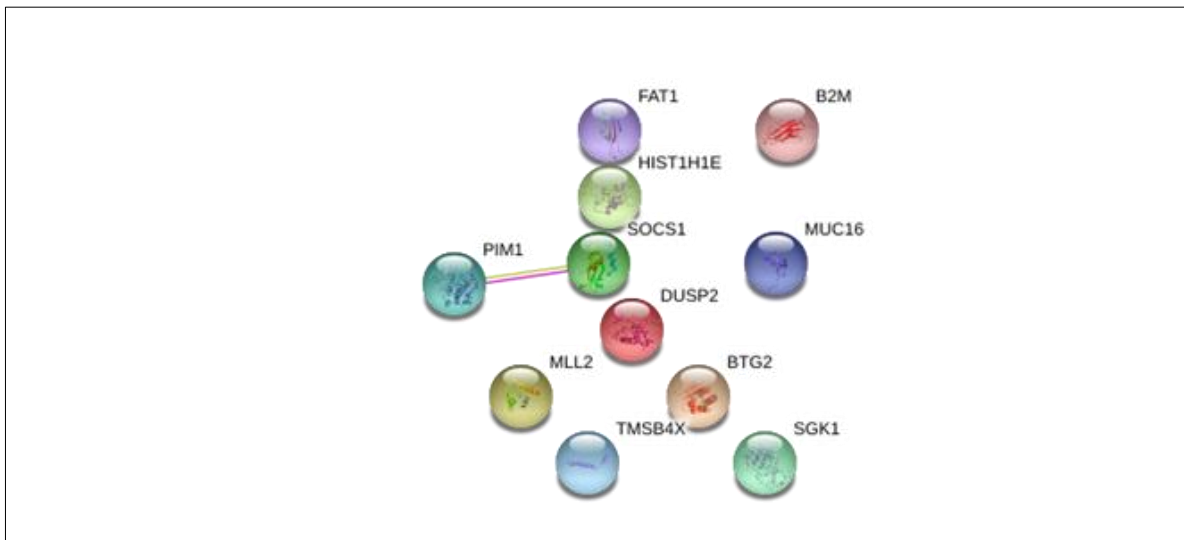


Abbildung 35: Stringanalyse der Top 11 Gene in DLBCL

Die Top 11 Gene in DLBCL wurden in dem webbasierten Tool String 10 analysiert. Die erzeugte Grafik zeigte, dass nur zwei der elf exprimierten Proteine miteinander interagieren (*PIM1*, *SOCS1*).

In der weiteren Mutationsanalyse wurden alle rekurrent mutierten Gene hinsichtlich des Ersterkrankungsalters der DLBCL-Patienten analysiert und als Überblick in Form eines Oncoprints visualisiert (siehe Abbildung A4 im Anhang). Im Oncoprint wurde für jedes Gen der identifizierte Mutationstyp in einem Farbcode dargestellt und zu den Patienten die identifizierten Translokationen aus den FISH-Analysen und die Gesamtanzahl an identifizierten Mutationen in den Target-Regionen des Lymphom-Panels dargestellt. Für die altersassoziierte Analyse wurde zusätzlich das Erstdiagnosealter der Patienten in vier Altersgruppen eingeteilt (0-39 Jahre, 40-59 Jahre, 60-79 Jahre und >80 Jahre) und im Oncoprint aufgetragen. Insgesamt zeigte die Betrachtung des Oncoprints keine eindeutige Assoziation zwischen einem rekurrent mutierten Gen oder der Anzahl an potenziell proteinverändernden Mutationen eines Patienten zum Ersterkrankungsalter des Patienten. Anschließend wurden die 50 rekurrent mutierten Gene in Form von Boxplots dargestellt und nach dem medianen Alter der entsprechend mutierten Patienten geordnet (Abbildung 36).

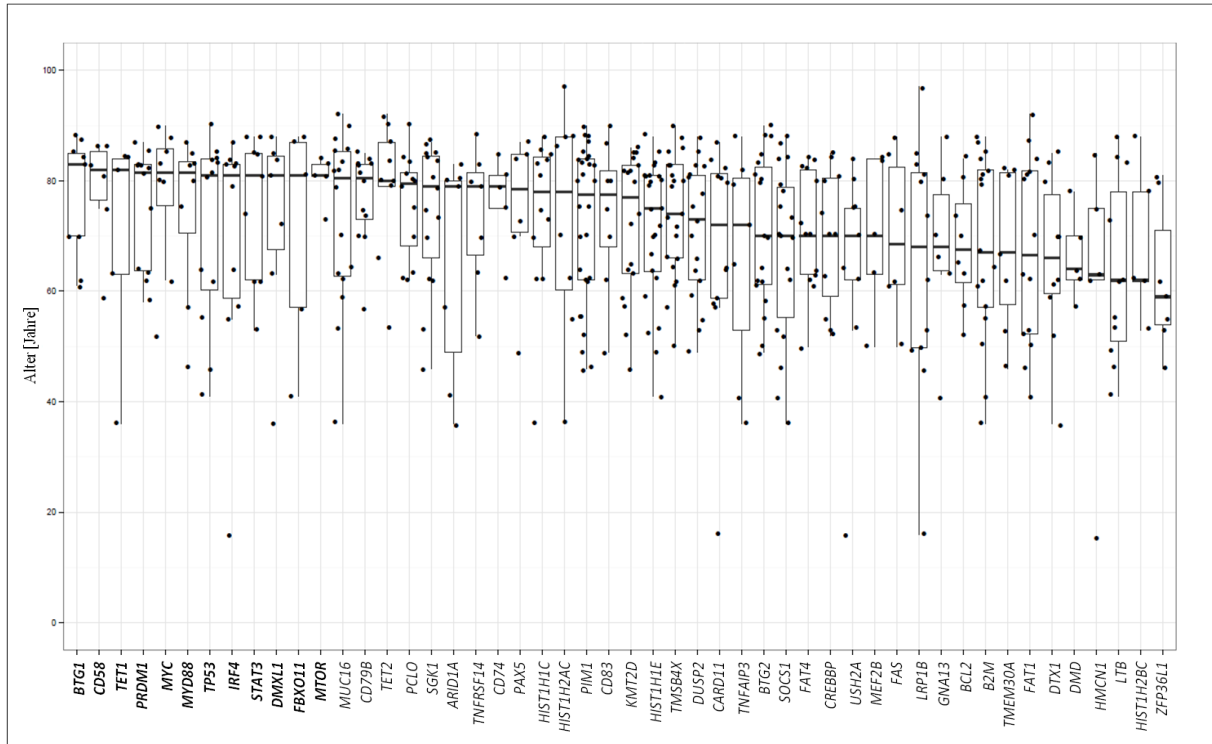


Abbildung 36: Altersverteilung der rekurrent mutierten Gene in DLBCL

In der Grafik wurden alle Gene aufgetragen, die in mind. 5 Tumor-Proben mit einer potentiell proteinverändernden Mutation identifiziert wurden. Insgesamt wurden in 84 analysierten DLBCL-Patienten 50 rekurrent mutierte Gene identifiziert. Die Gene mit dem höchsten bzw. zweithöchsten medianen Alter der betroffenen Patienten waren *BTG1* (Ø medianes Alter: 83 Jahre) bzw. *CD58*, *TET1*, *PRDM1*, *MYC* und *MYD88* (Ø medianes Alter: 82 Jahre) ermittelt. Insgesamt wurden 12 Gene mit einem medianen Alter der mutierten Patienten von über 80 Jahren identifiziert (fett markiert).

Das durchschnittliche mediane Alter der mutierten Patienten über die 50 Gene lag bei 75 Jahren. Die Gene mit dem höchsten Alter der betroffenen Patienten waren *BTG1* (Ø medianes Alter: 83 Jahre) und *CD58*, *TET1*, *PRDM1*, *MYC* sowie *MYD88* (jeweils Ø medianes Alter: 82 Jahre). Insgesamt wurden zwölf Gene identifiziert bei denen das mediane Alter der mutierten Patienten über 80 Jahren lag (*BTG1*, *CD58*, *MYC*, *MYD88*, *PRDM1*, *TET1*, *MTOR*, *DMXL1*, *STAT3*, *IRF4*, *TP53*, *FBXO11*). Diese 12 Gene werden im Folgenden als „Top 12 Gene“ bezeichnet. Die durchschnittliche Mutationsfrequenz (MF) der Top 12 Gene innerhalb der Gesamtkohorte variierte zwischen 6% (*TET1*, *MTOR*, *FBXO11*) und 14%, wobei die MF für die Gene *TP53*, *PRDM1* (jeweils 14%) am Höchsten waren.

Anschließend wurden die Top 12 Gene in einer Pathway-Analyse mit dem webbasierten Tool String 10 analysiert und das Ergebnis in der Abbildung 37 dargestellt.

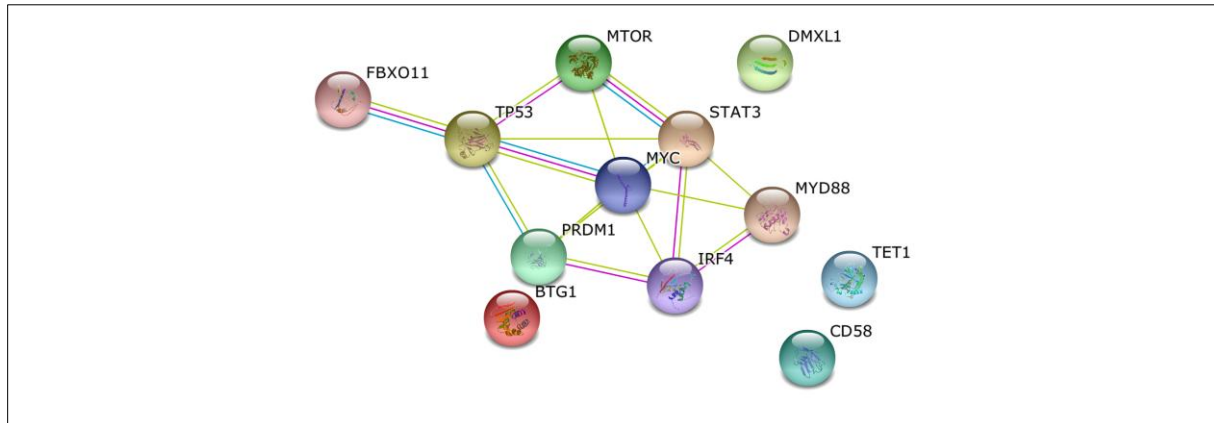


Abbildung 37: Stringanalyse der Top 12 Gene in senilen DLBCL

Die Top 12 Gene in DLBCL-Patienten wurden in dem webbasierten Tool String 12 analysiert. Die erzeugte Grafik zeigte, dass 8 der 12 exprimierten Gene miteinander interagieren.

Die durchgeführte Analyse zeigte, dass acht der zwölf exprimierten Proteine miteinander interagierten (*FBXO11*, *IRF4*, *MTOR*, *MYD88*, *MYC*, *PRDM1*, *STAT3*, *TP53*). Daraufhin wurde die Funktion der Top 12 Gene und Proteine entsprechend näher betrachtet und die wichtigsten Informationen in der nachfolgenden Tabelle 27 zusammengetragen. Die Ergebnisse zeigten, dass die identifizierten Top 12 Gene häufig Transkriptionsfaktoren sind überwiegend eine Rolle im Zellzyklus und der Immunantwort des menschlichen Organismus spielen.

Tabelle 27: Informationen zu den Top 12 Genen in senilen DLBCL

| Gen | Entrez Gene Name | Lokalisation | Typ | Funktion |
|---------------|----------------------------------------------------------------------------------|-----------------------|----------------------|----------------------------|
| <i>BTG1</i> | B-cell translocation gene 1, anti-proliferative | Zellkern | Transkriptionsfaktor | Zellzyklus |
| <i>CD58</i> | CD58 molecule | Plasmamembran | Transmembranrezeptor | Immunantwort |
| <i>DMXL1</i> | Dmx-like 1 | Extrazelluläre Matrix | Regulator | Regulatorische Funktion |
| <i>FBXO11</i> | F-box protein 11 | Cytoplasma | Enzym | Effektor von p53 Signalweg |
| <i>IRF4</i> | interferon regulatory factor 4 | Zellkern | Transkriptionsfaktor | Immunantwort |
| <i>MTOR</i> | mechanistic target of rapamycin (serine/threonine kinase) | Zellkern | Kinase | Zellantwort |
| <i>MYC</i> | v-myc avian myelocytomatosis viral oncogene homolog | Zellkern | Transkriptionsfaktor | Zellzyklus |
| <i>MYD88</i> | myeloid differentiation primary response 88 | Plasmamembran | Adapterprotein | Immunantwort |
| <i>PRDM1</i> | PR domain containing 1, with ZNF domain | Zellkern | Transkriptionsfaktor | Effektor von p53 |
| <i>STAT3</i> | signal transducer and activator of transcription 3 (acute-phase response factor) | Zellkern | Transkriptionsfaktor | Zellzyklus |
| <i>TET1</i> | tet methylcytosine dioxygenase 1 | Zellkern | Dioxygenase | DNA-Methylierung |
| <i>TP53</i> | tumor protein p53 | Zellkern | Transkriptionsfaktor | Tumorsuppressor |

Der Entrez Gene Name wurde nicht in die deutsche Sprache übersetzt und in ist Englischer Sprache entsprechend den Angaben des Ensemble Genome Browser Release 81, Stand 12/2015 aufgelistet.

Anschließend wurde ein Vergleich der rekurrent mutierten Gene in DLBCL unabhängig vom Ersterkrankungsalter der Patienten (Top 11 Gene, siehe Abbildung 35) mit den rekurrent mutierten Genen in senilen Patienten (Top 12 Gene, siehe Abbildung 37) vorgenommen. Die Ergebnisse zeigten, dass kein Gen in beiden analysierten Gen-Listen vorlag. Somit wurde kein Gen identifiziert, dass eine MF von über 15% aufwies und überwiegend in senilen DLBCL-Patienten rekurrent mutiert auftrat.

Daraufhin wurden die Aberrationen der Top 12 Gene in senilen DLBCL-Patienten detaillierter betrachtet. Hierzu wurden die drei Gene mit der höchsten MF ausgewählt (*TP53*, *PRDM1* und *IRF4*) und grafisch alle identifizierten Mutationen hinsichtlich der Mutationstypen und der Position im Protein dargestellt (Abbildung 38).

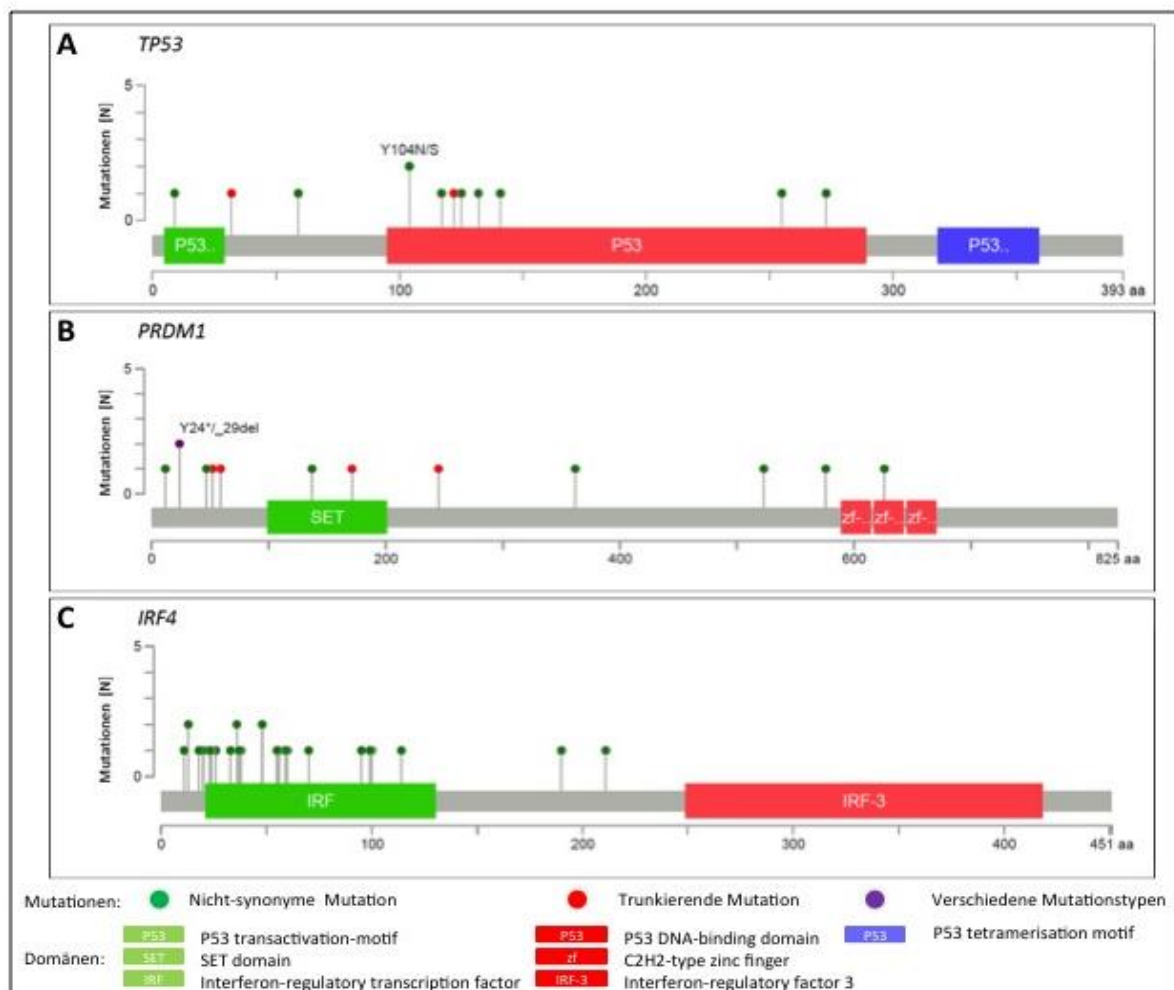


Abbildung 38: Schematische Darstellung der Mutationen in den Genen *TP53*, *PRDM1* und *IRF4* in DLBCL

Die Domänen-Namen wurden nicht in die deutsche Sprache übersetzt und sind in Englischer Sprache entsprechend den Angaben des webbasierten Tools MutationMapper übernommen. In der Grafik sind Positionen gekennzeichnet an denen zwei unterschiedliche Mutationen in der genomischen Sequenz identifiziert wurden. Die entsprechenden Mutationen sind darüber angegeben (z.B. *TP53*: Y104N/S). D.h. an dieser Position wurde die Mutation Y104N und Y104S in der analysierten Patienten-Kohorte detektiert. A: Im *TP53*-Gen wurden überwiegend nicht-synonyme Mutationen (10/12) identifiziert. Diese Mutationen lagen vorwiegend in der DNA-bindenden Domäne p53 (rot markiert). B: Die Mutationen in *PRDM1* verteilen sich über die kodierende Region des Gens, wobei 8/13 nicht-synonyme und 5/13 trunkierende Mutationen vorlagen. Nur 2/13 identifiziert Mutationen lagen in der SET-Domäne des Gens. C: Im *IRF4*-Gen wurden ausschließlich nicht-synonyme Mutation (26/26) detektiert, die vorwiegend in der IRF-Domäne des Gens lagen. Die detektierten Mutationen lagen insbesondere in der bekannten somatisch hypermutierten Region des *IRF4*-Gens.

Das Mutationsprofil im *TP53*-Gen der DLBCL-Patienten zeigte, dass die Aberrationen überwiegend nicht-synonyme Mutationen (10/12) waren, die in der DNA-bindenden Domäne des Proteins lagen (P53, rot markiert in Abbildung 38A). Zusätzlich wurden zwei trunkierende Mutationen in *TP53* detektiert. Im *PRDM1*-Gen wurden 8/13 nicht-synonyme Mutationen und 5/13 trunkierende Mutationen identifiziert, die sich über die komplette kodierende Region des Gens verteilten. Nur 2/13 Mutationen lagen in der SET-Domäne des PRDM1-Proteins (SET, grün markiert in Abbildung 38B). Im *IRF4*-Gen wurden ausschließlich nicht-synonyme Mutationen identifiziert (26/26). Die Mutationen verteilten sich auf 10 Patienten und lagen überwiegend in der IRF-Domäne, des Proteins (IRF, grün markiert in Abbildung 38C, N= 20/26). Genomisch liegen alle *IRF4*-Mutationen in der bekannten somatisch hypermutierten Region des Gens.

Nachfolgend wurde eine statistische Auswertung der 50 rekurrent mutierten Gene in Bezug auf das Ersterkrankungsalter der Patienten durchgeführt (Detaillierte Ergebnisse sind im Anhang in der Tabelle A5 zusammengestellt). Dazu wurde zuerst das statistische Verfahren der logistischen Regression (LR) angewandt. Die Ergebnisse zeigten, dass mittels LR in allen analysierten Genen keine signifikante Assoziation zwischen den Mutations-Ergebnissen und dem Alter der Patienten vorlag. Zusätzlich wurden alle rekurrent mutierten Gene mittels Mann-Whitney U Test analysiert. Hierzu wurden die 50 rekurrent mutierten Gene hinsichtlich des medianen Alters der mutierten Patienten in zwei Gruppen eingeteilt: Ersterkrankungsalter ≥ 80 Jahre und Ersterkrankungsalter < 80 Jahre. Anschließend wurde für jedes rekurrent mutierte Gen der P-Wert mittels U Test bestimmt. Die Ergebnisse der statistischen Auswertung der Top 12 Gene sind in der Tabelle 28 zusammengestellt.

Tabelle 28: Rekurrent mutierte Gene in senilen DLBCL-Patienten (Top 12 Gene)

| Gen | min. Alter [Jahre] | Medianes Alter [Jahre] | max. Alter [Jahre] | Interquartilabstand [Jahre] | LR P-Wert | U Test P-Wert | MF [%] |
|---------------------|--------------------|------------------------|--------------------|-----------------------------|-------------|---------------|-----------|
| <i>BTG1</i> | 70 | 83 | 85 | 15 | 0,25 | 0,25 | 11 |
| <i>CD58</i> | 76 | 82 | 85 | 9 | 0,28 | 0,24 | 7 |
| <i>MYC</i> | 76 | 82 | 86 | 10 | 0,08 | 0,22 | 10 |
| <i>MYD88</i> | 70 | 82 | 85 | 15 | 0,31 | 0,48 | 9 |
| <i>PRDM1</i> | 64 | 82 | 83 | 19 | 0,33 | 0,27 | 14 |
| <i>TET1</i> | 63 | 82 | 85 | 22 | 0,93 | 0,50 | 6 |
| <i>MTOR</i> | 81 | 81 | 83 | 2 | 0,15 | 0,12 | 6 |
| <i>DMXL1</i> | 68 | 81 | 85 | 17 | 0,52 | 0,71 | 8 |
| <i>STAT3</i> | 62 | 81 | 85 | 23 | 0,52 | 0,36 | 11 |
| <i>IRF4</i> | 59 | 81 | 83 | 24 | 0,75 | 0,79 | 12 |
| <i>TP53</i> | 60 | 81 | 85 | 25 | 0,33 | 0,85 | 14 |
| <i>FBXO11</i> | 57 | 81 | 87 | 30 | 0,50 | 0,96 | 6 |

Abkürzungen und Verweise: min. Alter: minimales Alter der entsprechend mutierten Patienten innerhalb des Boxplots (untere Quartilgrenze in Abbildung 37). max. Alter: maximales Alter der entsprechend mutierten Patienten innerhalb des Boxplots (obere Quartilgrenze in Abbildung 37). Medianes Alter: medianes Alter der entsprechend mutierten Patienten. LR: Logistische Regression. U Test: Mann-Whitney U Test. MF: Mutationsfrequenz innerhalb der analysierten Patienten-Kohorte. Zur besseren Übersicht wurden das Kandidatengen *CD58* und die drei Gene mit der höchsten MF (*TP53*, *PRDM1* und *IRF4*) fett markiert.

Die ermittelten P-Werte zeigten, dass keine signifikante Assoziation zwischen den Mutations-Ergebnissen und dem Alter der Patienten berechnet wurde. Da der Fokus der Arbeit auf Aberrationen in Patienten mit einem sehr hohen Alter lag, wurde neben dem medianen Alter der mutierten Patientengruppe auch der Interquartilabstand einbezogen. Dieser Abstand beschreibt die Differenz des max. und min. Ersterkrankungsalters der mutierten Patienten in Bezug auf die obere und untere Interquartilgrenze der ermittelten Boxplots (Abbildung 36). Unter Berücksichtigung des Interquartilabstands (<10) und einem medianen Alter der mutierten Patienten (>80 Jahre), wurde nur ein Gen (*CD58*) als Top Gen für Mutationen bei senilen Patienten identifiziert (Tabelle 28). Eine detaillierte Betrachtung des Kandidatengens *CD58* ist in dem nachfolgenden Abschnitt 3.2.1.2.2 aufgeführt.

Insgesamt konnten in den Mutationsanalysen dieser Arbeit 50 rekurrent mutierte Genen in DLBCL identifiziert werden, wovon elf Gene mit einer MF von über 15% in DLBCL nachgewiesen wurden (Top 11 Gene). Zwölf Gene konnten ermittelt werden, die vorwiegend bei senilen Patienten rekurrent mutierten auftraten (Top 12 Gene). Es wurde kein Gen identifiziert, das mit einer MF von über 15% vorwiegend in senilen DLBCL-Patienten mutiert vorliegt. Jedoch wiesen die Top 12 Gene alle eine MF von über 10% auf, wobei die höchste MF mit 14% bzw. 12% für *TP53*, *PRDM1* und *IRF4* ermittelt wurde, weshalb diese Gene als Kandidatengene für genetische Aberrationen in senilen Patienten mit einem DLBCL ausgewählt wurden. Wie bereits in diesem Kapitel erwähnt (siehe Tabelle 28), wurde zusätzlich *CD58* als Kandidatengen für genetische Aberrationen in senilen DLBCL-Patienten identifiziert und wird im Folgenden näher analysiert.

3.2.1.2.2 Kandidatengen *CD58*

Zeitgleich zu den altersassoziierten Analysen von genetischen Aberrationen in der vorliegenden Arbeit wurden im Rahmen des ICGC MMML-Seq Projektes Analysen zu strukturellen Varianten (SV) in den Keimzentrums-B-Zell-Lymphomen durchgeführt (N= 46, Median: 62 Jahre). Innerhalb dieser strukturellen Analysen wurde *CD58* als Kandidatengen mit rekurrent vorkommenden Deletionen in DLBCL-Patienten identifiziert. Die *CD58*-Analysen des ICGC MMML-Seq Projektes wurden von Frau Dr. Cristina Lopez Gonzalez durchgeführt, die zusätzlich die strukturellen Aberrationen von *CD58* mittels FISH-Analysen in einer unabhängigen Validierungskohorte von Patienten aus dem Untersuchungskollektiv des Instituts für Humangenetik validierte (N= 41, Median: 67 Jahre). Aufgrund der Erkenntnisse, dass offensichtlich auch strukturelle Aberrationen im *CD58*-Gen bei senilen DLBCL-Patienten auftreten, sollten in der *CD58*-Analyse dieser Arbeit sowohl Mutations-Ergebnisse als auch strukturelle Varianten innerhalb des Gens genauer betrachtet werden.

Dazu lagen die Sequenz-Daten aus den Mutationsanalysen mittels Lymphom-Panel von 17 senilen Patienten der Kohorte *Jack* (Median: 84 Jahre) vor. Nachdem in den Mutationsanalysen der vorliegenden Arbeit SNVs (engl. single nucleotide variant) im *CD58*-Gen bei senilen DLBCL-Patienten identifiziert wurden und im ICGC MMML-Seq Projekte strukturelle Aberrationen in *CD58* detektiert wurden, wurden zusätzliche Daten zu strukturellen Varianten im Analysen *CD58*-Gen bei senilen Patienten in die weiteren Analysen integriert.

Dazu lagen aus dem MMML-Verbundprojekt die SNV- und/oder SV-Daten von 239 Patienten vor (Median: 64 Jahre). Davon existierten von 232 Patienten (Median: 64 Jahre) die SNP-Array-Daten und zusätzlich lagen von 17 Patienten (Median: 84 Jahre) des MMML-Verbundprojektes die Sequenzdaten vor, da diese 17 Patienten der Kohorte *MMML-Senil* angehörten und im Rahmen der vorliegenden Arbeit mittels Lymphom-Panel analysiert wurden. Insgesamt lagen von 10 Patienten des MMML-Verbundprojektes sowohl Sequenz- als auch SNP-Array-Daten vor (Median: 83 Jahre).

Weiterhin wurden die FFPE-Tumorschnitte von 55 senilen DLBCL-Patienten der Kohorte *Jack* (Median: 83 Jahre) mit den gleichen *CD58*-Sonden analysiert, die zuvor im Rahmen des ICGC MMML-Seq Projektes zur Validierung der strukturellen Veränderungen in DLBCL verwendet wurden. Die Durchführung der FISH-Analysen bei den Patienten der *Jack*-Kohorte erfolgte durch Frau Dr. Cristina Lopez Gonzalez und Frau Reina Zühlke-Jenisch am Institut für Humangenetik. Die Ergebnisse der FISH-Analysen zeigten, dass die Daten von zehn Patienten nicht auswertbar waren und somit nur 45/55 *CD58*-FISH-Daten in die weitere Analyse eingingen.

Zur Auswertung von genetischen Aberrationen in DLBCL in dem *CD58*-Gen wurden insgesamt 373 Patienten (Median: 67 Jahre) aus den Kohorten ICGC, MMML, der Validierungskohorte und *Jack* genauer betrachtet. Eine Übersicht der analysierten Patienten in Bezug auf die Altersverteilung der Patienten in den einzelnen Kohorten und den entsprechend analysierten Daten ist in der nachfolgenden Abbildung 39 zusammengestellt.

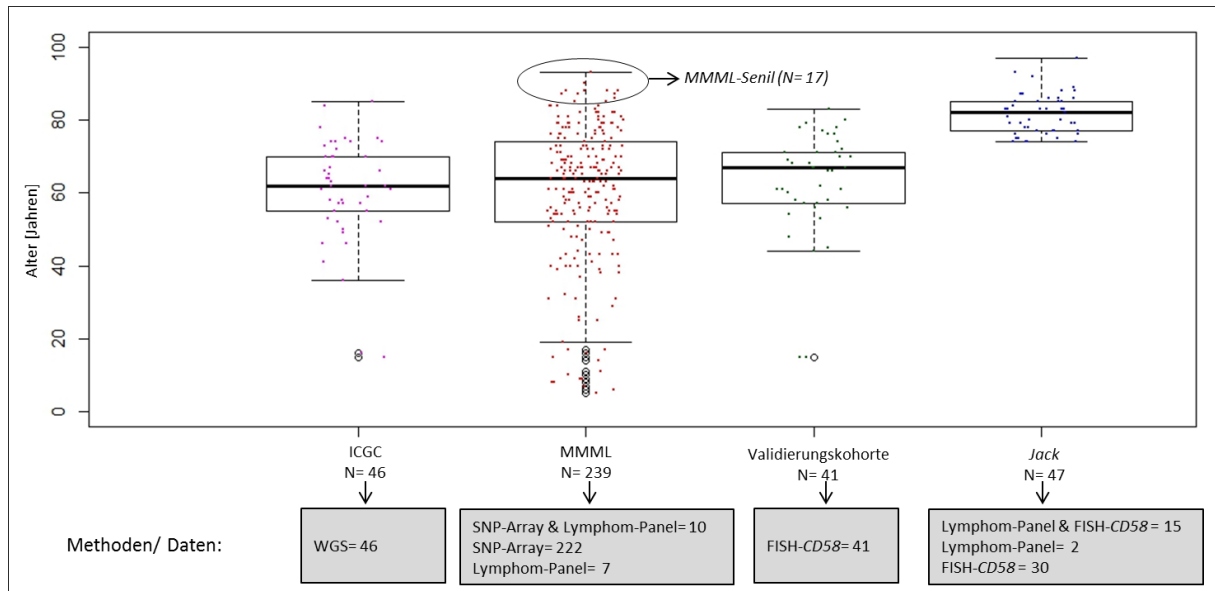


Abbildung 39: Altersverteilung der Patienten-Kohorte und analysierten molekulargenetischen Daten zur altersassoziierten Analyse der genetischen Aberrationen im Kandidatengen *CD58*

Die Betrachtung der analysierten Kohorte zeigte, dass innerhalb der Analyse des *CD58*-Gens DLBCL-Patienten im breiten Altersspektrum von 5-97 Jahren analysiert wurden (N=343, Median: 65 Jahre). Zur Analyse der genetischen Aberrationen in *CD58* wurden die WGS-Daten von 46 Patienten aus dem ICGC MMML-Seq Projekt, die SNP-Array-Daten von 232 Patienten aus dem MMML-Verbundprojekt, die Lymphom-Panel-Daten von 17 Patienten aus der Kohorte *MMML-Senil*, die FISH-Daten von *CD58* von 41 Patienten der Validierungskohorte sowie die Lymphom-Panel-Daten von 17 Patienten und die FISH-Daten von *CD58* von 55 Patienten der Kohorte *Jack* verwendet. Abkürzungen: WGS: (engl. whole genome sequencing) Ganz-Genom-Sequenzierung. MMML: Verbundprojekt "Molekulare Mechanismen bei malignen Lymphome". ICGC: ICGC MMML-Seq, ein Verbundprojekt des internationalen Krebsgenomkonsortiums. *Jack*: Kohorte von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphome, die im Rahmen der vorliegenden Arbeit neu identifiziert wurden. *MMML-Senil*: Kohorte von senilen Patienten aus dem MMML-Verbundprojekt.

In der weiteren Analyse wurden zuerst die SNV- und SV-Daten der Gesamtkohorte (N=373) unabhängig vom Ersterkrankungsalter der Patienten ausgewertet. Die Ergebnisse zeigten, dass in 39/373 Patienten (~10%, Median: 74 Jahre) eine genetische Aberration in *CD58* identifiziert wurde (ICGC: 8/46, MMML: 18/239, Validierungskohorte: 6/41 und *Jack*: 7/47). Eine detaillierte Auflistung der genetischen Aberrationen im *CD58*-Gen (1p13) in der analysierten Gesamtkohorte ist in der Tabelle 29 zusammengestellt. Die genaue Betrachtung der genetischen Aberrationen in *CD58* zeigte, dass in 33/373 der Patienten nur eine genetische Aberration (SNV: 3/373 und SV: 30/373) identifiziert wurde. Bei 6/373 Patienten (PID: 4188398, 4131257, 4145528, 4134434 *Jack*-39, MPI-756) wurden mehr als eine, aber max. drei genetische Aberrationen in *CD58* nachgewiesen.

Tabelle 29: Untersuchungskollektiv für die altersassoziierte Auswertung der genetischen Aberrationen im *CD58*-Gen

| Kohorte | Probe | Alter | Start [bp]-Ende [bp] | Nukleotid- | Protein | Funktion | Typ |
|----------------------|--------------|------------------------------------|-----------------------------------|-------------------|-------------|----------|-------|
| ICGC | 4107137 | 59 | 117086941 | Ex2: c.352-355del | p.Leu118fs* | FSD | Indel |
| | 4188398 | 75 | 117078668 | Ex3: c.C547T | p.Gln183fs* | Stop | SNV |
| | | | Chr9 75818305 - Chr1 117109275 | | | T | SV |
| | 4131257 | 72 | 117112256 - 155725148 | | | D | SV |
| | | | 76466660 - 119235528 | | | D | SV |
| | 4134434 | 84 | 117077979 - 117104493 | | | D | SV |
| | | | 72883609 - 197158206 | | | Dup | SV |
| | | | 106404081 - 167897917 | | | Dup | SV |
| | 4145528 | 62 | 90061732 - 161744240 | | | D | SV |
| | | | 117061041 - 117065855 | | | D | SV |
| 4163639 | 75 | Chr3 147056731 - Chr1:117088523 | | | T | SV | |
| 4135099 | 49 | 115924522 - 117604545 | | | D | SV | |
| 4176133 | 61 | 115799139 - 120282609 | | | D | SV | |
| MMML | Jack-MPI-243 | 86 | 117113593 | Ex1: c.T2C | p.Met1Thr | NS | SNV |
| | Jack-MPI-756 | 81 | 117113585 | Ex1: c.G10A | p.Gly4Arg | NS | SNV |
| | | | 117113592 | Ex1: c.G3T | p.Met1Ile | | Indel |
| | MPI-014 | 76 | 75174093 - 117317605 | | | D | SV |
| | MPI-045 | 66 | 112771847 - 142548088 | | | I | SV |
| | MPI-106 | 70 | 20352768 - 119571017 | | | D | SV |
| | MPI-190 | 67 | 26904865 - 142548088 | | | D | SV |
| | MPI-304 | 50 | 86250745 - 119571017 | | | D | SV |
| | MPI-418 | 88 | 110807795 - 243539815 | | | I | SV |
| | MPI-022 | 74 | 69815163 - 119571017 | | | D | SV |
| | MPI-109 | 42 | 62070630 - 119571017 | | | D | SV |
| | MPI-142 | 60 | 112771847 - 117317605 | | | I | SV |
| | MPI-355 | 78 | 95516921 - 119571017 | | | D | SV |
| | MPI-393 | 51 | 46823998 - 119571017 | | | D | SV |
| | MPI-399 | 72 | 109362443 - 211553226 | | | I | SV |
| | MPI-205 | 82 | 1036213 - 243539815 | | | I | SV |
| MPI-389 | 65 | 100644730 - 154333751 | | | D | SV | |
| MPI-403 | 57 | 114919565 - 118596361 | | | I | SV | |
| MPI-412 | 58 | 66616707 - 119571017 | | | I | SV | |
| Validierungs-kohorte | 1060-8 | 74 | | | | D | SV |
| | 1206-12 | 77 | | | | D | SV |
| | 16-8 | 71 | | | | D | SV |
| | 1684-7 | 76 | | | | D | SV |
| | 531-6 | 56 | | | | D | SV |
| | 58-9 | 83 | | | | D | SV |
| Jack | Jack-39 | 83 | 117087122 | Ex2: c.173dupA | p.Lys58fs* | FSI | Indel |
| | | | 117087140 | Ex2: c.G157T | p.Glu53fs* | Stop | SNV |
| | Jack-47 | 86 | 117087008 | Ex2: c.286-288del | p.Thr96del | FSD | Indel |
| | Jack-31 | 83 | | | | I | SV |
| | Jack-34 | 85 | | | | M | SV |
| | Jack-36 | 97 | | | | D | SV |
| | Jack-134 | 79 | | | | D | SV |
| | Jack-22 | 85 | | | | D | SV |

Abkürzungen und Verweise: Kohorte: Kohorte/ Studie aus dem der analysierte Patient stammt. Alter: Ersterkrankungsalter in Jahren. Start/ Ende: Genomische Start-/Endposition der DNA-Sequenz an der die genetische Aberration identifiziert wurde. MAF: Mutationsallelfrequenz. Nukleotid: Nukleotidveränderung nach HGVS. Protein: Proteinveränderung nach HGVS. I: Insertion. D: Deletion. Stop: Stopmutation. T: Translokation. SV: strukturelle Veränderung. FSD: Frameshift Deletion. FSI: Frameshift Insertion. NS: Nicht-synonyme Mutation. SNV: (engl. single nucleotide variant) Punktmutation. Indel: Insertion oder Deletion. Chr: Chromosom. Ex: Exon. Jack-MPI: Fälle senile Patienten aus der Kohorte *MMML-Senil*.

Eine genaue Betrachtung der detektierten Punktmutationen (SNV) im *CD58*-Gen zeigte, dass 4/7 trunkierende Mutationen, 3/7 nicht-synonyme Mutationen und eine In-Frameshift-Deletion vorlagen. 3/7 Mutationen lagen in der CD2-bindenden Domäne des *CD58*-Gens (Abbildung 40). CD2 ist ein Adhäsionsmolekül und wurde als einer der ersten T-Zell-Marker identifiziert. Das Protein wird von über 95% aller Thymozyten, einigen natürlichen Killerzellen, aber nicht von B-Zellen exprimiert²⁵². Die CD2-Domäne interagiert mit dem CD2-Glycoprotein, welches wiederum für Antigen-unabhängige und -abhängige Interaktionen von T-Zellen mit deren Zielmolekülen verantwortlich ist^{252,253}.

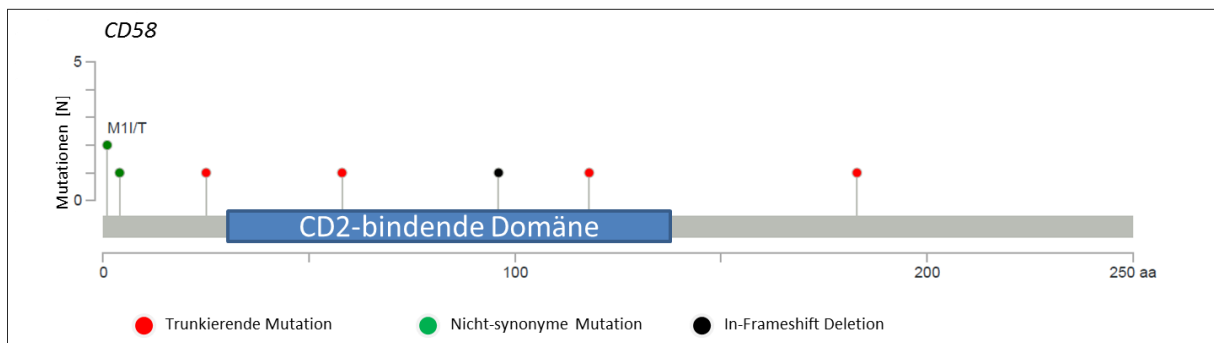


Abbildung 40: Schematische Darstellung der Punktmutationen im *CD58*-Gen bei DLBCL-Patienten

In der Grafik sind Positionen gekennzeichnet an denen zwei unterschiedliche Mutationen in der genomischen Sequenz identifiziert wurden. Die entsprechenden Mutationen sind darüber angegeben (z.B. M11/T). D.h. an dieser Position wurde die Mutation M1I und M1T in der analysierten Patienten-Kohorte detektiert. *CD58*: Im *CD58*-Gen wurden 4/7 trunkierende und 3/7 nicht-synonyme Mutationen identifiziert. Davon lagen drei Mutationen in der CD2-bindenden Domäne (blau markiert).

Abschließend wurde eine statistische Auswertung der Gesamtkohorte (N=373, Median: 67 Jahre) hinsichtlich der Altersassoziation von genetischen Aberrationen im *CD58*-Gen vorgenommen. Hierzu wurde das Verfahren der logistischen Regression (LR, Abbildung 41A) und der Kruskal-Wallis Test (KW, Abbildung 41B) angewandt. Es wurden in beiden Verfahren signifikante P-Werte für das Erstdiagnosealter der betroffenen (mutierten) Patienten im Vergleich zu den nicht betroffenen (unmutierten) Patienten ermittelt (LR: P-Wert= 0,0035 und KW: P-Wert= 0,0035).

Es ist zu beachten, dass nicht von allen Patienten sowohl SNV- als auch SV-Daten für das *CD58*-Gen vorlagen und demnach ggf. falsch negative Ergebnisse in den Analysen enthalten sein. D.h. möglicherweise haben noch mehr DLBCL-Patienten eine genetische Aberration in *CD58*, die in den durchgeführten Analysen dieser Arbeit als negativ (Wildtyp) gewertet wurden.

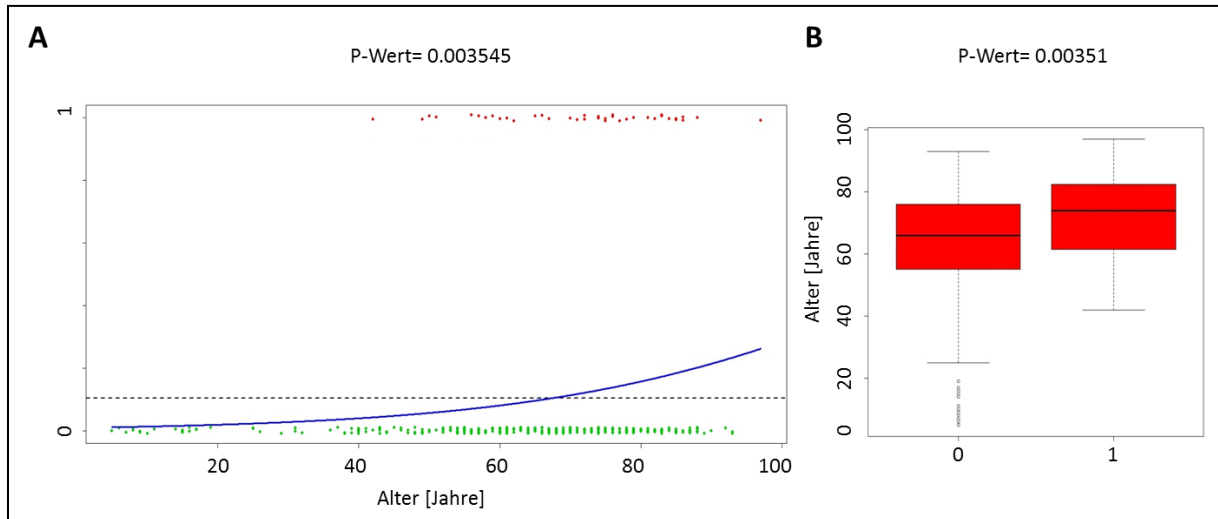


Abbildung 41: Statistische Auswertung von altersassoziierten genetischen Aberrationen im CD58-Gen

A: In der Darstellung der logistischen Regression stellt jeder Punkt einen analysierten Patienten dar. Eine Aberration wurde in Form von roten Punkten in der Grafik eingefügt. Keine Aberration/ Wildtyp ist durch grüne Punkte gekennzeichnet. Die gestrichelte Linie stellt das mediane Alter des Patienten-Kollektives dar. Auf der x-Achse ist das Ersterkrankungsalter der Patienten mit zunehmendem Alter aufgetragen. Die Logistische Regression der Ergebnisse der SNV- und SV-Analyse im *CD58*-Gen im Zusammenhang mit dem Ersterkrankungsalter von DLBCL-Patienten (N= 373, Median: 67 Jahre) zeigte eine signifikante Assoziation von genetischen Aberrationen in *CD58* bei Patienten im hohen Alter (P: 0,0035). B: In der Darstellung des Kruskal-Wallis Tests wurden auf der x-Achse die Werte 0 und 1 dargestellt, wobei 0 für negativ (Wildtyp) und 1 für positiv steht. Der Kruskal-Wallis Test der Ergebnisse der SNV- und SV-Analyse im *CD58*-Gen im Zusammenhang mit dem Ersterkrankungsalter von DLBCL-Patienten (N= 373, Median: 67 Jahre) zeigte eine signifikante Assoziation von genetischen Aberrationen in *CD58* bei Patienten im hohen Alter (P: 0,0035).

Zusammenfassend konnte in der vorliegenden Arbeit das *CD58*-Gen als signifikant altersassoziiertes Marker für genetische Aberrationen bei DLBCL-Patienten mit einem hohen Ersterkrankungsalter identifiziert werden (Medianes Alter der betroffenen Patienten: 74 Jahre).

3.2.1.2.3 Assoziation und Zusammenfassung der altersassoziierten genetischen Marker in Patienten mit einem diffus großzelligen B-Zell-Lymphom

Inwieweit die verschiedenen molekulargenetischen Marker, die in den bisherigen Kapiteln zur Basis-Charakterisierung und Mutationsanalyse analysiert wurden, miteinander assoziiert vorlagen, sollte anschließend überprüft werden. Als Einschlusskriterien für die anschließende Assoziationsanalyse der molekulargenetischen Daten der DLBCL-Patienten wurde das Vorhandensein von Basis-Charakterisierungs- und Hochdurchsatz-Sequenzierungs-Daten gewählt. Von den 1609 Patienten mit einem DLBCL, die im Rahmen der vorliegenden Arbeit analysiert wurden, lagen zu 84 DLBCL-Patienten im Alter von 15 bis 97 Jahren (75 Jahre) die Daten der Basis-Charakterisierung und Hochdurchsatz-Sequenzierung vor. Von 1525 Patienten waren ausschließlich die Basis-Charakterisierungs-Daten vorhanden. Zur Betrachtung der Assoziation zwischen den genetischen Markern in DLBCL wurden die bisherigen Ergebnisse aus der Basis-Charakterisierung und der Mutationsanalyse der Lymphom-Panel-Target-Regionen in der analysierten Kohorte zusammengestellt und miteinander verglichen. Dazu wurden die analysierten DLBCL-Patienten nach dem Erstdiagnosealter (0-39 Jahre, 40-59 Jahre, 60-79 Jahre und >80 Jahre) in vier Altersgruppen eingeteilt. Anschließend wurde ein Überblick über die im Rahmen dieser Arbeit erhobenen Ergebnisse und in Form einer Heatmap erstellt (Abbildung 42). Im Detail wurden die Altersgruppen, FISH-Ergebnisse (*BCL2*, *BCL6*, *MYC*, *IGH*), IHC-Ergebnisse (*BCL6*, *BCL2*, Ki67, MUM1, CD10), Mutations-Ergebnisse von bereits bekannten rekurrent mutierten Genen in DLBCL (*MYD88*, *EP300*, *CCND3*, *EZH2*, *PRDM1*, *TP53*, *CARD11*, *CD79B*, *CREBBP*, *BCL2*, *SOCS1*, *PIM1*) und die DLBCL-Subtypen (*Hans-Klassifikator*: GCB, nonGCB; *GEP-Klassifikator*: GCB, ABC) in einem Farbcode dargestellt und eine hierarchische Clusteranalyse der analysierten Patienten mit dem Statistikprogramm R vorgenommen.

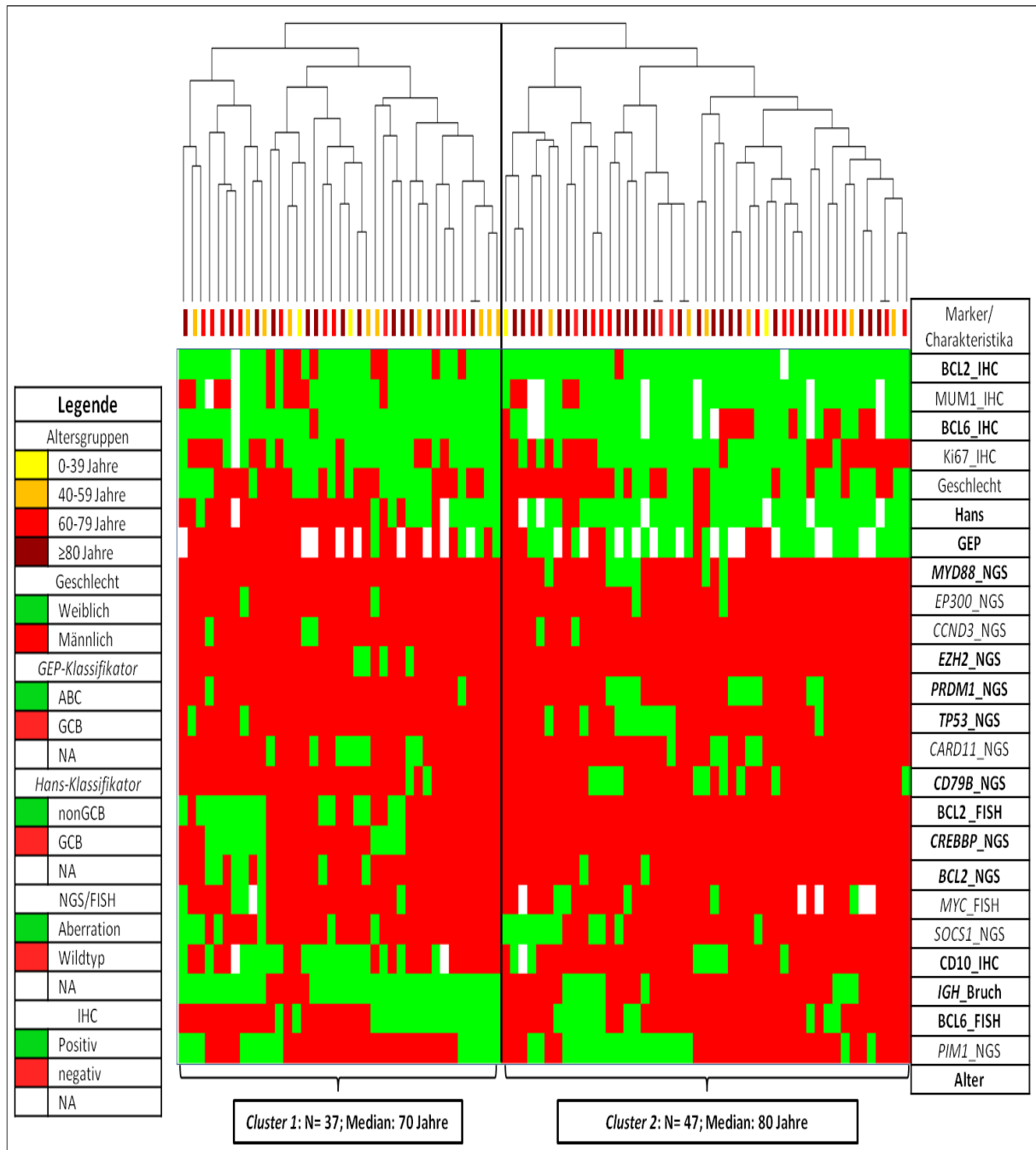


Abbildung 42: Assoziationen der analysierten genetischen Marker bei DLBCL-Patienten

Insgesamt wurden 84 DLBCL-Patienten im Alter von 15 bis 97 Jahren (Median: 75 Jahre) mittels Basis-Charakterisierung und Hochdurchsatz-Sequenzierung analysiert. In einer hierarchischen Clusteranalyse der Patienten wurden 2 Gruppen identifiziert (Cluster 1 und Cluster 2). In der statistischen Auswertung mittels Wilcoxon Test wurden die analysierten Marker zwischen den Patienten aus Cluster 1 mit denen aus Cluster 2 verglichen. Dabei wurde gezeigt, dass die Patienten vom Cluster 1 vorwiegend vom DLBCL-Subtyp GCB waren und signifikant häufiger Mutationen in *EZH2*, *CREBBP* und *BCL2* aufwiesen. Weiterhin hatten Patienten, die zum Cluster 1 gehörten, signifikant häufiger *BCL2*-, *BCL6*- und *IGH*- Brüche und die Genexpression der Proteine *CD10* und *BCL6* war signifikant höher. Patienten vom Cluster 2 gehörten eher dem DLBCL-Subtyp ABC (*GEP-Klassifikator*) bzw. nonGCB (*Hans-Klassifikator*) an und wiesen signifikant häufiger Mutationen in *MYD88*, *CD79B*, *PRDM1* und *TP53* auf. Zusätzlich war die Genexpression des Proteins *BCL2* signifikant höher bei Patienten, die zum Cluster 2 gehörten. Insgesamt hatten die Patienten vom Cluster 2 ein signifikant höheres Ersterkrankungsalter als die Patienten vom Cluster 1. Abkürzungen: FISH: Fluoreszenz-in-situ-Hybridisierung. IHC: Immunhistochemie. NGS: (engl. Next generation sequencing) Hochdurchsatz-Sequenzierung. GCB: (engl. Germinal center B-cell). Keimzentrumsähnliche B-Zellen). nonGCB: (engl. Non-Germinal center B cell) Aktivierte B-Zellen. ABC: (engl. Activated b-cell) Aktivierte B-Zelle. NA: nicht verfügbar bzw. nicht analysiert. Hans: *Hans-Klassifikator*. GEP: *GEP-Klassifikator*.

Tabelle 30: Analyierte genetischen Marker in Assoziation zu den identifizierten Clustern 1 und 2

| Marker | Gen/Protein | Veränderung Cluster 1 mut[N]/ wt[N], mut[%] | Veränderung Cluster 2 mut[N]/ wt[N], mut[%] | P-Wert |
|---------------------------|---------------|------------------------------------------------|------------------------------------------------|--------|
| Brüche | <i>BCL2</i> | 15/37, 41% | 0/47, 0% | 0,03 |
| | <i>BCL6</i> | 17/37, 46% | 11/47, 23% | 0,03 |
| | <i>IGH</i> | 32/37, 86% | 9/47, 19% | <0,01 |
| Proteinexpression | <i>BCL6</i> | 35/36, 97% | 9/47, 19% | 0,01 |
| | <i>CD10</i> | 18/35, 51% | 7/45, 16% | <0,01 |
| | <i>BCL2</i> | 28/36, 78% | 45/46, 98% | <0,01 |
| <i>Hans-Klassifikator</i> | GCB | 23/35, 66% | 6/41, 15% | <0,01 |
| | nonGCB | 12/35, 34% | 35/41, 85% | <0,01 |
| <i>GEP-Klassifikator</i> | GCB | 24/28, 86% | 10/33, 30% | <0,01 |
| | ACB | 4/28, 14% | 23/33, 70% | <0,01 |
| Mutationen | <i>EZH2</i> | 4/37, 11% | 0/47, 0% | 0,02 |
| | <i>CREBBP</i> | 11/37, 30% | 0/47, 0% | <0,01 |
| | <i>BCL2</i> | 7/37, 19% | 2/47, 4% | 0,03 |
| | <i>MYD88</i> | 0/37, 0% | 7/47, 15% | <0,01 |
| | <i>CD79B</i> | 2/37, 5% | 10/47, 21% | 0,02 |
| | <i>TP53</i> | 2/37, 5% | 10/47, 21% | 0,04 |

Abkürzungen: mut: mutiert, wt: unmutiert (Wildtyp). N: Anzahl. In der Tabelle sind nur Marker aufgelistet, die als signifikant verändert zwischen Cluster 1 und 2 identifiziert wurden.

In der Analyse wurden zwei große Cluster identifiziert, die nachfolgend als *Cluster 1* (N= 37, Median: 70 Jahre) und *Cluster 2* (N= 47, Median: 80 Jahre) bezeichnet werden. Zur weiteren Analyse der Assoziationen wurde ein Vergleich der genetischen Marker der Patienten aus dem *Cluster 1* mit den genetischen Markern der Patienten aus dem *Cluster 2* durchgeführt (siehe Tabelle 30). Dazu wurde eine statistische Auswertung vorgenommen und der P-Wert mittels Mann-Whitney U Test für jeden analysierten Marker berechnet.

Eine detaillierte Auswertung von *Cluster 1* zeigte, dass sich diese Gruppe vorwiegend aus Patienten mit einem DLBCL vom Subtyp GCB zusammensetzte. Diese Patienten hatten im Vergleich zu den Patienten vom *Cluster 2* signifikant häufiger Mutationen in den Genen *EZH2*, *CREBBP* und *BCL2*. Zusätzlich wurden bei den DLBCL-Patienten vom *Cluster 1* signifikant häufiger chromosomale Imbalancen in den Genen *BCL2*, *BCL6* und *IGH* ermittelt und die Genexpression der Proteine *BCL6* und *CD10* war signifikant höher als bei Patienten vom *Cluster 2*.

Das *Cluster 2* wurde eher von DLBCL-Patienten vom Subtyp nonGCB bzw. ABC gebildet. Diese Patienten hatten im Vergleich zu den Patienten vom *Cluster 1* signifikant häufiger Mutationen in den Genen *MYD88*, *CD79B*, *PRDM1* und *TP53*. Zusätzlich war die Genexpression des Proteins *BCL2* bei den DLBCL-Patienten vom *Cluster 2* signifikant höher als bei Patienten vom *Cluster 1*. Zusätzlich wurde das Alter als signifikanter Marker zwischen dem *Cluster 1* und *2* identifiziert, wobei die Patienten aus dem *Cluster 2* ein signifikant höheres Ersterkrankungsalter aufwiesen als die Patienten aus dem *Cluster 1* (*Cluster 1*: Median von 70 Jahren & *Cluster 2*: Median von 80 Jahren).

Für das Geschlecht, chromosomale Imbalancen im *MYC*-Gen, der Genexpression von *Ki67* und *MUM1* und Mutationen in den Genen *EP300*, *CCND3*, *PIMI* sowie *SOCS1* wurden keine signifikanten Unterschiede zwischen den Patienten aus dem Cluster 1 und 2 identifiziert.

Zusammenfassend wurden in dieser Arbeit im Rahmen der altersassoziierten Analysen genetischer Aberrationen bei DLBCL-Patienten gezeigt, dass im hohen Alter mehr Frauen betroffen waren und diese überwiegend vom ABC- bzw. nonGCB-Subtyp waren. Mittels statistischer Verfahren konnte gezeigt werden, dass mit zunehmendem Alter der DLBCL-Patienten die Genexpression der Proteine *MUM1*, *Ki67* und *BCL2* signifikant zu und die Expression von *CD10* mit dem Alter der DLBCL-Patienten signifikant abnahm (siehe Abschnitt 3.2.1.2). Zusätzlich wurden 12 Gene ermittelt bei denen das mediane Alter der mutierten Patienten über 80 Jahren lag (siehe Abschnitt 3.2.1.2.1) und es konnten strukturelle Aberrationen und Punktmutationen im *CD58*-Gen insbesondere bei DLBCL-Patienten im sehr hohen Alter identifiziert werden (MF: ~10%, Median: 74 Jahre, siehe Abschnitt 3.2.1.2.2). In einer abschließenden Assoziationsanalyse von genetischen Markern bei DLBCL konnten zwei Cluster identifiziert werden, die sich signifikant im Ersterkrankungsalter der Patienten innerhalb der Cluster unterschieden. Die Patienten aus *Cluster 2* waren signifikant älter und wiesen häufiger genetische Aberrationen in den Genen *MYD88*, *CD79B*, *PRDM1* und *TP53* auf. Zudem waren die Patienten aus *Cluster 2* signifikant häufiger vom DLBCL-Subtyp ABC bzw. nonGCB und exprimierten deutlich stärker das Protein *BCL2* (siehe Abbildung 42).

3.2.1.3 Vergleich von altersassoziierten genetischen Markern bei folliculären und diffus großzelligen B-Zell-Lymphomen

Im Rahmen dieser Arbeit wurden altersassoziierte Analysen von genetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen durchgeführt, um Aberrationen in Patienten mit einem sehr hohen Ersterkrankungsalter von vorwiegend über 80 Jahren zu identifizieren. In der vorliegenden Arbeit wurden altersassoziierte molekulargenetische Untersuchungen an insgesamt 1609 DLBCL-Patienten im Alter von 3 bis 97 Jahren (Median: 66 Jahre) und 283 FL-Patienten im Alter von 10 bis 95 Jahre (Median: 60 Jahre) durchgeführt. Eine Übersicht der Ergebnisse der altersassoziierten Analysen hinsichtlich genetischer Aberrationen in DLBCL und FL sind in der nachfolgenden Tabelle 31 zusammengetragen.

Tabelle 31: Vergleich der im Rahmen dieser Arbeit identifizierten altersassoziierten genetischen Marker

| Charakteristika | | DLBCL (N= 1609) | | FL (N= 283) | |
|----------------------------|---------|---------------------------------|------------------------------|---------------------------------|------------------------------|
| Median Alter (min, – max,) | | 66 (3-97) | | 60 (10-95) | |
| Geschlecht (m / w) | | 841 / 768 | | 145 / 138 | |
| | Methode | Logistische Regression (P-Wert) | Kruskal-Wallis Test (P-Wert) | Logistische Regression (P-Wert) | Kruskal-Wallis Test (P-Wert) |
| Geschlecht | | <0,001 | <0,001 | <0,001 | <0,001 |
| <i>BCL2</i> | FISH | 0,16 | 0,92 | 0,99 | 0,05 |
| <i>BCL6</i> | | 0,08 | 0,37 | 0,64 | 0,53 |
| <i>MYC</i> | | 0,43 | 0,63 | 0,07 | 0,04 |
| <i>IGH</i> | | 0,18 | 0,13 | 0,86 | 0,77 |
| <i>BCL2</i> | IHC | 0,01 | 0,03 | 0,70 | 0,64 |
| <i>BCL6</i> | | 0,34 | 0,4134 | 0,03 | 0,05 |
| Ki67 | | <0,01 | <0,01 | 0,48 | 0,26 |
| MUM1 | | 0,29 | 0,27 | 0,62 | 0,65 |
| CD10 | | <0,001 | <0,001 | 0,78 | 0,64 |
| <i>KMT2D</i> | | 0,65 | NA | 0,05 | NA |
| <i>CREBBP</i> | | 0,53 | NA | 0,01 | NA |

Abkürzungen und Verweise: Signifikante P-Werte ($\leq 0,050$) sind in grau markiert wurden. m: Männlich. w: Weiblich. Alter: in Jahren angegeben. NA: Nicht verfügbar, weil kein Kruskal-Wallis Test mit den Ergebnissen durchgeführt wurde. In der Tabelle wurden die Gene *KMT2D* und *CREBBP* als altersassoziierte genetische Marker aus den Sequenz-Analysen aufgelistet, da diese zwei Gene sowohl in DLBCL als auch in FL eine MF von über 10% aufwiesen. Für alle anderen rekurrent mutierten Gene wurde entweder nur bei DLBCL oder FL eine MF von über 10% identifiziert.

Die Altersverteilungen der analysierten DLBCL- und FL-Patienten zeigte, dass bei sehr alten Patienten überwiegend DLBCL diagnostiziert wurden (137/161). Das mediane Alter der analysierten Patientengruppen lag in der FL-Kohorte bei 60 Jahren und in der DLBCL-Kohorte bei 66 Jahren (siehe Tabelle 31). Der Vergleich der Altersstruktur mit dem Geschlecht in der analysierten Keimzentrums-B-Zell-Lymphom-Kohorte zeigte, dass in der FL-Kohorte mit zunehmendem Alter mehr Männern betroffen waren. Hingegen erkrankten im hohen Alter der analysierten DLBCL-Kohorte mehr Frauen. In den FISH- und IHC-Analysen dieser Arbeit (siehe Abschnitt 3.2.1.1 und 3.2.1.2) konnten keine gemeinsamen altersassoziierten Marker für genetische Aberrationen in FL und DLBCL-Patienten identifiziert werden. Hingegen wurden Unterschiede zwischen DLBCL und FL identifiziert, die z.T. bereits bekannt sind und auch in der molekularen Diagnostik der Lymphome Anwendung finden (z.B. t(14;18) bei FL). In der DLBCL-Kohorte wurden altersassoziierte Aberrationen in den analysierten Genprodukten Ki67, CD10, und BCL2 identifiziert. Währenddessen wurden innerhalb der FL-Kohorte altersassoziierte Signifikanzen für chromosomale Imbalancen unter Beteiligung der analysierten Gene *BCL2* und *MYC* ermittelt. Zudem wurde in der FL-Kohorte eine signifikante Assoziation zwischen dem Ersterkrankungsalter der Patienten und der Expression des Proteins BCL6 identifiziert, wobei im höheren Alter weniger BCL6 bei FL-Patienten exprimiert wurde (siehe Tabelle 31).

In den Mutationsanalysen dieser Arbeit wurden neun Gene in FL (siehe Abschnitt 3.2.1.1) und 50 Gene in DLBCL (siehe Abschnitt 3.2.1.2) als rekurrent mutiert identifiziert. D.h. es wurden mind. fünf Patienten mit einer potentiell proteinverändernden Mutation in der analysierten FL- bzw. DLBCL-Kohorte identifiziert. Ein Vergleich aller rekurrent mutierten Gene mit einer MF von über 10% in DLBCL mit den rekurrent mutierten Gene mit einer MF von über 10% in FL zeigte, dass zwei Gene (*KMT2D* und *CREBBP*) häufig in DLBCL und FL mutiert vorlagen (Abbildung 43). Mittels logistischer Regression konnte gezeigt werden, dass bei FL-Patienten mit dem Alter signifikant weniger *KMT2D*-Mutationen vorlagen. Für Mutationen im *KMT2D*-Gen bei DLBCL-Patienten sowie Mutationen im *CREBBP*-Gen bei DLBCL- und FL-Patienten lag keine Signifikanz zwischen dem Ersterkrankungsalter und dem Mutationsprofil vor. Die MF von sieben Genen war in der FL-Kohorte höher als in der DLBCL-Kohorte. Demgegenüber waren 24 Gene ausschließlich in DLBCL mit einem MF von über 10% mutiert (Abbildung 43).

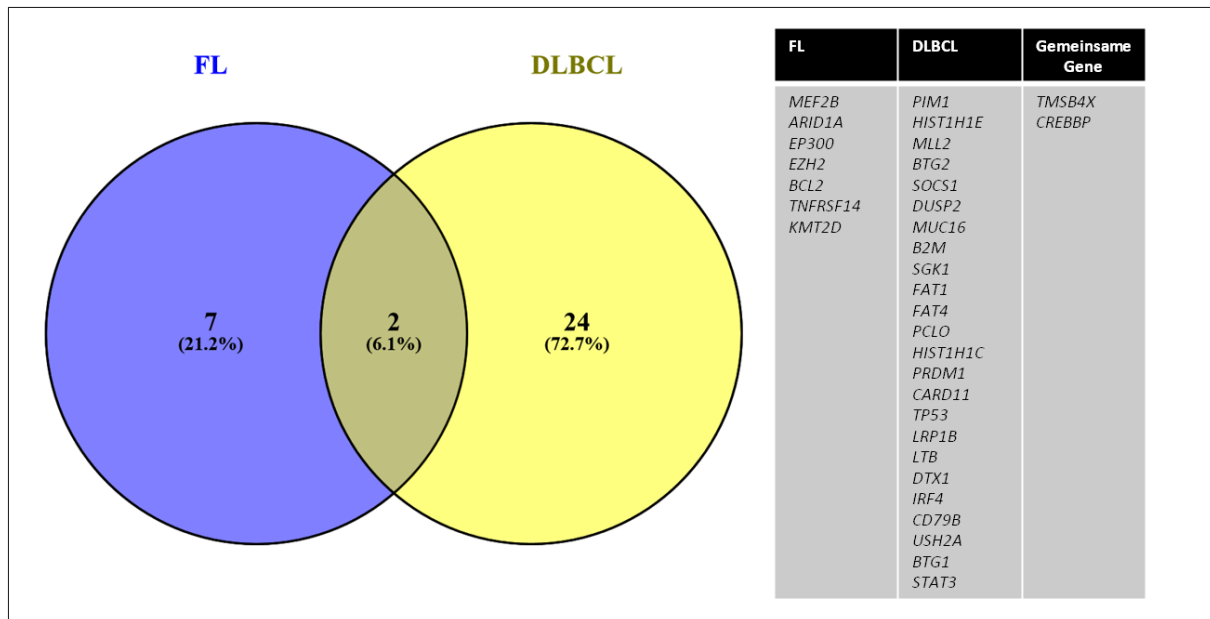


Abbildung 43: Venn-Diagramm der rekurrent mutierten Gene mit einer Mutationsfrequenz von über 10% in DLBCL und FL. Ein Vergleich der rekurrent mutierten Gene mit einer Mutationfrequenz (MF) von über 10% aus den Mutationsanalysen in FL (N=9, siehe 3.2.1.1) und DLBCL (N= 26, siehe 3.2.1.2) zeigten, dass *KMT2D* und *CREBBP* sowohl in DLBCL als auch in FL rekurrent mutiert vorlagen. In den Mutationsanalysen dieser Arbeit lagen 7 Gene ausschließlich in FL und 24 Gene ausschließlich in DLBCL rekurrent mutiert mit einem MF von über 10% vor.

Zusammenfassend wurde im Rahmen der vorliegenden Arbeit gezeigt, dass die analysierten Lymphom-Entitäten DLBCL und FL verschiedene Altersverteilungen der jeweils betroffenen Patienten aufwiesen und bei senilen Patienten überwiegend ein DLBCL diagnostiziert wurde. Es wurde gezeigt, dass bei senilen DLBCL-Patienten häufiger der ABC- bzw. nonGCB-Subtyp auftritt und die Genexpression der Proteine BCL2, und Ki67 zu- sowie CD10 abnimmt. Insgesamt wurden zwölf rekurrent mutierte Gene mit einem medianen Alter der jeweils mutierten DLBCL-Patienten von über 80 Jahren identifiziert. Für vier der zwölf Gene (*MYD88*, *IRF4*, *TP53* und *CD79B*) konnte eine MF von über 10% in DLBCL berechnet werden, weswegen diese Gene als altersassoziierte Marker in senilen DLBCL-Patienten identifiziert wurden. Zusätzlich wurde *CD58* als Kandidatengen für genetische Aberrationen bei senilen Patienten mit einem DLBCL ermittelt (P-Wert $\leq 0,050$).

3.2.2 Epigenetische Aberrationen bei reifzelligem Keimzentrums-B-Zell-Lymphomen

Es ist bekannt, dass neben den genetischen Aberrationen auch epigenetischen Modifikationen der DNA die Genexpression regulieren und verändern können. In Tumoren führt z.B. eine Veränderung des DNA-Methylierungsmusters (Hypermethylierung) in der Promotorregion von Tumorsuppressorgenen häufig zu einer Inaktivierung dieser Gene. Daraufhin verliert das Tumorsuppressorgen seine Funktion und es kommt zur unkontrollierten Proliferation von Tumorzellen^{82,254,255}. Analysen zur Altersassoziation von epigenetischen Aberrationen bei Keimzentrums-B-Zell-Lymphomen sind bisher auf wenige Studien beschränkt. Daraufhin wurden in der vorliegenden Arbeit altersassoziierte epigenetischen Veränderungen des DNA-Methylierungsmusters von Keimzentrums-B-Zell-Lymphomen durchgeführt. Im Vergleich zu Kontroll-Proben weisen Tumor-Proben des selben Ursprungs ein verändertes Methylierungsmuster auf. Zusätzlich verändert sich das DNA-Methylierungsmuster von jedem Individuen während des Alternsprozesses des menschlichen Organismus. Daher wurden in die DNA-Methylierungsanalysen dieser Arbeit verschiedene Kontroll-Proben und Tumor-Entitäten aus unterschiedlichen Altersgruppen eingeschlossen, um innerhalb des Datensatzes zu überprüfen, ob sich die Proben nach ihrem DNA-Methylierungsprofil in die zugehörigen Entitäten auftrennen. Da keine Methylierungsdaten von senilen FL-Patienten vorlagen, wurde die Analyse altersassoziierter epigenetischer Aberrationen ausschließlich an DLBCL durchgeführt. Eine Übersicht des analysierten Datensatzes im Zusammenhang mit der Altersverteilung der analysierten Tumor- und Kontroll-Proben ist in der nachfolgenden Abbildung 44 zusammengestellt.

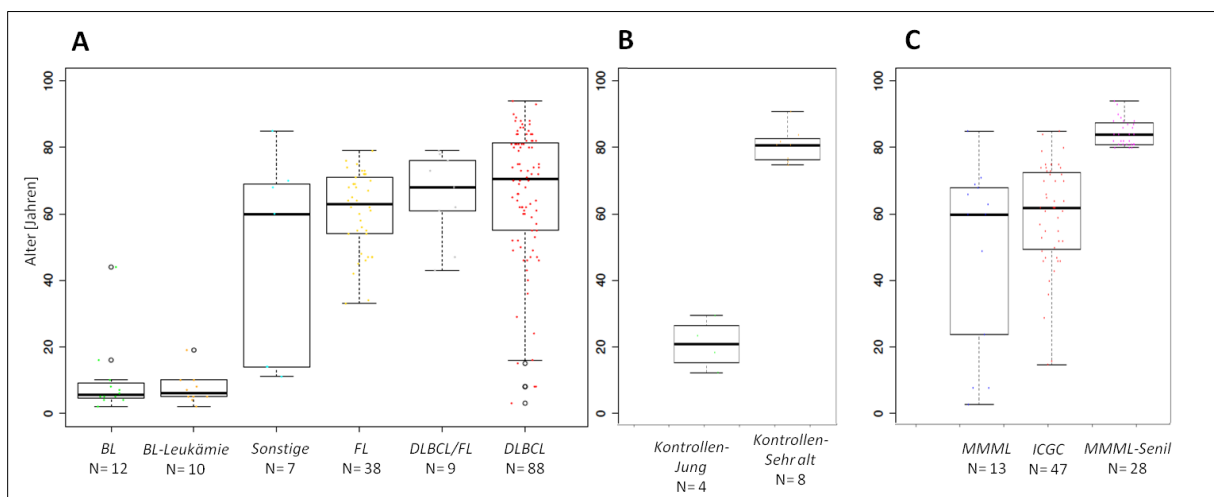


Abbildung 44: Altersverteilung der Patienten-Kohorten die in den DNA-Methylierungsanalysen betrachtet wurden

A: Die Altersverteilung der 164 Tumor-Proben des analysierten Datensatzes B: Die Altersverteilung der 12 Kontroll-Proben des analysierten Datensatzes C: Die Altersverteilung der DLBCL-Patienten des analysierten Datensatzes. Im Rahmen der vorliegenden Arbeit wurden DNA-Methylierungsanalysen von 88 DLBCL-Patienten im Alter von 3 bis 94 Jahren (Median: 71 Jahre) durchgeführt. Abkürzungen: BL: Burkitt-Lymphom. BL-Leukämie: Burkitt-Leukämie. FL: Follikuläres Lymphom. DLBCL: Diffus großzelliges B-Zell-Lymphom. DLBCL/FL: Diffus großzelliges B-Zell-Lymphom mit einem Anteil von Tumorzellen eines Follikulären Lymphoms. *Kontrollen-Jung*: FACS-Sortierte Zellen von gesunden Spendern. *Kontrollen-Sehr alt*: Lymphadenitisproben. MMML: MMML-Verbundprojekt. ICGC: ICGC MMML-Seq Projekt. *MMML-Senil*: In dieser Arbeit neu zusammengestellte Kohorte von senilen DLBCL-Patienten aus der Kohorte des MMML-Verbundprojektes.

An dieser Stelle wird noch einmal erwähnt, dass die Methylierungsdaten von allen analysierten Proben aus extrahierter DNA von Frischmaterial erhoben wurden, um technische Artefakte aufgrund des Ausgangsmaterials auszuschließen. Zur DNA-Methylierungsanalyse bei Keimzentrums-B-Zell-Lymphomen wurden im Rahmen dieser Arbeit die Methylierungsdaten von 27 senilen DLBCL-Patienten der Kohorte *MMML-Senil* (Median: 84 Jahre) mittels des Infinium HumanMethylation450 BeadChip (450K) neu erhoben und zusammen mit weiteren 137 Tumor-Proben in einem Datensatz von insgesamt 164 Tumor- (Median: 63 Jahre) und 12 Kontroll-Proben (Median: 77 Jahre) hinsichtlich altersassoziierter epigenetischer Aberrationen in DLBCL analysiert. Die insgesamt 164 Tumor-Proben des analysierten Datensatzes (Abbildung 44A) setzten sich aus folgenden Entitäten zusammen: diffus großzelligen B-Zell-Lymphomen (*DLBCL*, N= 88, Median: 71 Jahre, Abbildung 44C), folliculären Lymphomen (*FL*, N= 38, Median: 63 Jahre), diffus großzellige B-Zell-Lymphome mit einem Anteil von Tumorzellen eines folliculären Lymphoms (*DLBCL/FL*, N= 9, Median: 9 Jahre), Burkitt-Lymphomen (*BL*, N=12, Median: 6 Jahre), Burkitt-Leukämien (*BL-Leukämie*, N=10, Median: 6 Jahre) und sonstigen B-Zell-Lymphomen (4 hochmaligne-, 1 niedermaligne- und 2 Marginalzonen-Lymphome, N= 7, Median: 60 Jahre). Als Kontroll-Proben (Abbildung 44B) für die DNA-Methylierungsanalysen wurden die FACS-sortierten Zellen (*Kontrollen-Jung*, N= 4, Median: 22 Jahre) und Lymphadenitis-Proben (*Kontrollen-Sehr alt*, N= 8, Median: 81 Jahre) verwendet.

3.2.2.1 Qualitätskontrolle der DNA-Methylierungsdaten des gesamten Datensatzes

Zur Qualitätskontrolle des analysierte Datensatzes wurde zuerst die Loci detection rate (LDR) für alle Proben ermittelt, die für alle 164 Tumor- und 12 Kontroll-Proben bei über 99% (LDR >99%) lag, woraufhin alle Proben in die weitere Auswertung gingen. Anschließend wurde eine SNP- und Gesamt-Korrelationsanalyse für den gesamten analysierten Datensatz erstellt (Durchführung siehe unter Abschnitt 2.2.7.1). Die erzeugte SNP-Korrelationsmatrix (Abbildung 45A) zeigte, dass unterschiedliche Proben eines Individuums ähnliche Korrelationskoeffizienten aufwiesen (*Kontrollen-Jung*). Für die Tumor- und Lymphadenitis-Proben (*Kontrollen-Sehr alt*) wurden keine Mehrfachhybridisierungen durchgeführt und jeweils nur 1 Probe hybridisiert. Die ermittelte SNP-Korrelationsmatrix zeigte ein erwartet sehr heterogenes Muster von geringen Korrelationskoeffizienten für die Tumor- und Lymphadenitis-Proben. Damit konnte in dem analysierten Datensatz eine Mehrfachhybridisierung des gleichen Patienten, wobei es sich um eine Probenvertauschung gehandelt hätte, weitestgehend ausgeschlossen werden. Die Gesamt-Korrelationsmatrix (Abbildung 45B) zeigte für den analysierten Datensatz zwei Muster: niedrige Korrelationskoeffizienten (blau/gelb) und hohe Korrelationskoeffizienten (rot). In den Tumor-Proben waren zumeist geringe Korrelationskoeffizienten zu erkennen, was darauf schließen ließ, dass eine hohe Variabilität in den AVG-Beta-Werten vorlag. Bei den Kontrollen wurden Blöcke mit Werten von hohen Korrelationskoeffizienten ermittelt, was bedeutet, dass die Kontroll-Proben eine geringere Variabilität in den AVG-Beta-Werten aufwiesen.

Es gab einzelne Hybridisierungen von FL-Patienten deren ermittelter Korrelationskoeffizient nicht in das Muster der anderen Tumor-Proben passte und ähnlich dem Muster der Normalkontrollproben war. Da diese FL-Proben jedoch in der SNP-Korrelationsmatrix unauffällig aussahen (ähnlich dem Korrelationskoeffizientenmuster der anderen Tumor-Proben), wurde keine Probe aus den weiteren Analysen ausgeschlossen. Insgesamt wurden in den Qualitätskontrollen des analysierten Datensatzes mittels LDR und Korrelationsmatrizen keine Proben oder Loci ermittelt, die eine unzureichende Qualität aufwiesen. Daraufhin wurden alle 176 Proben mit jeweils 485.577 Loci zusammen im Statistikprogramm R normalisiert.

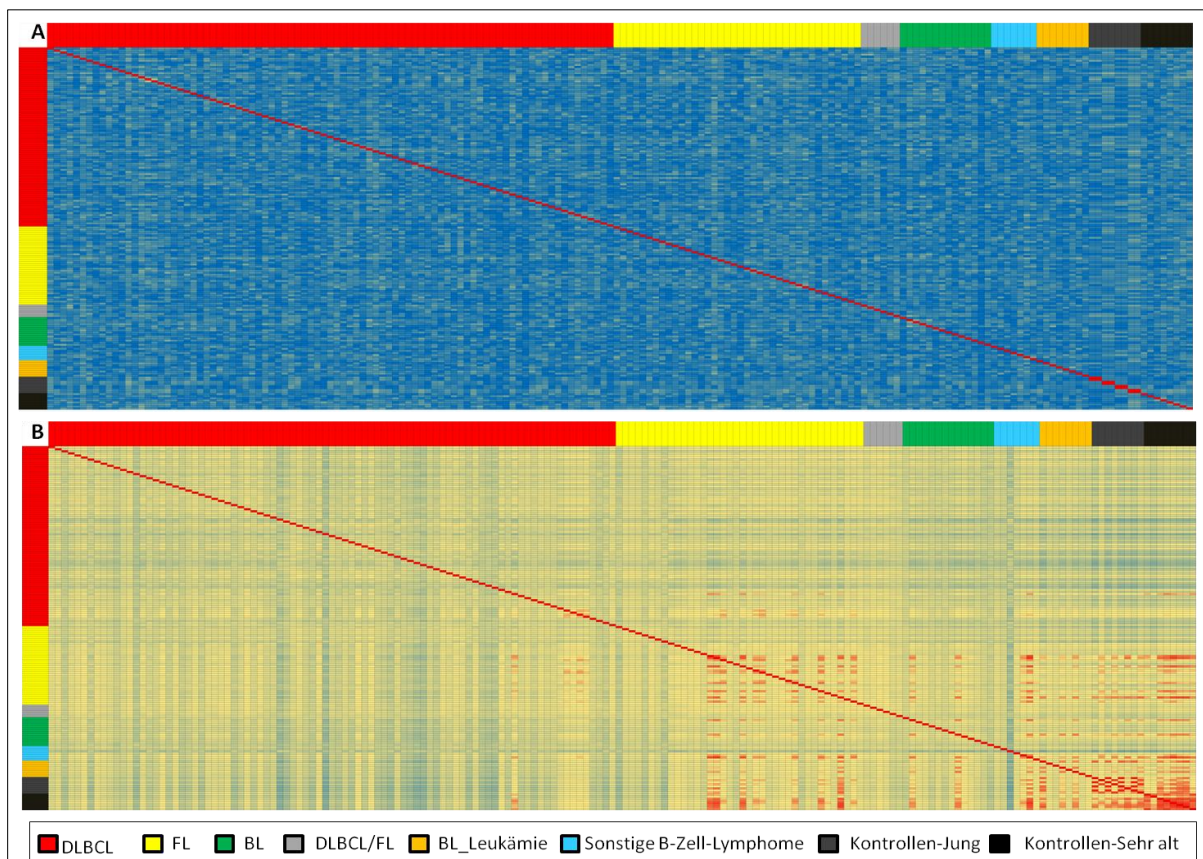


Abbildung 45: SNP- und Gesamt-Korrelationsmatrix des analysierten Datensatzes in den DNA-Methylierungsanalysen. Die erhobenen AVG-Beta-Werte des 450K von 164 Tumor- (farbig nach Lymphom-Subtypen markiert) und 12 Kontroll-Proben (dunkelgrau/schwarz markiert) wurden mittels Korrelationsmatrizen analysiert. Hierbei ist zu beachten, dass von den Kontroll-Proben der Kohorte *Kontrollen-Jung* jeweils zwei Zell-Populationen eines Individuums analysiert wurden, die als separat Hybridisierungen in den SNP- und Korrelationsmatrizen eingegangen und gekennzeichnet sind (dunkelgrau markiert). Von den Tumor- und Lymphadenitis-Proben wurden keine Mehrfachhybridisierungen durchgeführt. Die Rechtecke horizontal bzw. vertikal neben der farbigen Matrix repräsentieren jeweils eine Probe, wobei die Reihenfolge der Proben in beiden Richtungen identisch ist. Jedes farbige Rechteck innerhalb der Matrix spiegelt den Korrelationskoeffizienten wider. Der unsymmetrische Farbcode für die Werte der Korrelationskoeffizienten ist: 0= blau; ab 0,95= gelb; 1= rot. A: Die SNP-Korrelationsmatrix zeigt die Korrelation aller 65 auf dem 450K lokalisierten “rs“-Loci für jede analysierte Probe. Für die hybridisierten Proben vom gleichen Individuum wurden ähnliche Korrelationskoeffizienten ermittelt (*Kontrollen-Jung*, dunkelgrau markiert). Von den analysierten Tumor- und Lymphadenitis-Proben wurden keine Mehrfachhybridisierungen durchgeführt, wonach in der ermittelte Korrelationsmatrix ein erwartungsgemäß heterogenes Muster von geringen Korrelationskoeffizient über alle “rs“-Loci zu erkennen war. B: Die Gesamt-Korrelationsmatrix zeigt die Korrelation der 485.512 “nicht-rs“-Loci für jede analysierte Probe. Insgesamt zeigte die Korrelationsmatrix für die Tumor-Proben größtenteils sehr ein heterogenes Muster (gelb/blau markiert, hohe Variation in den AVG-Beta-Werten), wohingegen für die Kontroll-Proben einzelne Blöcke von hohen Korrelationskoeffizienten zu erkennen waren (rot markiert, geringe Variation in den AVG-Beta-Werten). Abkürzungen: DLBCL: Diffus großzellige B-Zell-Lymphome. FL: Follikuläre Lymphome. BL: Burkitt-Lymphome. DLBCL/FL: Diffus großzellige B-Zell-Lymphome mit einem Anteil von Tumorzellen eines Follikulären Lymphoms.

3.2.2.2 Differenziell methylierte Regionen in diffus großzelligen B-Zell-Lymphomen

Nach dem Filtern (detailliert beschrieben im Abschnitt 2.2.7.2) wurden die normalisierten AVG-Beta-Werte der 176 Proben mit jeweils 444.165 Loci in die Software Qlucore Omics eingeladen. Zuerst wurde eine unsupervidierte Hauptkomponentenanalyse (engl. principal component analysis, PCA) mit den AVG-Beta-Werten von allen 176 Proben durchgeführt. In einem Multigruppenvergleich (ANOVA) bei einer Varianz von 0,3 wurden 250.929 CpG-Loci als differentiell methyliert zwischen den Gruppen *BL*, *BL-Leukämie*, *DLBCL*, *FL*, *DLBCL/FL*, *Kontrollen-Jung* und *Kontrollen-Sehr alt* identifiziert. In dieser PCA stellten die analysierten Tumor-Proben eine sehr heterogene Gruppe hinsichtlich des DNA-Methylierungsmusters dar und es war keine eindeutige Gruppierung der Proben in Tumor- und Kontrollen zu erkennen (Abbildung 46A). Eine supervidierte Analyse der 176 Proben mit einer Varianz von 0,3 und $FDR \leq 1e-18$ zeigte eine deutliche, wenn auch nicht vollständige, Gruppierung des Datensatzes in *DLBCL*, *FL* und *DLBCL/FL*-Proben gegenüber den *BL*-, *BL-Leukämie*- und Kontrollproben aufgrund der differentiellen DNA-Methylierung von 80 CpG-Loci (Abbildung 46B).

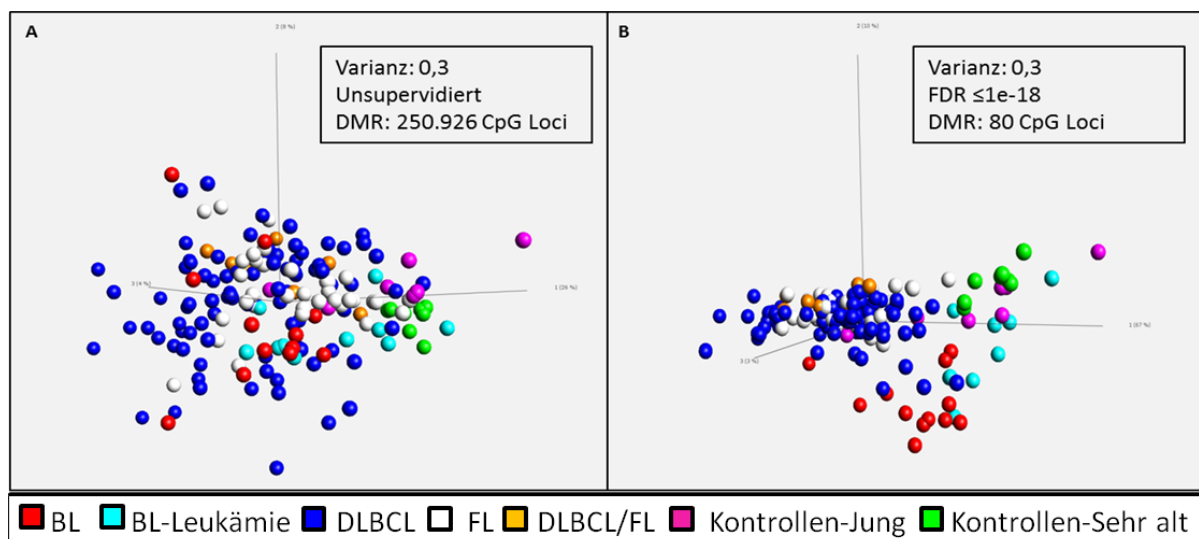


Abbildung 46: Hauptkomponentenanalyse des analysierten Datensatzes

In der Analyse sind die AVG-Beta-Werte von 176 Proben in jeweils 444.165 Loci in einem Multigruppenvergleich ausgewertet worden. A: Die unsupervidierte PCA-Analyse der Tumor- und Kontrollproben bei einer Varianz von 0,3 zeigte keine Gruppierung der analysierten Proben in Tumoren und Kontrollen hinsichtlich des DNA-Methylierungsmusters. B: Die supervidierte PCA-Analyse der Tumor- und Kontrollproben bei einer Varianz von 0,3 und einer $FDR \leq 1e-18$ zeigte eine deutliche, wenn auch nicht vollständige Gruppierung des Datensatzes in *DLBCL*, *FL* und *DLBCL/FL*-Proben gegenüber den *BL*-, *BL-Leukämie*- und Kontrollproben aufgrund der differentiellen DNA-Methylierung von 80 CpG-Loci. Abkürzungen: DLBCL: Diffus großzellige B-Zell-Lymphome. FL: Follikuläre Lymphome. BL: Burkitt-Lymphome. DLBCL/FL: Diffus großzellige B-Zell-Lymphome mit einem Anteil von Tumorzellen eines Follikulären Lymphoms.

Um die Heterogenität der DNA-Methylierung innerhalb der Gruppe der reifzelligigen Keimzentrums-B-Zell-Lymphome (N= 126) beurteilen zu können, wurde die PCA wiederholt und nur die AVG-Beta-Werte der Lymphom-Entitäten DLBCL (N= 88) und FL (N=38) analysiert. Die unsupervidierte PCA-Analyse für den Gruppenvergleich zwischen den Entitäten DLBCL und FL (insgesamt N=126) und einer Varianz von 0,3 ermittelte 255.557 differentiell methyliert CpG-Loci. In dieser PCA stellten die analysierten DLBCL-Proben in ihrem DNA-Methylierungsmusters eine deutlich heterogenere Gruppe im Vergleich zu den FL-Proben dar. Eine eindeutige Gruppierung der analysierten Proben nach den AVG-Beta-Werten in DLBCL und FL war nicht möglich (Abbildung 47A). Eine supervidierte Analyse der 126 Proben mit einer Varianz von 0,3 und FDR $\leq 0,01$ zeigte keine Gruppierung des Datensatzes in DLBCL- und FL-Proben. Unter Anwendung dieser Auswertungsparameter wurden 316 differentiell methylierte CpG-Loci zwischen DLBCL und FL identifiziert (Abbildung 47B). Die anschließende Auswertung der erstellten Liste der differentiell methylierten CpG-Loci zeigte, dass nur 6/316 CpG-Loci einem Gen zugeordnet wurden und davon kein Gen mehrfach betroffen war. Alle weiteren 310/316 CpG-Loci waren keinem Gen zugeordnet und wurden mit NA von der GenomeStudio Software, woraus die Information zur Annotation der CpG-Loci exportiert wurden (detailliert beschrieben im Abschnitt 2.2.7.1), bezeichnet. Diese als NA bezeichneten Loci lagen genomweit verteilt und es konnte keine Region mit mind. 3 differentiell methylierten Loci in unmittelbarer Nähe (<500 bp) identifiziert werden.

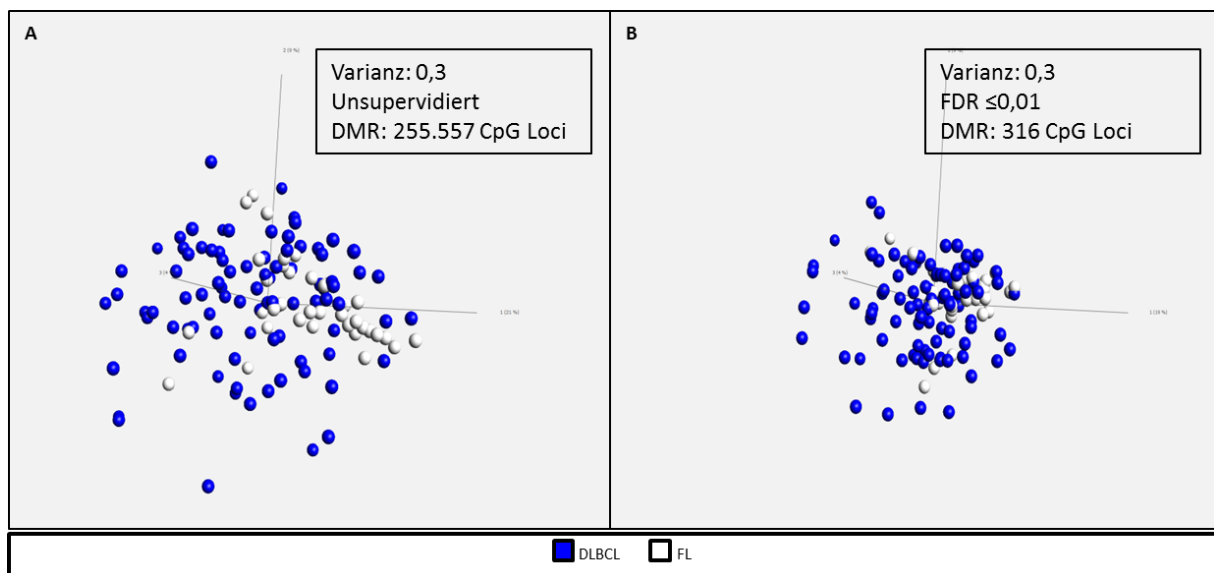


Abbildung 47: Hauptkomponentenanalyse der analysierten DLBCL- und FL-Proben

In der Analyse sind die AVG-Beta-Werte von 126 DLBCL- und FL-Proben in jeweils 444,165 Loci in einem Multigruppenvergleich ausgewertet worden. A: Die unsupervidierte PCA-Analyse der Tumor-Proben bei einer Varianz von 0,3 zeigte keine Gruppierung der analysierten Proben in DLBCL und FL hinsichtlich des DNA-Methylierungsmusters. B: Die supervidierte PCA-Analyse der Tumor-Proben bei einer Varianz von 0,3 und einer FDR $\leq 0,01$ zeigte, dass die analysierten DLBCL-Patienten eine deutlich heterogenere Gruppe im DNA-Methylierungsmuster im Vergleich zu den FL-Patienten darstellen. Unter den angewandten Auswertungsparametern wurden 316 differentiell methylierte CpG-Loci zwischen DLBCL und FL identifiziert. Abkürzungen: DLBCL: Diffus großzelliges B-Zell-Lymphom. FL: Follikuläres Lymphom.

Da der Fokus der vorliegenden Arbeit in den altersassoziierten Analysen von Keimzentrums-B-Zell-Lymphomen lag, wurde anschließend das DNA-Methylierungsprofil der Tumore hinsichtlich des Erstdiagnosealters der Patienten, analysiert. Zur altersassoziierten DNA-Methylierungsanalyse von Keimzentrums-B-Zell-Lymphomen wurden die 450K-Daten von 88 DLBCL-Patienten im Alter von 3 bis 94 Jahren (Median: 71 Jahre) ausgewertet. Zuerst wurden alle analysierten DLBCL-Patienten nach dem Ersterkrankungsalter in folgende Altersgruppen eingeteilt: Gruppe 1: 0-39 Jahre (N=9, Median: 16 Jahre), Gruppe 2: 40-59 Jahre (N= 16, Median: 50 Jahre), Gruppe 3: 60-79 Jahre (N= 31, Median: 69 Jahre) und Gruppe 4 ≥ 80 Jahre (N= 32, Median: 84 Jahre). Anschließend wurde eine lineare Regression der AVG-Beta-Werte der CpG-Loci gegenüber dem Ersterkrankungsalter der Patienten in der Software Qlucore Omics durchgeführt. Die Ergebnisse der linearen Regression wurden in der zugehörigen PCA grafisch dargestellt, wobei die zuvor erstellten DLBCL-Altersgruppen durch einen unterschiedlichen Farbcode dargestellt wurden (Abbildung 48A). Die logistische Regression bei einer FDR ≤ 1 (Standardeinstellung für logistische Regression in der Software Qlucore Omics) und einer Varianz von 0,3 ermittelte insgesamt 258,564 differentiell methylierte CpG-Loci in DLBCL-Patienten gegenüber ihrem Ersterkrankungsalter. Nachfolgend wurde die FDR verringert, um potenziell signifikant differenziell methylierte CpG-Loci in DLBCL zu identifizieren. Bei einer FDR $\leq 0,5$ wurden bereits nur noch 74 differentiell methylierte CpG-Loci in DLBCL-Patienten gegenüber ihrem Ersterkrankungsalter berechnet (Abbildung 48B).

Zur genaueren Betrachtung wurden die 74 differentiell methylierten CpG-Loci in einer supervidierten, hierarchischen Clusteranalyse in Form einer Heatmap dargestellt. Hierzu wurden die Proben hinsichtlich ihres Alters aufsteigend geordnet (Abbildung 48C). Die grafische Darstellung zeigte, dass es Unterschiede im DNA-Methylierungsmuster der jüngeren (<60 Jahre) gegenüber den älteren (>60 Jahre) DLBCL-Patienten gab. Jedoch zeigte das weitere Reduzieren der FDR, dass in den durchgeführten Methylierungsanalysen mittels logistischer Regression in DLBCL keine signifikant differenziell methylierten CpG-Loci vorlagen (P-Wert $\leq 0,050$).

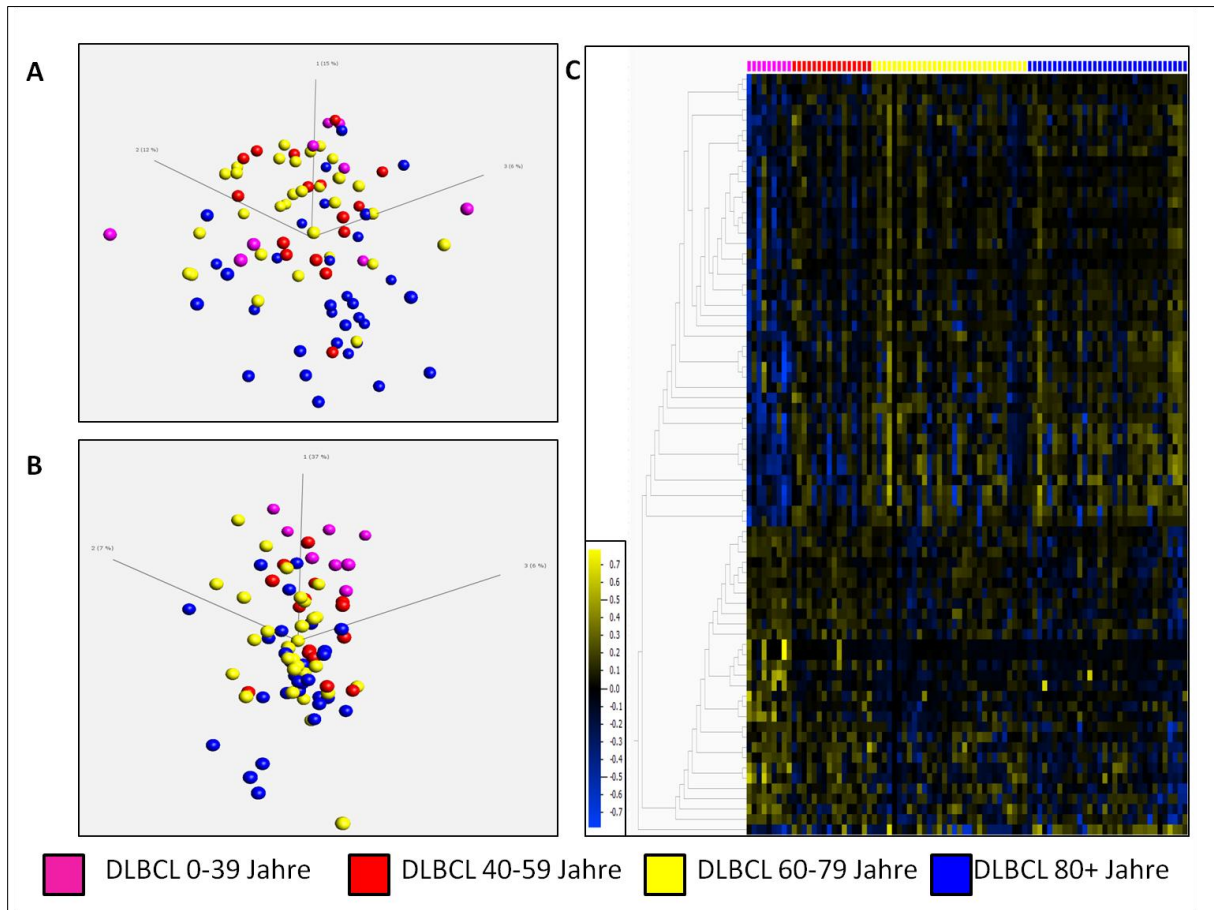


Abbildung 48: Lineare Regressionen und supervidierte, hierarchische Clusteranalyse von DLBCL-Patienten

In den DNA-Methylierungsanalysen wurden epigenetische Aberrationen mittels 450K von 88 DLBCL-Patienten im Alter von 3 bis 97 Jahren (Median: 71 Jahre) betrachtet. Dazu wurden von jedem Patienten die AVG-Beta-Werte von 444.165 Loci mit dem Ersterkrankungsalter des Patienten in einer linearen Regression ausgewertet. Die Patienten wurden nach ihrem Erstdiagnosealter in vier Altersgruppen eingeteilt und grafisch dargestellt: 0-39 Jahre (pink, N= 9), 40-59 Jahre (rot, N= 16), 60-79 Jahre (gelb, N= 31) und <80 Jahre (blau, N= 32). A: Die supervidierte PCA-Analyse mittels linearer Regression bei einer Varianz von 0,3 und einer FDR ≤ 1 zeigte keinen Zusammenhang des Ersterkrankungsalters und der DNA-Methylierung bei den analysierten DLBCL-Patienten. B: Die supervidierte PCA-Analyse mittels linearer Regression bei einer Varianz von 0,3 und einer FDR $\leq 0,5$ zeigte keinen Zusammenhang des Ersterkrankungsalters und der DNA-Methylierung bei den analysierten DLBCL-Patienten. C: Die supervidierte hierarchische Clusteranalyse der 74 differentiell methylierten CpG-Loci aus der PCA (B) zeigte, dass es Unterschiede im DNA-Methylierungsmuster zwischen jüngeren (<60 Jahre) und älteren (>60 Jahre) DLBCL-Patienten gab, diese jedoch nicht signifikant waren.

Zusammenfassend wurde in der linearen Regressionsanalyse keine signifikanten Zusammenhänge zwischen dem DNA-Methylierungsmuster und dem Ersterkrankungsalter von DLBCL-Patienten ermittelt.

3.2.2.3 Epigenetisches Alter nach Horvath von Patienten mit einem diffus großzelligen B-Zell-Lymphom und Normalkontroll-Proben

Im Folgenden wurde das epigenetische Alter der 88 DLBCL-Patienten und der 12 Kontroll-Proben (12 Individuen mit insgesamt 16 Subpopulationen: 8 Lymphadenitis, 4 naive B-Zellen, 3 GCB-Zellen, 1 T-Zellen) nach dem Horvath-Modell bestimmt. Dieses von Steve Horvath entwickelte Modell ermöglicht es mittels der AVG-Beta-Werte von 353 altersassoziierten CpG-Loci des 450K das epigenetische Alter der analysierten Probe zu berechnen. Für zwei Tumor-Proben war es aus bisher ungeklärten Gründen nicht möglich das epigenetische Alter nach Horvath zu bestimmen. Diese Proben wurden in der weiteren Analyse ausgeschlossen.

Von allen weiteren 86 DLBCL-Patienten und den 16 Subpopulationen der Kontroll-Proben wurde das biologische mit dem epigenetischen Alter verglichen und grafisch dargestellt (Abbildung 49).

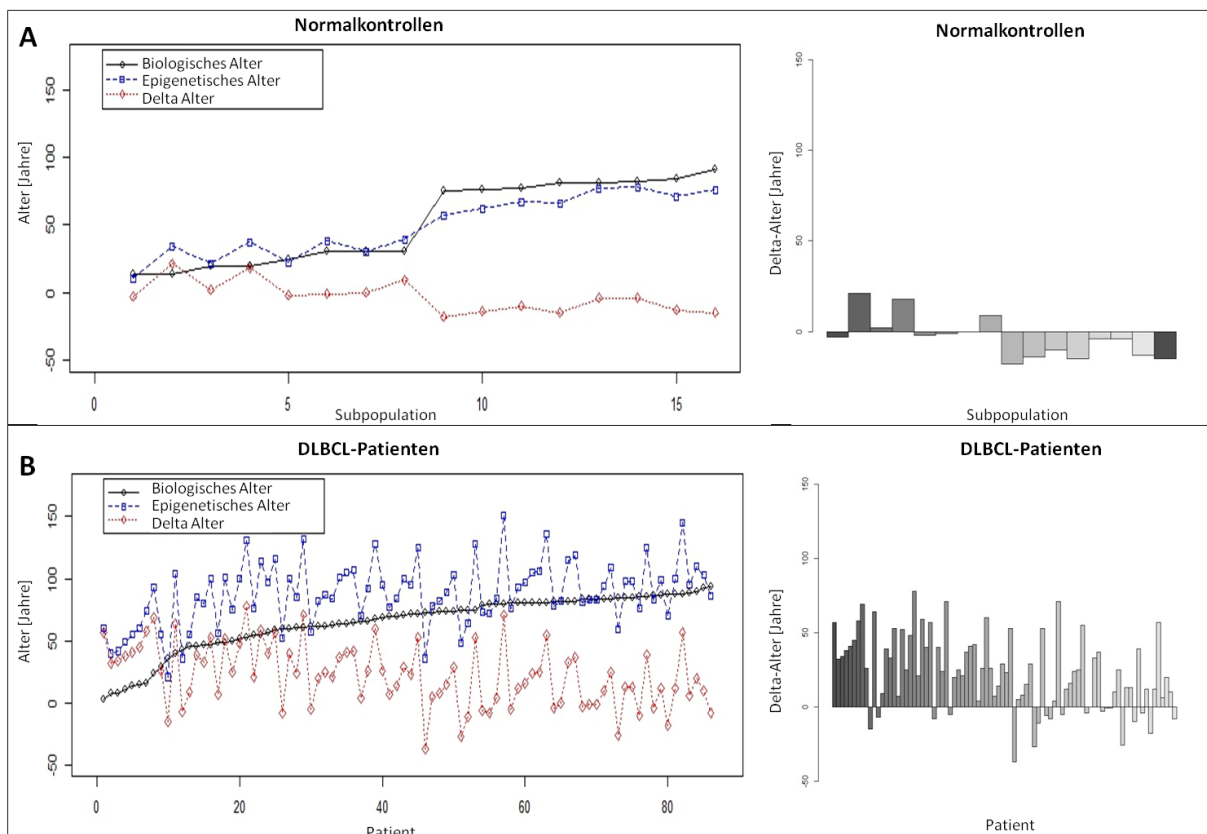


Abbildung 49: Vergleich des biologischen und epigenetischen Alters von DLBCL- und Kontroll-Proben

In der Grafik stellt jeder Punkt bzw. Balken eine analysierte Probe dar. Für die Beurteilung der Altersstruktur wurden zur besseren Visualisierung die Punkte durch eine Linie verbunden. Die Differenz des Alters aus dem biologischen und epigenetischen Alter (Delta-Alter) wurde zusätzlich separat in einem Balkendiagramm dargestellt. Die Proben sind jeweils nach dem biologischen Alter geordnet, damit ist die Reihenfolge der Proben im Punktediagramm analog zum Balkendiagramm. A: Vergleich des biologischen Alters mit dem ermittelten epigenetischen Alter nach Horvath der 12 analysierten Kontroll-Proben von insgesamt 16 Subpopulationen. Die Kontroll-Proben zeigten keine signifikanten Differenzen im biologischen und epigenetischen Alter. B: Vergleich des biologischen Alters mit dem ermittelten epigenetischen Alter nach Horvath der 86 analysierten DLBCL-Proben. Die DLBCL-Proben zeigten signifikante Differenzen im biologischen und epigenetischen Alter (P -Wert $< 0,050$). Bei 60/86 DLBCL-Patienten war das berechnete epigenetische Alter der DLBCL-Proben deutlich höher als das biologische Alter.

Bei den Subpopulationen der Normalkontrollen zeigte der Vergleich des biologischen (13-91 Jahre; Median: 75 Jahre) und epigenetischen Alters (10-78 Jahre; Median: 48 Jahre), dass die Differenzen der Altersangaben (Delta-Alter) im Bereich von 0 bis 21 Jahren (Median: 4 Jahre) lagen und somit sehr gering waren (Abbildung 49A). Die statistische Auswertung mittels Mann-Whitney U Test zeigte, dass die Subpopulationen der Kontroll-Proben keine signifikanten Unterschiede zwischen dem biologischen und epigenetischen Alter aufwiesen (P-Wert >0,050).

Die detaillierte Betrachtung der einzelnen Subpopulationen der Normalkontrollen zeigte, dass alle analysierten Lymphadenitisproben (N= 8) ein geringeres epigenetisches Alter nach Horvath im Vergleich zum biologischen Alter hatten (siehe Tabelle 32) und im Durchschnitt 12 Jahre jünger waren. Von den acht sortierten Subpopulationen der Kontroll-Proben hatten vier ein höheres epigenetisches als biologisches Alter, wobei es sich um die analysierten GCB-Zellen der gesunden Individuen handelte. Für die FACS-sortierten naiven B- und T-Zellen der gesunden Individuen wurde für 3/8 ein jüngerer und für 1/8 ein identisches epigenetisches Alter im Vergleich zum biologischen Alter berechnet.

Tabelle 32: Biologisches und epigenetisches Alter der analysierten Subpopulationen der Normalkontrollen

| Probe | Zellpopulation | Biologisches Alter [Jahre] | Epigenetisches Alter [Jahre] | Delta-Alter [Jahre] |
|--------------|----------------|----------------------------|------------------------------|---------------------|
| Jack-3-NK | Lymphadenitis | 75 | 57 | -18 |
| Jack-7-NK | Lymphadenitis | 81 | 66 | -15 |
| Jack-5-Nk | Lymphadenitis | 91 | 76 | -15 |
| Jack-8-NK | Lymphadenitis | 76 | 62 | -14 |
| Jack-1-NK | Lymphadenitis | 84 | 71 | -13 |
| Jack-4-NK | Lymphadenitis | 77 | 67 | -10 |
| Jack-9-NK | Lymphadenitis | 81 | 77 | -4 |
| Jack-2-NK | Lymphadenitis | 82 | 78 | -4 |
| 4160735-B | Naive B-Zellen | 13 | 10 | -3 |
| 4118819-B | Naive B-Zellen | 24 | 22 | -2 |
| 4118819-nonB | T-Zellen | 24 | 38 | -1 |
| 4174884-B | Naive B-Zellen | 30 | 30 | 0 |
| 4122131-B | Naive B-Zellen | 19 | 21 | 2 |
| 4174884-GCB | GCB-Zellen | 30 | 39 | 9 |
| 4122131-GCB | GCB-Zellen | 19 | 37 | 18 |
| 4160735-GCB | GCB-Zellen | 13 | 34 | 21 |

Abkürzungen: GCB-Zellen: Keimzentrums-B-Zellen. Epigenetisches Alter: Ermitteltes Alter aus den AVG-Beta-Werten von 353 altersassoziierten CpG-Loci des Infinium HumanMethylation450 Bead Chips nach dem Modell von Steve Horvath. Delta-Alter: Differenz aus dem biologischen und epigenetischen Alter.

Bei den DLBCL-Proben zeigte der Vergleich des biologischen Alters (3-94 Jahre, Median: 72 Jahre) mit dem epigenetischen Alter nach Horvath (21-151 Jahre, Median: 87 Jahre) eine große Differenz von einem Alter von -37 bis 78 Jahre (Median: 24 Jahre, Abbildung 49B). Die Differenz des biologischen und epigenetischen Alters wird im Folgenden als Delta-Alter bezeichnet.

Die anschließende statistische Auswertung aller DLBCL-Proben (Gesamtkohorte N= 86) mittels Mann-Whitney U Test identifizierte signifikante Unterschiede zwischen dem biologischen und epigenetischen Alter (P-Wert= 1,79e-08). Das biologische Alter und epigenetische Alter nach Horvath sind im Anhang in der Tabelle A6 zusammengestellt.

Die detaillierte Betrachtung der einzelnen DLBCL-Proben zeigte, dass für die meisten DLBCL-Proben das ermittelte epigenetische Alter deutlich über dem biologischen Alter (66/86, Median: 29 Jahre) lag. 19/86 Proben hatten ein jüngeres epigenetisches Alter im Vergleich zu deren biologischen Alter (Median: -8 Jahre) und eine DLBCL-Probe zeigte ein identisches biologisches Alter und epigenetischen Alter nach Horvath (MPI-749).

Zusammenfassend wurde für die analysierten DLBCL-Proben größtenteils ein signifikant deutlich höheres epigenetisches Alter im Vergleich zum zugehörigen biologischen Alter mittels Mann-Whitney U Test berechnet. Hingegen wurden für die Kontroll-Proben keine signifikanten Unterschiede zwischen dem biologischen und epigenetischen Alter identifiziert.

3.2.2.4 Rekurrent mutierte Gene mit einem hohen biologischen und epigenetischen Alter bei Patienten mit einem diffus großzelligen B-Zell-Lymphom

Im Rahmen der vorliegenden Arbeit wurden das Mutationsprofil und die DNA-Methylierung von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom analysiert. Abschließend wurden die Ergebnisse dieser vorangegangenen Auswertungen kombiniert und die rekurrent mutierten Gene in senilen DLBCL-Patienten (siehe Abschnitt 3.2.1.2) mit dem ermittelten epigenetischen Alter nach Horvath (siehe Abschnitt 3.2.2.3) verglichen. Dazu wurden für die 50 rekurrent mutierten Gene in DLBCL, neben dem biologische Ersterkrankungsalter, das ermittelte epigenetische Alter nach Horvath und die Differenz des biologischen und epigenetischen Alters (Delta-Alter) für jeden Patienten hinzugefügt und das mediane Alter der jeweils mutierten Patienten für jedes Gen betrachtet. Bei dieser Auswertung ist zu beachten, dass nur zu den DLBCL-Proben aus Frischmaterial die 450K-Daten vorlagen. Somit konnte für die DLBCL-Patienten der Kohorte *Jack* kein epigenetisches Alter nach Horvath berechnet werden und diese Fälle sind in den folgenden Analysen ausgeschlossen. Insgesamt standen für die Vergleichsanalysen der Mutationen und der DNA-Methylierung die Sequenz- und 450K-Daten von 49 DLBCL-Patienten mit einem biologischen Ersterkrankungsalter von 15 bis 90 Jahren zur Verfügung (Median: 68 Jahre). Diese Patienten waren aus den Forschungsstudien MMLL (Kohorte: *MMLL-Senil*, N= 21, Median: 84 Jahre) und ICGC MMLL-Seq (N= 28, Median: 57 Jahre).

Für den Vergleich der genetischen und epigenetischen Aberrationen wurden die 50 rekurrent mutierten Gene aus der Mutationsanalyse in Form von Boxplots dargestellt und nach dem epigenetischen Alter der entsprechend mutierten Patienten geordnet (Abbildung 50).

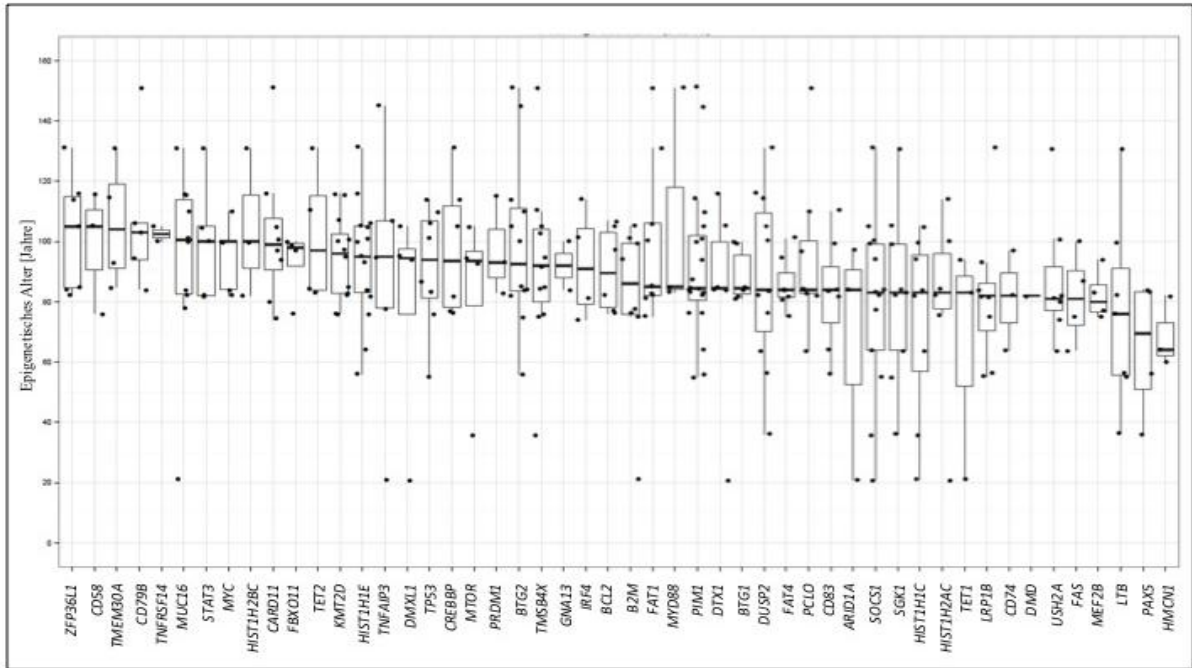


Abbildung 50: Altersverteilung der rekurrent mutierten Gene in DLBCL entsprechend dem epigenetischen Alter

In der Grafik wurden alle 50 rekurrent mutierten Gene aus den Mutationsanalysen von DLBCL gegen das ermittelte epigenetische Alter nach Horvath aufgetragen (N= 49). Die Gene *ZFP36L1*, *CD58*, *TMEM30A*, *CD79B* und *TNFRSF14* zeigten das höchste mediane epigenetische Alter in der Gruppe von Patienten mit einer Mutation.

In den bei DLBCL rekurrent mutierten Genen (N= 50) lag das mediane Alter der mutierten Patienten bei durchschnittlich 85 Jahren. Das höchste mediane Alter wurde für DLBCL-Patienten mit einer potentiell proteinverändernden Mutation in den Genen *ZFP36L1*, *CD58* (Median jeweils: 105 Jahre), *TMEM30A* (Median: 104 Jahre), *CD79B* (Median: 103 Jahre) und *TNFRSF14* (Median: 102,5 Jahre) ermittelt. In der weiteren Auswertung wurden jeweils die 20 rekurrent mutierten Gene mit dem höchsten medianen biologischen Alter im Vergleich zu den Patienten mit denen des höchsten epigenetischen Alter miteinander verglichen (Abbildung 51).

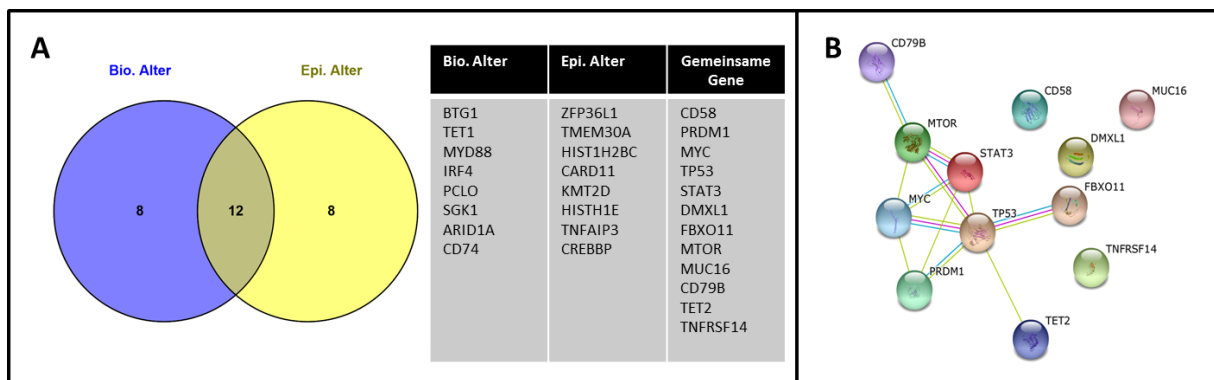


Abbildung 51: Vergleich der gemeinsamen Gene nach dem biologischen und epigenetischen Alter in senilen DLBCL

A: In einem Vergleich der jeweils 20 rekurrent mutierten Gene mit dem höchsten medianen Alter der mutierten Patienten aus den Mutationsanalysen nach dem biologischen und epigenetischen Alter wurden insgesamt 12 gemeinsame Gene identifiziert. Jeweils 8 Gene waren exklusiv in der Kandidatengenliste nach dem biologischen oder epigenetischen Alter in senilen DLBCL-Patienten. B: Die Pathway-Analyse der 12 gemeinsamen Gene aus Grafik A zeigte, dass 8 der 12 exprimierten Proteine miteinander interagieren.

Der Vergleich der Top 20 Gene ermittelte 12 gemeinsame Gene, die in den Mutationsanalysen sowohl ein hohes biologisches als auch epigenetisches Alter aufwiesen (siehe Abbildung 51A). Als Top Gen der zwölf gemeinsamen Gene wurde mit dem höchsten medianen Alter der mutierten Patienten *CD58* ermittelt. Das *CD58*-Gen wurde in dieser Arbeit bereits als Kandidatengenen für genetische Aberrationen bei senilen DLBCL-Patienten identifiziert (siehe Abschnitt 3.2.1.2) und ist ein Adhäsionsprotein, das eine wichtige Rolle in der Immunantwort des menschlichen Organismus spielt. *CD58* wird von Antigen-präsentierenden Zellen (APC) u.a. B-Zellen exprimiert und bindet an das Oberflächenprotein *CD2*, woraufhin es über eine komplexe Signalkaskade zur Aktivierung der T-Zellen kommt. Eine anschließende Pathway-Analyse der zwölf gemeinsamen Gene mit dem webbasierten Tool String 10 zeigte, dass acht der zwölf exprimierten Proteine miteinander interagierten (siehe Abbildung 51B). Neben den gemeinsamen Genen wurden jeweils 8 Gene identifiziert, die nur in der Gruppe nach dem biologischen oder epigenetischen Alter auftraten (Abbildung 51A).

Zusammenfassend konnten 12 Gene ermittelt werden, die sowohl nach dem biologischen als auch epigenetischen Alter vorwiegend in senilen DLBCL-Patienten rekurrent mutiert vorliegen. Das Gen mit dem höchsten medianen biologischen und epigenetischen Alter der mutierten Patientengruppen war *CD58*.

3.2.2.5 Rekurrent mutierte Gene die das epigenetischen Alter bei Patienten mit einem diffus großzelligen B-Zell-Lymphom beeinflussen

In der weiteren Auswertung der rekurrent mutierten Gene in DLBCL-Patienten wurden Gene identifiziert, die potenziell das epigenetische Alter der betroffenen Patienten beeinflussen (sog. „epigenetic modifiers“). Hierzu wurden die rekurrent mutierten Gene in senilen DLBCL-Patienten hinsichtlich des biologischen Alters und des epigenetischen Alters verglichen und die Differenz der beiden Altersangaben gebildet, die im Folgenden als Delta-Alter bezeichnet wurde. Das Delta-Alter wurde analog zum vorherigen Kapitel 3.2.2.5 für alle 49 DLBCL-Proben berechnet zu denen Sequenz- und 450K-Daten vorlagen. Die 50 rekurrent mutierten Gene in DLBCL (siehe Abschnitt 3.2.1.2) wurden erneut in Form von Boxplots dargestellt und in diesem Fall nach dem Delta-Alter der entsprechend mutierten Patientengruppe geordnet (Abbildung 52). In den rekurrent mutierten Genen (N= 50) bei DLBCL lag das mediane Delta-Alter der mutierten Patienten bei durchschnittlich 19 Jahren. Das höchste mediane Delta-Alter wurde für DLBCL-Patienten mit einer potentiell proteinverändernden Mutation in den Genen *IRF4* (Median 47,5 Jahre), *CARD11* (Median: 40 Jahre), *ZFP36L1* und *MYD88* (jeweils Median: 39 Jahre) ermittelt. Das geringste mediane Delta-Alter wurde für DLBCL-Patienten mit einer potentiell proteinverändernden Mutation in den Genen *TET1* (Median:-1 Jahr), *PAX5* (Median: 1,5 Jahre) und *ARID1A* (Median: 4 Jahre) ermittelt.

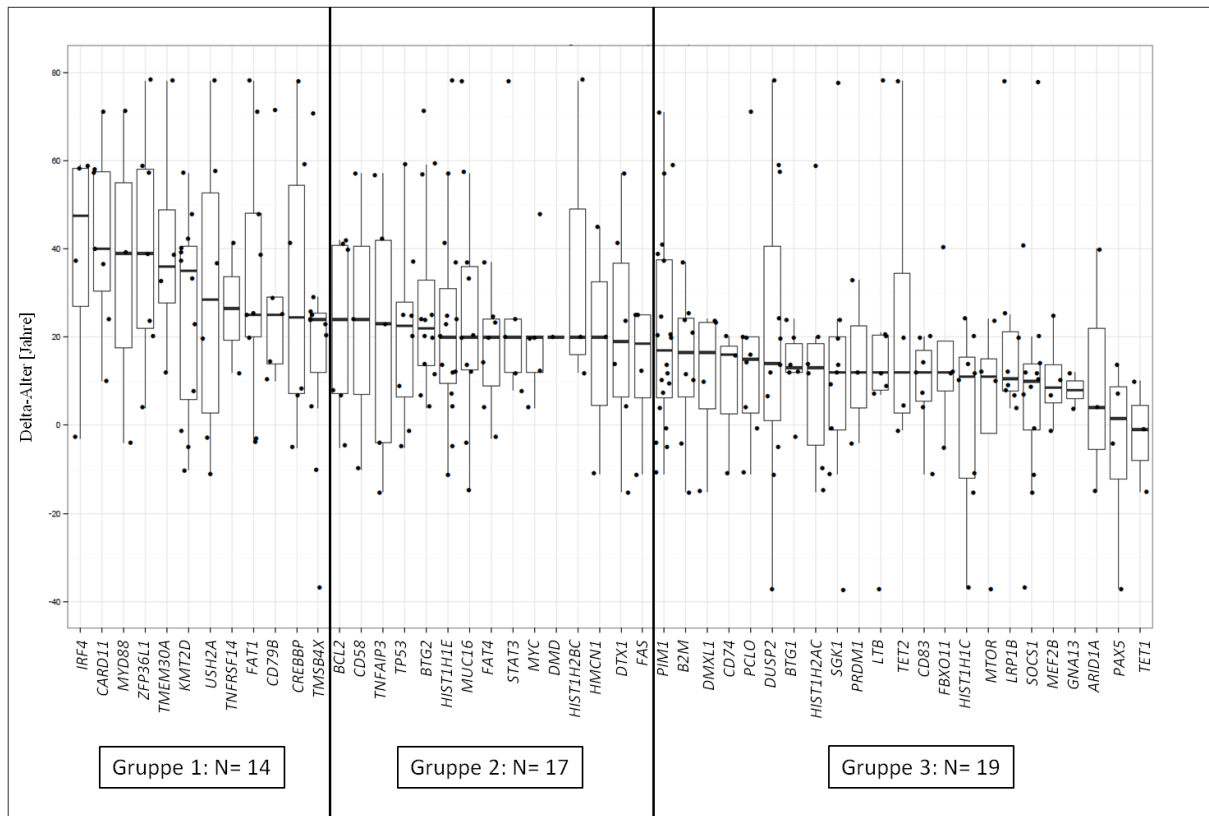


Abbildung 52: Altersverteilung der rekurrent mutierten Gene in DLBCL nach dem Delta-Alter

In der Grafik wurden alle 50 rekurrent mutierten Gene in den 49 analysierten DLBCL-Patienten gegen die Differenz aus dem biologischen und epigenetische Alter (Delta-Alter) aufgetragen. Die Gene *IRF4*, *CARD11*, *MYD88* und *ZFP36L1* zeigten das höchste mediane Delta-Alter in der Gruppe von Patienten mit einer Mutation. Hingegen wurde bei den Genen *ARID1A*, *PAX5* und *TET1* das geringste mediane Delta-Alter in der Gruppe von Patienten mit einer Mutation identifiziert.

In der weiteren Auswertung der Mutations-Ergebnisse in DLBCL im Zusammenhang mit dem Delta-Alter der Patienten wurde die Standardabweichung über die medianen Alter der mutierten Patienten über alle Gene ermittelt (Details siehe Abschnitt 2.2.8.2). Anschließend wurden die rekurrent mutierten Gen in drei Gruppen eingeteilt (siehe Abbildung 52): Gruppe 1: hohes Delta-Alter (d.h. das epigenetische Alter ist deutlich höher als das biologische Alter), Gruppe 2: mittleres Delta-Alter (d.h. das epigenetische und biologische Alter sind ähnlich) und Gruppe 3: geringes Delta-Alter (d.h. das epigenetische Alter ist deutlich geringer als das biologische Alter). Da der Fokus auf der Identifizierung von Genen lag, die das epigenetische Alter beeinflussen, wurden nachfolgend ausschließlich die Gruppen 1 und 3 genauer betrachtet. Insgesamt wurden zwei Gene aus der Gruppe 1 (*CREBBP*, *KMT2D*) bzw. vier Gene aus der Gruppe 3 (*HIST1H2AC*, *HIST1H1C*, *TET1*, *TET2*) identifiziert, deren Funktionen bereits mit epigenetischen Aberrationen assoziiert wurden (Abbildung 53A). Die Pathway-Analyse der Gene aus der Gruppe 1 zeigte, dass 7 der 14 exprimierten Proteine (Gruppe 1) miteinander interagierten (siehe Abbildung 53B). Bei den Genen aus der Gruppe 3 wurden keine komplexen Interaktionen zwischen den 19 exprimierten Proteinen ermittelt (Abbildung 53C).

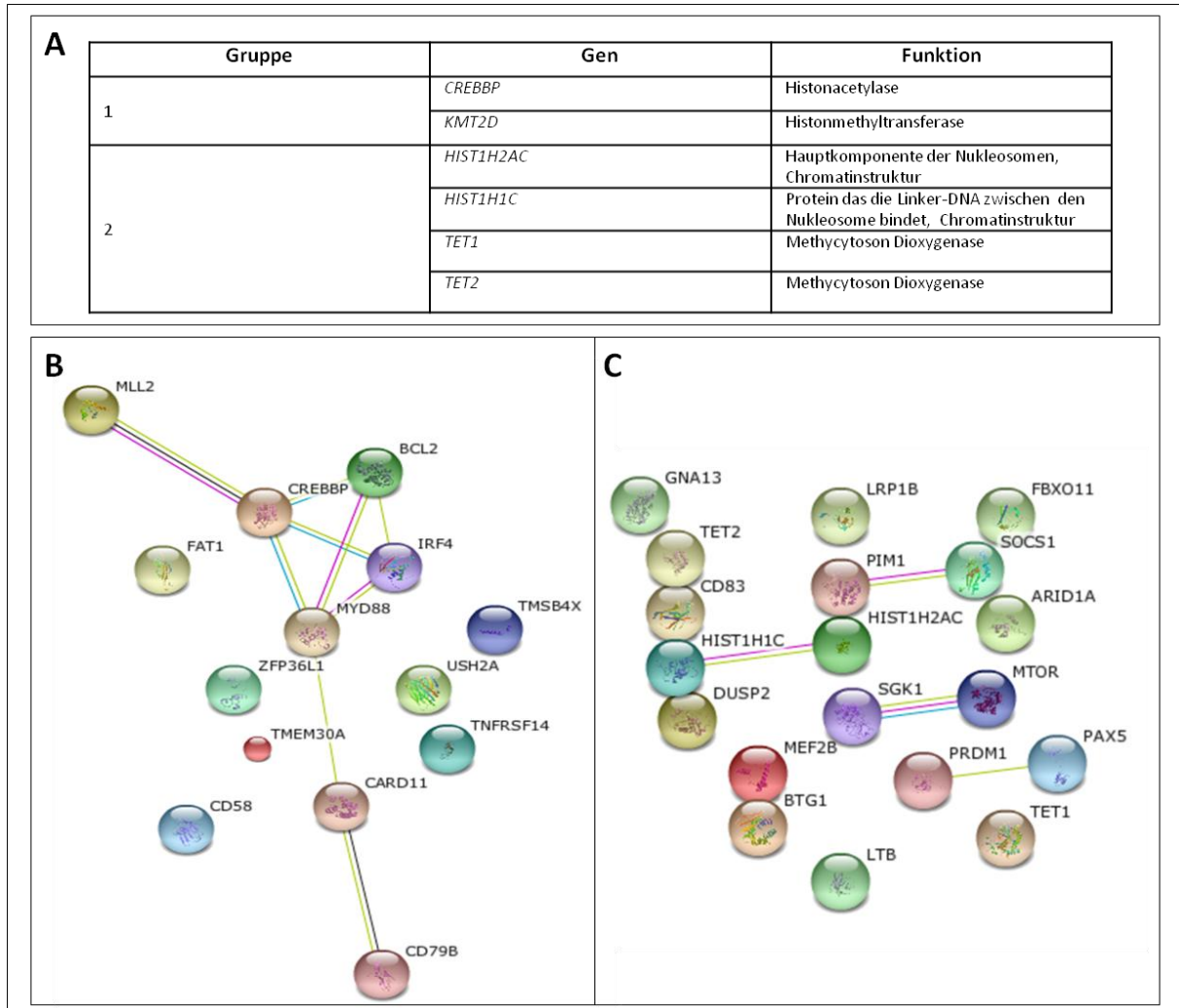


Abbildung 53: Informationen und Stringanalyse zu Genen die das epigenetische Alter in DLBCL-Patienten beeinflussen. A: In der Gruppe 1 wurden die Gene *CREBBP* und *KMT2D* bzw. in der Gruppe 2 die Gene *HIST1H2AC*, *HIST1H1C*, *TET1* und *TET2* identifiziert, deren Funktionen bereits mit epigenetischen Aberrationen assoziiert wurde. B: Die Stringanalyse der Gene aus Gruppe 1 zeigte, dass 7 der 14 exprimierten Proteine miteinander interagieren. C: Die Stringanalyse der Gene aus Gruppe 3 zeigte, dass keine komplexen Interaktionen zwischen den 19 exprimierten Proteinen erkennbar waren.

Zusammenfassend wurden drei Gruppen von Genen in dem Vergleich der Mutationsanalyse der rekurrent mutierten Gene nach dem biologischen und epigenetischen Alter der mutierten Patienten identifiziert. Es wurden Gene ermittelt, deren mutierte Patienten eine ähnliche Differenz zwischen dem biologischen und epigenetischen Alter aufwiesen (Gruppe 2). Außerdem wurden sog. „epigenetic modifiers“ identifiziert d.h. Gene die möglicherweise das epigenetische Alter der DLBCL-Patienten beeinflussen. Bei diesen „epigenetic modifiers“ war das epigenetische Alter der DLBCL-Patienten entweder deutlich höher (Gruppe 1) oder deutlich geringer (Gruppe 3) als das biologische Alter. Insgesamt konnten sechs Gene identifiziert werden von denen bereits bekannt ist, dass sie eine wichtige Rolle im Zusammenhang mit epigenetischen Modifikationen spielen können (Gruppe 1: *CREBBP*, *KMT2D* und Gruppe 2: *HIST1H2AC*, *HIST1H1C*, *TET1*, *TET2*).

4 Diskussion

Das Ziel der vorliegenden Arbeit war die Identifizierung von genetischen und epigenetischen Aberrationen bei Patienten mit einem Keimzentrums-B-Zell-Lymphom im Senium. Zur Identifizierung der genetischen Aberrationen auf Mutationsebene wurde zu Beginn der Arbeit das Verfahren der Hochdurchsatz-Sequenzierung am Institut für Humangenetik optimiert. Im Rahmen der Methoden-Optimierung konnte in dieser Arbeit an Frisch- und FFPE-Material ein Hybridisierungs-basiertes Custom Panel zur Analyse von rekurrent mutierten Genen in Lymphomen mittels Hochdurchsatz-Sequenzierung optimiert werden. Anschließend wurden Tumor-Proben von Patienten mit einem hohen Ersterkrankungsalter (≥ 80 Jahre) mittels Fluoreszenz-in-situ-Hybridisierung (FISH), Immunhistochemie (IHC), Hochdurchsatz-Sequenzierung (NGS) und Infinium Human Methylation 450K BeadChip (450K) analysiert und die erhobenen Daten der senilen Patienten mit bereits existierenden molekulargenetischen Daten aus verschiedenen klinischen und Forschungsstudien verglichen. Es folgte eine statistische Auswertung der Assoziation von identifizierten genetischen und epigenetischen Markern bei Patienten mit einem Keimzentrums-B-Zell-Lymphom gegenüber dem Ersterkrankungsalter der Patienten. Insgesamt wurden zwölf Gene identifiziert, die vorwiegend in DLBCL-Patienten im hohen Alter mutiert vorlagen. Als Top Kandidatengen für genetische Aberrationen bei senilen DLBCL-Patienten wurde das *CD58*-Gen ermittelt. Zusätzlich wurden Gene identifiziert die das epigenetische Alter der DLBCL-Patienten beeinflussen (sog. „epigenetic modifiers“), wobei für einzelne dieser Gene bereits eine Assoziation mit Histon- oder Chromatinmodifikationen bekannt war (z.B. *TET2*, *CREBBP*).

Im folgenden Teil sollen die im Rahmen dieser Arbeit erzielten Ergebnisse in den Zusammenhang mit dem gegenwärtigen Stand der Wissenschaft gestellt werden. Dazu werden zuerst die wichtigsten technischen Aspekte der Methoden-Optimierung und anschließend die inhaltlichen Aspekte der Identifizierung der altersassoziierten molekulargenetischen Aberrationen der Patienten mit einem Keimzentrums-B-Zell-Lymphom genauer betrachtet.

4.1 Methodische Aspekte zur Probenvorbereitung für die Hochdurchsatz-Sequenzierung

Im Rahmen der vorliegenden Arbeit wurde zuerst die Methode der Hochdurchsatz-Sequenzierung am Institut für Humangenetik optimiert. Das Ziel im Bereich der Methodik war die Identifizierung und Optimierung von Parametern in der Probenvorbereitung, die zur Verbesserung der Sequenzier-Ergebnisse führen. Insbesondere sollten die Fragen beantwortet werden, ob es möglich ist FFPE-Proben mittels Hochdurchsatz-Sequenzierung zu analysieren und welchen Einfluss die verwendeten Polymerasen in der Probenvorbereitung auf die abschließenden Sequenzier-Daten haben. Zur Bearbeitung dieser Fragestellungen wurden in der vorliegenden Arbeit verschiedene Target-Resequenzierungen mittels PCR- und Hybridisierungs-basierter kommerzieller und Custom Panel unter der Verwendung des MiSeq-Geräts durchgeführt. Die Verifizierung der identifizierten Mutationen erfolgte in dieser Arbeit mittels Sanger-Sequenzierung. Zur Optimierung der Probenvorbereitung wurden verschiedene Parameter z.B. Input, Normalisierung oder Sequenzier-Kit-Komponenten variiert, um möglichst viele Sequenzen mit bestmöglicher Qualität zu generieren. Im Rahmen der Austestung der einzelnen Panel/Assay wurden in dieser Arbeit insbesondere drei Charakteristika identifiziert, die einen Einfluss auf die Qualität der Sequenzier-Ergebnisse haben: das Ausgangsmaterial, die Polymerasen und die Sequenzierung.

Die Ergebnisse der vorliegenden Arbeit zeigen, dass die Hochdurchsatz-Sequenzierung von FFPE-Material sowohl mittels PCR- als auch Hybridisierungs-basierten Panel/Assays möglich ist. Jedoch ist ein entscheidender Faktor für optimale Ergebnisse der Hochdurchsatz-Sequenzierung die Qualität des Ausgangsmaterials. Es ist bekannt, dass die Analysen von FFPE-Material im Gegensatz zu Frischmaterial häufig mit einer höheren Fehlerrate in der Probenvorbereitung und Sequenzierung sowie einer geringeren durchschnittlichen On-Target-Rate (\emptyset -OTR) und On-Target-Coverage (\emptyset -OTC) assoziiert sind, was vermutlich auf Deaminierungs-Artefakte während der DNA-Extraktion zurückzuführen ist²⁵⁶. Aufgrund dieser Problematik sind viele kommerziell erhältliche Panel nur für die Anreicherung von extrahierter DNA aus Frischmaterial ausgelegt und nicht FFPE-geeignet. In der vorliegenden Arbeit wurden zwei Custom-Panel/Assay (PCR-basierter *ID3*-Assay: siehe Abschnitt 3.1.2 und Hybridisierungs-basiertes Lymphom-Panel: siehe Abschnitt 3.1.4) optimiert, die es ermöglichen eine Mutationsanalyse an Frisch- und FFPE-Proben durchzuführen. Insbesondere die Ergebnisse der Sequenzierung der Lymphom-Panel-Proben zeigen jedoch, dass die Coverage der FFPE-Proben von der DNA-Qualität abhängig ist (siehe Abschnitt 3.1.4). An degradierte DNA-Fragmente können die spezifischen Oligonukleotide schlechter hybridisieren, weswegen folglich keine Library Präparation und anschließende Sequenzierung der degradierten DNA-Fragmente möglich ist.

Die sequenzierte Fragmentlänge von extrahierter DNA aus FFPE-Material sollte bei max. 200 bp liegen, da die Probenvorbereitung und Sequenzierung von längeren DNA-Fragmenten aufgrund der Degradierung der DNA aus FFPE-Material häufig problematisch ist und nicht funktioniert²⁵⁷. Zudem wird davon ausgegangen, dass die chemische Behandlung des FFPE-Materials mittels Rotihistol zur Deparaffinierung, verstärkt zur Degradierung der DNA und somit zur schlechteren Qualität der DNA führt.

Die Hochdurchsatz-Sequenzierung ist ein hoch sensitives Verfahren, das alle Fehler die während der Amplifikation in der Sequenzabfolge eingebaut werden, detektiert. Demnach ist es wichtig in der Library Präparation eine Polymerase zu nutzen, die eine hohe Ausbeute bei verschiedenen Target-Regionen erzielt und eine geringe Fehlerrate aufweist. Die kommerziellen und Custom Panel die von einer Firma entworfen werden, enthalten im Library Präparation Kit bereits ausgetestete Polymerasen. Für die Erstellung des Custom-*ID3*-Assay sollte eine Austestung verschiedener Polymerasen erfolgen, um die Anzahl an Polymerase- und Sequenzierartefakten möglichst gering zu halten. Hierzu wurden in den Testversuchen verschiedene Polymerase Systeme an Frisch- und FFPE-Material eingesetzt. Die Ergebnisse der PCR-Amplifikationen, Hochdurchsatz-Sequenzierungen sowie die ermittelten Werte der durchschnittlichen On-Target-Rate (\emptyset -OTR) der einzelnen Polymerasen und Testproben zeigten, dass die Probenvorbereitung mittels Accu Prime Polymerase für die *ID3*-Assay-Proben am erfolgreichsten war und die meisten Sequenzierungs-Daten unter Verwendung dieses Polymerase Systems erzielt wurden.

Bei dem Prozess der Sequenzierung ist insbesondere die Analyse von Regionen in repetitiven Sequenzen, Homopolymerbereichen und GC-reichen Regionen problematisch und in derartigen Regionen können häufig keine oder nur unspezifische PCR-Produkte generiert und sequenziert werden²⁵⁸⁻²⁶⁰. Bei jeder Hochdurchsatz-Sequenzierung sollte unabhängig von dem analysierten Panel/Assay ein möglichst hoher Output an Reads erzielt werden, was eine optimale Cluster Density erfordert. Eine zu hohe Cluster Density führt zu einer schlechten Qualität der Reads und einem geringen Output²⁶¹. Alle Panel/Assay wurden in dieser Arbeit zuerst nach dem Standard-Protokoll von Illumina normalisiert. Zusätzlich wurde in dieser Arbeit ein *Small-Scale-Protokoll* zur Normalisierung von Libraries optimiert, wodurch der Output an Reads für einzelne Panel-Proben erhöht wurde (siehe Abschnitt 3.1). Die Spezifität und Sensitivität der getesteten Assay/Panel wurden in dieser Arbeit aus dem Vergleich der Hochdurchsatz- und Sanger-Sequenzierungsdaten ermittelt und lagen jeweils bei 100%. Zudem wurde beim *ID3*-Assay die Reproduzierbarkeit des Assays durch die Sequenzierung von Duplikaten überprüft, wobei das Bestimmtheitsmaß bei 0,996 lag.

Insgesamt konnten im Rahmen der vorliegenden Arbeit vier verschiedene PCR- und Hybridisierungs-basierte kommerzielle und Custom Panel optimiert und die Methode der Hochdurchsatz-Sequenzierung etabliert werden, womit die Aufgabenstellung erfolgreich bearbeitet wurde. Zur Probenvorbereitung von FFPE-Material konnten in dieser Arbeit insbesondere zwei Custom Panel/Assay erfolgreich optimiert: *ID3*-Assay und Lymphom-Panel .

Der *ID3*-Assay ermöglicht eine Mutationsanalyse im Exon 1, der am häufigsten mutierten Region bei molekularen Burkitt-Lymphome (mBL). Beispielsweise stellen mBL die häufigste Form von Lymphomen im Kindesalter dar und in den letzten Jahren konnte u.a. im Rahmen des ICGC MMML-Seq Projektes gezeigt werden, dass 68% aller mBL eine Mutation im *ID3*-Gen (Inhibitor of DNA binding 3) aufweisen und diese nicht in nicht-molekularen Burkitt Lymphomen (non-mBL) und nur extrem selten in intermediären Lymphomen vorkommt²³⁰⁻²³². Dadurch können Mutationen dieses Gens zu einem potentiellen molekularen Marker zur Diagnostik von Burkitt-Lymphomen werden und zur Unterscheidung von anderen Lymphom-Entitäten dienen. Im Hinblick auf die wachsende Bedeutung der Panel-Diagnostik mittels Hochdurchsatz-Sequenzierung in der Krankenversorgung bietet dieser *ID3*-Assay möglicherweise eine schnelle und kostengünstige Alternative zur Routine-Diagnostik von mBL in der Zukunft.

Das Lymphom-Panel ist ein Custom Panel, dass die Mutationsanalyse an Frisch- und FFPE-Material von rekurrent mutierten Genen in Lymphomen mittels Hochdurchsatz-Sequenzierung ermöglicht. Dieses Lymphom-Panel konnte in der vorliegenden Arbeit für die Mutationsanalyse der sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom erfolgreich angewandt werden.

Zu Beginn dieser Arbeit wurde die DNA von FFPE-Material am Institut für Humangenetik standardmäßig mit dem QIAamp DNA FFPE Tissue Kit der Firma Qiagen isoliert, weshalb dieses Kit auch für die DNA-Extraktion der FFPE-Proben in der vorliegenden Arbeit verwendet wurde (basiert auf Deparaffinierung mittels Rotihistol, siehe Abschnitt 2.2.1.2). Aufgrund der zunehmenden Bedeutung der Analyse von FFPE-Material in der Forschung und Klinik wurde inzwischen von verschiedenen Firmen nach alternativen Methoden zur Verbesserung der DNA-Extraktion von FFPE-Material geforscht. Eine neue Methode zur Extraktion von DNA aus FFPE-Material wurde insbesondere von den Firmen Covaris und Diagnode entwickelt und basiert auf der Deparaffinierung mittels Ultraschall. Seitens Covaris und Diagnode besteht die Hypothese, dass die Extraktion von DNA durch mechanische Kräfte wie ultraschall deutlich „schonender“ für die DNA ist als eine chemische Behandlungen durchzuführen (z.B. mittels Rotihistol). Die extrahierte DNA aus FFPE-Material mittels Ultraschall soll demnach weniger stark degradiert sein^{262,263}. Parallel dazu entwickelte die Firma Qiagen das neue GeneRead DNA FFPE Kit (chemische DNA-Extraktion; kein Ultraschall) unter Verwendung des Enzyms Uracil-N-Glycosylase, wodurch die Deaminierungsartefakte bei der DNA-Extraktion aus FFPE-Material verringert werden sollen^{256,264}.

Erste Austestungen der neuen Extraktions-Kits GeneRead DNA FFPE Kit (Qiagen) und truXTRAC FFPE DNA Kit (Covaris) im Vergleich mit dem bisher verwendeten QIAamp DNA FFPE Tissue Kit (Qiagen) zeigen, dass die Qualität der extrahierten DNA aus FFPE-Material mittels der neuen Kits deutlich besser ist (interne Daten unter Verwendung des Infinium HD FFPE QC Assay der Firma Illumina, Daten in der vorliegenden Arbeit nicht gezeigt). Aufgrund dieser neuen Ergebnisse werden inzwischen am Institut für Humangenetik und am Institut für Pathologie/ Sektion Hämatopathologie für die DNA-Extraktion aus FFPE-Proben das GeneRead DNA FFPE Kit (Qiagen) oder truXTRAC FFPE DNA Kit (Covaris) verwendet. Es ist davon auszugehen, dass die verbesserte DNA-Qualität zukünftig auch zur Verbesserung der Qualität der Sequenzdaten von FFPE-Proben führt und in den Hochdurchsatz-Sequenzierungen eine höheren Ø-OTR und Ø-OTC erzielt wird (Testversuche sind in der Bearbeitung).

Zu Beginn der vorliegenden Arbeit waren noch keine Panel verfügbar, die explizit für die Probenvorbereitung von FFPE-Material entwickelt wurden. Die damals erhältlichen Panel und der angebotene Service zur Erstellung von Custom Panel beinhalteten vorwiegend die Analyse von Frischmaterial. Die Erstellung von größeren Hybridisierungs-basierten Custom Panel (> 500 Mb) wurde zu diesem Zeitpunkt insbesondere von der Firma Agilent angeboten, wobei die Verwendung von FFPE-Material als Ausgangsmaterial ausgetestet werden musste, was ein Ziel der vorliegenden Arbeit war. Die Sequenzier-Ergebnisse der Lymphom-Panel-Proben in dieser Arbeit zeigen, dass es möglich ist unter Verwendung des SureSelectXT Verfahrens der Firma Agilent FFPE-Proben mittels Hochdurchsatz-Sequenzierung zu analysieren. Jedoch ist in den Daten auch erkennbar, dass nur etwa 50% der Target-Regionen des Lymphom-Panels in die abschließende Analyse eingingen. Über die weiteren 50% der Target-Regionen konnte in den Analysen dieser Arbeit keine Aussage zum Mutationsstatus in Keimzentrums-B-Zell-Lymphomen vorgenommen werden. Verschiedene Firmen forschen derzeit intensiv im Bereich der Probenvorbereitung von FFPE-Material zur weiteren Analyse mittels Hochdurchsatz-Sequenzierung. Inzwischen gibt es ein PCR-basiertes Verfahren der Firma Illumina, das explizit zur Analyse von FFPE-Proben entwickelt und auf den Markt gebracht wurde. Besonders ist hierbei, dass die Library Präparation von jeder Probe zweifach in separaten Ansätzen (Pool A und B) erfolgt, wodurch die Anzahl an Artefakten minimiert werden soll. Zudem entwickelt die Firma Illumina aktuell ein Hybridisierungs-basiertes Verfahren zur Analyse von FFPE-Material, das noch im Jahre 2016 auf den Markt kommen soll. Mit Hilfe dieser neuen FFPE-geeigneten Verfahren kann möglicherweise ein neues Panel erstellt werden, das die Target-Regionen des Lymphom-Panels beinhaltet und wodurch die Ø-OTR verbessert werden kann. Eventuell ist es mit den neuen Sequenzdaten möglich den Mutationsstatus von allen analysierten Target-Regionen zu ermitteln. Die Ergebnisse der Mutationsanalysen der sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom dieser Arbeit im Vergleich zu den bereits publizierten Daten anderer Arbeitsgruppen werden im nachfolgenden Abschnitt 4.2 genauer betrachtet und diskutiert.

4.2 Senile Patienten mit einem reifzelligem Keimzentrums-Lymphom

DLBCL und FL sind in Zentraleuropa mit bis zu 80% die häufigsten Lymphome im Erwachsenenalter und Senium⁸⁹. Die Inzidenz von Keimzentrums-B-Zell-Lymphomen hat in den letzten Jahren kontinuierlich zugenommen und die aktuellen SEER-Daten lassen vermuten, dass die Anzahl an Patienten mit einem Keimzentrums-B-Zell-Lymphom auch in den nächsten Jahren stetig zunimmt²⁶⁵. Da zusätzlich die Lebenserwartung der Bevölkerung in den westlichen Ländern weiter ansteigt und mittlerweile in Deutschland bei etwa 80 Jahren liegt (Stand 2011, Robert Koch Institut, Deutschland, Berlin²⁶⁶), nimmt die Behandlung von senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom an Bedeutung in der onkologischen Versorgung zu. In der aktuellen Literatur gibt es keine exakte Definition für die Altersgruppe des „Seniums“, sodass in dieser Arbeit eine Extremdefinition gewählt wird und alle Patienten mit einem Ersterkrankungsalter von mind. 80 Jahren als „senil“ bezeichnet werden^{166,167}. An dieser Stelle wird noch einmal explizit drauf hingewiesen, dass der Begriff „senil“ in der vorliegenden Arbeit ausschließlich auf das Ersterkrankungsalter der Patienten von mind. 80 Jahren basiert und mit keiner geistigen oder körperlichen Gebrechlichkeit bzw. anderen „Altersschwächen“ der sehr alten Patienten assoziiert wird. In den letzten Jahren wurden in verschiedenen klinischen Studien (z.B. RICOVER-60^{118,226}, NHLB^{116,117}) und Forschungsprojekten (MMML, ICGC MMML-Seq) Patienten in Deutschland mit Keimzentrums-B-Zell-Lymphomen molekulargenetisch analysiert. Jedoch wurden in diesen Studien vorwiegend Patienten bis zu einem Ersterkrankungsalter von 80 Jahren molekulargenetisch charakterisiert, wodurch die senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom in den bisherigen Analysen völlig unterrepräsentiert sind. Da die Behandlung von senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom in der onkologischen Versorgung an Bedeutung zunimmt, das vorliegende Wissen über die molekularen Eigenschaften der Erkrankung bei älteren Patienten jedoch gering ist, lag der Fokus der vorliegenden Arbeit in der Analyse von altersassoziierten molekulargenetischen Veränderungen bei senilen Patienten mit einem aggressiven Keimzentrums-B-Zell-Lymphom: diffus großzellige B-Zell-Lymphome (DLBCL) und folliculäre Lymphome (FL). Hierzu wurde gezielt eine Kohorte von senilen Patienten zusammengestellt und molekulargenetisch charakterisiert. Anschließend wurden die Tumor-Proben von verschiedenen Altersgruppen hinsichtlich chromosomaler Imbalancen, der Expression von spezifischen Proteinen, Mutationen und dem DNA-Methylierungsmuster analysiert. In den altersassoziierten Analysen wurden genetische und epigenetische Aberrationen identifiziert, die vorwiegend bei Patienten im hohen Alter auftreten und als mögliche Biomarker für eine altersangepasste Therapie bei senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom fungieren können. Die Hoffnung besteht darin, aus diesen Analysen neue Möglichkeiten für eine zielgerichtete Therapie aufzuzeigen, die für Patienten des höheren Alters verträglich ist.

4.2.1 Lymphom-Entitäten und Geschlechterverteilung

Zur Beurteilung der Repräsentativität der analysierten Patienten-Kohorte gegenüber Patienten mit einem Keimzentrums-B-Zell-Lymphom in den westlichen Ländern wurden zuerst die Verteilungen der Lymphom-Entitäten und Geschlechter der analysierten Kohorte der vorliegenden Arbeit mit bereits publizierten Daten verglichen.

Die erhobenen Daten der vorliegenden Arbeit zeigen, dass in der neu zusammengestellten Kohorte von senilen Patienten größtenteils DLBCL und nur wenige FL bei Patienten im hohen Alter diagnostiziert wurden. Wie bereits in der Einleitung im Kapitel 1.5 dargestellt, verändert sich in den westlichen Ländern die Inzidenz der Lymphom-Entitäten mit dem Ersterkrankungsalter der Patienten. So zeigen die aktuellen SEER-Daten, dass ab dem 65. Lebensjahr die Inzidenz von FL konstant bleibt, wohingegen die Anzahl an DLBCL bis ins hohe Alter stetig zunimmt¹⁷⁰, was durch die Ergebnisse der vorliegenden Arbeit bestätigt wurde.

Die Betrachtung der analysierten DLBCL-Kohorte in der vorliegenden Arbeit (N= 1609) veranschaulichten, dass etwa 75% der Patienten ein Ersterkrankungsalter von über 60 Jahren aufwiesen (60-79 Jahre: 65% und ≥ 80 Jahre: 9%). Insgesamt zeigt der Vergleich der Geschlechter- und Altersverteilung der analysierten sehr alten und senilen Patienten dieser Arbeit mit den Kohorten aus bereits veröffentlichten Arbeiten, dass die analysierte Patienten-Kohorte in der vorliegenden Arbeit repräsentativ für die Verteilung von Patienten mit einem Keimzentrums-B-Zell-Lymphom der westlichen Länder ist. In der Studie von Williams et al. wurden 4635 DLBCL-Patienten mit einem Ersterkrankungsalter von mind. 66 Jahre zwischen 2002 und 2009 analysiert und in zwei Gruppen eingeteilt: 66-80 Jahre (75% der Kohorte) und >80 Jahre (25% der Kohorte). Ein weiterer Vergleich der senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom aus verschiedenen Studien gegenüber der analysierten Patienten-Kohorte in der vorliegenden Arbeit hinsichtlich klinischer Daten wie z.B. die Überlebensrate oder Dosierung ist nicht möglich, da zu den senilen Patienten der vorliegenden Arbeit keine klinischen Informationen vorliegen.

Die Daten der analysierten FL-Kohorte dieser Arbeit lassen vermuten, dass FL von jüngeren Patienten unter 40 Jahren häufiger beim weiblichen Geschlecht vorkommen und mit zunehmendem Alter bei mehr Männern ein FL diagnostiziert wird. Die Geschlechterverteilung der gesamten analysierten FL-Kohorte in dieser Arbeit (Median: 60 Jahre) liegt bei etwa 1:1 (siehe Abschnitt 3.2.1.1). Hingegen beschreibt die aktuelle Literatur, dass das Verhältnis von Frauen und Männern etwa 1,7:1 beträgt und das mediane Alter von FL-Patienten in den westlichen Ländern bei etwa 59 Jahren liegt^{89,267}.

Ein altersassoziierter Vergleich der Geschlechterverteilung bei FL-Patienten mit bereits publizierten Daten ist nicht möglich, da es zurzeit keine exakten Daten aus veröffentlichten Studien zu verschiedenen Altersgruppen von FL-Patienten gibt.

Die Geschlechterverteilung in der neu erstellten DLBCL-Kohorte dieser Arbeit (Frauen: 63% und Männer: 37%) zeigen, dass senile Patienten häufiger Frauen als Männer sind. Hierbei ist jedoch zu beachten, dass die Lebenserwartung der Frauen in Deutschland etwa fünf Jahre höher ist als von Männern (Stand 2011, Robert Koch Institut, Deutschland, Berlin²⁶⁶). Es ist demnach nicht auszuschließen, dass die Geschlechterverteilung der senilen Patienten mit einem Keimzentrums-B-Zell-Lymphom durch einen biologischen Effekt beeinflusst wird und nicht ausschließlich ein lymphom-spezifisches Phänomen darstellt. Der Vergleich der Geschlechterverteilung der senilen Patienten der Kohorten aus den Veröffentlichungen von Peyrade et al. (Frauen: 66%¹¹¹) und Williams et al. (mehr Frauen, keine genauen Zahlen angegeben²⁶⁸) bestätigt die Geschlechterverteilung in der analysierten Kohorte der vorliegenden Arbeit.

Nachdem gezeigt werden konnte, dass die analysierte Patienten-Kohorte in der vorliegenden Arbeit repräsentativ für die Verteilung von Patienten mit einem Keimzentrums-B-Zell-Lymphom der westlichen Länder ist, sollte die Fragestellung beantwortet werden, ob es Unterschiede in den genetischen Aberrationen zwischen jungen und älteren Patienten mit einem Keimzentrums-B-Zell-Lymphom gibt.

Die Ergebnisse der FISH-Analysen dieser Arbeit zeigen, dass chromosomale Imbalancen im *BCL2*-Gen bei FL signifikant häufiger bei jüngeren Patienten auftreten und mit dem Ersterkrankungsalter abnehmen. Diese Erkenntnis führt zu der Annahme, dass die Translokation t(14;18) signifikant häufiger bei jüngeren FL-Patienten vorkommt. Ein direkter Vergleich der erhobenen Daten der analysierten senilen Patienten dieser Arbeit mit publizierten Daten hinsichtlich t(14;18) ist nicht möglich, da unterschiedliche Sonden für die FISH-Analysen verwendet wurden und teilweise nur separate FISH-Daten für die chromosomalen Imbalancen der Gene *IGH* und *BCL2* vorliegen. Aufgrund der Häufigkeit der t(14;18) bei FL ist jedoch davon auszugehen, dass vorwiegend das *IGH*-Gen den Translokationspartner von *BCL2* darstellt. Die bereits publizierten Daten für DLBCL-Patienten der westlichen Länder zeigen, dass das Verhältnis von Männern und Frauen etwa 1,6:1¹⁷⁰ beträgt und das mediane Alter der betroffenen Patienten bei 64 Jahren liegt^{91,100}. Bei den Keimzentrums-Lymphomen des Erwachsenenalters ist die Zuordnung der potentiellen molekularen Marker zu den Entitäten FL und DLBCL etwas schwieriger, da für Lymphome in einigen Arbeiten keine exakte Aufteilung entsprechend der Entitäten vorliegt, sondern die Gesamt-Kohorte als Non-Hodgkin-Lymphome betrachtet wurde⁶². Nichtsdestotrotz geht aus der Literatur hervor, dass bei 70-80% aller FL die Translokation t(14;18) nachweisbar ist, wohingegen nur 30-50% der DLBCL diese Translokation besitzen und diese zumeist dem GCB-Subtyp angehören^{143,172,269,270}.

In den Mutationsanalysen der vorliegenden Arbeit wurden nur Gene als rekurrent mutiert bezeichnet für die eine MF von mind. 10% ermittelt wurde. Insgesamt wurden 9 Gene als rekurrent mutiert in FL identifiziert, wovon fünf als m7-FLIPI-Gene (*EZH2*, *ARID1A*, *MEF2B*, *EP300*, *CREBBP*) identifiziert wurden. Als m7-FLIPI werden sieben Gene bezeichnet (*EZH2*, *ARID1A*, *MEF2B*, *EP300*, *FOXO1*, *CREBBP*, *CARD11*) die Pastore et al. in ihren Mutationsanalysen von 74 Genen bei 151 FL-Patienten als prognostisch relevante Marker identifizierten und in einem klinischen Risikomodell zusammenfassten²⁷¹. In m7-FLIPI konnte gezeigt werden, dass mit Hilfe des Mutationsprofils dieser sieben Gene in Assoziation mit dem *Follicular Lymphoma International Prognostic Index* (FLIPI) eine Prognose für den Therapieverlauf und das Überleben des Patienten ermittelt werden kann und das klinische Risikomodell zukünftig eventuell als Option zur Berechnung bzw. Strategie einer Risikoadaptierten und personalisierten Therapie bei FL-Patienten verwendet werden kann¹⁰⁴. Da in der vorliegenden Arbeit keine klinischen Daten für die neu erstellte Kohorte von senilen Patienten vorlagen, war die Berechnung des FLIPI für diese Patienten nicht möglich. Die m7-FLIPI-Gene *CARD11* und *FOXO1* wurden in der vorliegenden Arbeit nicht als rekurrent mutiert in FL identifiziert, wiesen jedoch auch in den Analysen von Pastore et al. die geringsten Mutationsfrequenzen (MF) auf²⁷¹. Demnach ist davon auszugehen, dass genetische Aberrationen in *CARD11* und *FOXO1* prognostisch relevant sind, jedoch deutlich seltener auftreten als Mutationen in den Genen *EZH2*, *ARID1A*, *MEF2B*, *EP300*, *CREBBP*.

Acht der zehn am häufigsten mutierten Gene aus der Arbeit von Pastore et al. wurden in dieser Arbeit als rekurrent mutiert in FL identifiziert (*KMT2D*, *CREBBP*, *BCL2*, *TNFRSF14*, *EZH2*, *MEF2B*, *ARID1A*, *EP300*). Mutationen in den Genen *IRF8* und *STAT6* konnten nicht nachgewiesen werden, da unter Verwendung des Lymphom-Panel die Target-Regionen nicht (*STAT6*) bzw. nur somatisch hypermutiert Regionen (*IRF8*) betrachtet wurden. Insgesamt stehen die erzielten Ergebnisse der FL-Mutationsanalysen der vorliegenden Arbeit im Einklang mit den Daten anderen Studien. Es ist bereits aus anderen wissenschaftlichen Arbeiten bekannt, dass in FL vorwiegend Aberrationen in Genen vorliegen, die zur Modifikation des Chromatins führen können (z.B. *KMT2D*, *CREBBP*, *EZH2*, *EP300*)^{128,129,132,272,273}.

Zusätzlich wurde in der vorliegenden Arbeit das *TMSB4X*-Gen als rekurrent mutiert in FL identifiziert. Vor Beginn dieser Arbeit waren genetische Aberrationen in *TMSB4X* bei FL nicht bekannt und ein Vergleich mit den Daten von Pastore et al. ist nicht möglich, da *TMSB4X* nicht in den analysierten Target-Regionen enthalten war. *TMSB4X* ist von zentraler Bedeutung des Zellwachstums, der -bewegung und der -differenzierung. Khodabakhshi et al. konnten bereits zeigen, dass u.a. *TMSB4X* somatisch hypermutiert in DLBCL vorliegt und diese Lymphome größtenteils vom GCB-Subtyp sind¹⁴⁰. Aus verschiedenen Veröffentlichungen sind weitere rekurrent mutierte Gene in FL bekannt u.a. *CDKN2A*, *CDKN2B*, *MYC*, *BCL6*, die z.T auch bei DLBCL nachweisbar sind^{50,57,59-62}.

Zudem konnte in der Vergangenheit gezeigt werden, dass etwa ein Drittel aller FL mit der Zeit zu aggressiven B-Zell-Lymphomen insbesondere DLBCL transformieren⁹². Dies steht im Einklang mit der Erkenntnis, dass die transformierten FL (tFL) den DLBCL vom GCB-Subtyp ähneln und viele genetische Aberrationen von tFL auch bei DLBCL-GCB nachweisbar sind²⁷⁴⁻²⁷⁷.

Die Mutationsdaten der vorliegenden Arbeit zeigen, dass Mutationen im *KMT2D*-Gen bei FL-Patienten mit dem Alter signifikant abnehmen (P-Wert <0,050). Analog dazu nehmen Aberrationen in *KMT2D* sowie Mutationen in anderen Genen die häufig in FL mutiert sind (z.B. *CREBBP*, *TNFRSF14*) bei DLBCL-Patienten im hohen Alter ab (siehe Abschnitt 3.2.1.2, Abbildung 36). Es gibt bisher keine Studie die explizit genetische Aberrationen bei senilen Patienten analysiert hat. In den publizierten Daten zu Mutationsanalysen bei Keimzentrums-B-Zell-Lymphomen wurden ausschließlich die rekurrent mutierten Gene unabhängig vom Alter betrachtet, sodass ein Vergleich zwischen den zwölf Genen die in dieser Arbeit vorwiegend in senilen Patient rekurrent mutiert waren (siehe Abschnitt 3.2.1.2, Tabelle 27) mit bereits publizierten Aberrationen in senilen DLBCL-Patienten nicht möglich ist. Die Mutationsdaten der vorliegenden Arbeit bestätigen die Hypothese, dass mit Zunahme des Ersterkrankungsalters der Patienten mit einem Keimzentrums-B-Zell-Lymphom die Anzahl an FL und DLBCL-GCB abnimmt und bei Patienten des Seniums häufiger DLBCL vom ABC-Subtyp vorliegen. Im folgenden werden die Ergebnisse der altersassoziierten genetischen Analysen dieser Arbeit im Vergleich zu bislang publizierten Ergebnissen zu molekulargenetischen Analysen der DLBCL-Subtypen genauer betrachtet und diskutiert.

4.2.2 Genetische Aberrationen in diffus großzelligen B-Zell-Lymphomen vom ABC-Subtyp

In den Mutationsanalysen der vorliegenden Arbeit wurden 50 rekurrent mutierte Gene in DLBCL-Patienten identifiziert, wobei u.a. die Gene *PCLO*, *PIM1*, *KMT2D*, *GNA13* und *CARD11* ermittelt wurden (Abschnitt 3.2.1.2., Abbildung 34). Ein Vergleich mit bereits publizierten Daten zeigt, dass das Mutationsprofil von DLBCL in der vorliegenden Arbeit mit den Ergebnissen anderer Arbeitsgruppen übereinstimmt^{62,127,135}. Hierbei ist jedoch zu beachten, dass die Mutationsanalysen in dieser Arbeit ausschließlich auf die Resequenzierung von spezifischen Target-Regionen beruht und diese Regionen gezielt für Mutationen in Keimzentrums-B-Zell-Lymphomen ausgewählt wurden. Demnach war es nicht möglich neue rekurrent mutierte Gene in DLBCL bzw. FL zu identifizieren.

In den Analysen dieser Arbeit konnten zwölf Gene identifiziert werden, die vorwiegend bei Patienten mit einem hohen Ersterkrankungsalter rekurrent mutiert vorliegen (Medianes Alter der mutierten Patienten über 80 Jahre, siehe Abschnitt 3.2.1.2, Tabelle 27). Zu diesen zwölf Genen zählen u.a. *TP53*, *PRDM1*, *IRF4* und *MYD88*. Zudem zeigten die erhobenen Daten der vorliegenden Arbeit, dass sehr alte DLBCL-Patienten im Vergleich zu adulten DLBCL-Patienten signifikant häufiger vom ABC-Subtyp sind, signifikant mehr Ki67 und signifikant weniger CD10 exprimieren.

*TP53*²⁷⁸⁻²⁸¹, *PRDM1*¹⁵⁵, *IRF4*⁵⁷ und *MYD88*¹⁶⁰ sind Gene für die bereits bekannt ist, dass Mutationen in diesen Genen verstärkt mit dem ABC-Subtyp assoziiert sind, was mit den Ergebnissen der Mutationsanalysen dieser Arbeit übereinstimmt (P-Wert <0,050).

Das Tumorsuppressorgen *TP53* besitzt eine zentrale Rolle in der Genese vieler Tumore²⁸²⁻²⁸⁵ und es wurde bereits gezeigt, dass genetische Aberrationen im *TP53*-Gen zu einer erhöhten Therapieresistenz führen und mit einem ungünstigen Krankheitsverlauf assoziiert sind²⁸⁶. In den Analysen der vorliegenden Arbeit wurden in 14 % der analysierten DLBCL-Kohorte Mutationen im *TP53*-Gen nachgewiesen, die vorwiegend bei älteren Patienten diagnostiziert wurden (Median: 81 Jahre). In den Mutationsanalysen dieser Arbeit liegen die Aberrationen größtenteils in der DNA-bindenden Domäne von *TP53*, was vermutlich zur Inaktivierung des Gens und der Inhibierung der nachfolgenden Signalkaskaden führt. Infolgedessen setzt die unkontrollierte Proliferation der Tumorzellen sowie die Bildung von Lymphomen ein.

PRDM1 (*BLIMP1*) ist ein transkriptioneller Repressor der von zentraler Bedeutung in der terminalen B-Zell-Differenzierung ist und von einigen Zentrozyten mit plasmazell-spezifischen Markern im Keimzentrum sowie von allen Plasmazellen exprimiert wird²⁸⁷⁻²⁸⁹. Es ist bereits bekannt, dass *PRDM1* ein Tumorsuppressorgen ist und genetische Aberrationen im *PRDM1*-Gen die Plasmazell-Differenzierung der B-Zellen unterbinden und zur Lymphogenese führen können. Pasqualucci et al. konnten in ihrer Arbeit zeigen, dass Mutationen in *PRDM1* vorwiegend bei DLBCL vom ABC-Subtyp auftreten und bei ~25% der DLBCL-ABC eine Deregulation von *BCL6* und eine Inaktivierung von *PRDM1* vorliegt, wodurch die terminale Differenzierung der Keimzentrums-B-Zellen blockiert wird¹⁵⁵. In der vorliegenden Arbeit wurde in 14% der analysierten DLBCL-Kohorte eine Mutation in *PRDM1* identifiziert. Die genaue Betrachtung des Mutationsprofils zeigte, dass die Aberrationen in keiner spezifischen Domäne sondern verteilt über die vollständige kodierende Region des *PRDM1*-Gens auftraten. Interessanterweise wurden in dieser Arbeit Mutationen im *PRDM1* vorwiegend bei älteren Patienten identifiziert (Median: 82 Jahre). Zudem waren DLBCL der Patienten mit einer *PRDM1*-Mutation größtenteils von ABC-Typ, was im Einklang mit der aktuell Literatur steht¹⁵⁵.

IRF4 (*MUM1*) ist ein Transkriptionsfaktor dessen Funktion u.a. in der Regulation des Immunglobulin-Klassenwechsels sowie der terminalen B-Zell-Differenzierung liegt. Analog zu *PRDM1* wird *IRF4* von einigen Keimzentrums-B-Zellen sowie von allen Plasmazellen exprimiert. Zudem konnten Sciammas et al. zeigen, dass *IRF4* den *PRDM1* Lokus aktivieren kann und somit eine entscheidende Bedeutung in der Regulation der Genexpression von *PRDM1* spielt²⁹⁰. In den Ergebnissen dieser Arbeit konnte in 12% aller DLBCL-Patienten eine Mutation im *IRF4*-Gen identifiziert werden, die vorwiegend in der bekannten somatisch hypermutierten Region der *IRF*-Domäne des Gens liegen.

Zudem wurden *IRF4*-Mutationen größtenteils bei Patienten mit einem hohen Ersterkrankungsalters nachgewiesen (Median: 83 Jahre). Die immunhistochemischen Färbungen veranschaulichten, dass die Expression von IRF4/MUM1 mit dem Alter zu und bei senilen Patienten wieder abnimmt. Es besteht daher die Hypothese, dass Mutationen im *IRF4*-Gen bei den älteren Patienten zur Inaktivierung der Genexpression führen.

MYD88 ist ein Adapterprotein, dass in etwa 30% aller DLBCL-ABC die Mutation L265P aufweist und infolgedessen über eine komplexe Signalkaskade die Aktivierung des NFκB-Signalweges erfolgt¹⁶⁰. Vor kurzem konnten Škunca et al.²⁹¹ zeigen, dass DLBCL-Patienten die NFκB im Cytosol exprimieren häufiger vom nonGCB-Subtyp sind und die Prognose der Patienten deutlich schlechter ist als von DLBCL-Patienten die NFκB nicht cytosolisch exprimieren, wobei diese häufiger dem GCB-Subtyp zugeordnet wurden²⁹¹. Die Mutationsanalysen dieser Arbeit bestätigten diese Erkenntnis, wobei 9% der analysierten DLBCL-Patienten eine L265P-Mutation aufwiesen und die *MYD88*-Mutation größtenteils bei älteren Patienten (Median: 82 Jahre) auftrat die vom ABC- bzw. nonGCB-Subtyp waren.

Aus verschiedenen Studien existieren bereits Ergebnisse die vermuten lassen, dass sich die DLBCL-Subtypen ABC und GCB nicht nur in den Ursprungszellen des Tumors sondern auch in den onkogenetischen Mechanismen und dem klinischen Verlauf der Erkrankung signifikant unterscheiden^{143,292-294}. In diesem Zusammenhang ist bekannt, dass DLBCL-GCB häufiger chromosomale Aberrationen in Genen aufweisen, die zur Inaktivierung der Apoptose führen können (z.B. *BCL2*¹³⁴, *BCL6*¹⁴⁷) oder genetische Aberrationen in Genen besitzen, die zur epigenetischen Modifikation von Histonen oder der Chromatinstruktur (z.B. *EZH2*¹³⁰, *CREBBP*¹³²) führen. Hingegen sind bei DLBCL-ABC häufiger genetische Aberrationen in Genen nachweisbar, die zur Aktivierung des NFκB-Signalweg führen (z.B. *MYD88*¹⁶⁰, *CD79B*¹⁶²), die Steuerung der B-Zell-Differenzierung (z.B. *PRDM1*^{154,155}) oder die Zellzyklusregulation (z.B. *CDKN2A*¹³⁴) beeinflussen.

Weiterhin zeigten die altersassoziierten molekulargenetischen Analysen dieser Arbeit, dass sehr alte DLBCL-Patienten im Vergleich zu adulten DLBCL-Patienten signifikant mehr Ki67 und signifikant weniger CD10 exprimieren (P-Wert <0,050).

CD10 ist ein proteolytisches Enzym, das u.a. auf der Zelloberfläche von Keimzentrums-B-Zellen exprimiert wird. Verschiedene in vitro und in vivo Studien zeigten, dass CD10 eine Rolle in der Regulation der B-Zell-Entwicklung spielt²⁹⁵. In diesem Zusammenhang konnte bereits gezeigt werden, dass die Expression von CD10 charakteristisch für Keimzentrums-B-Zell-Lymphome wie z.B. BL, FL und DLBCL vom GCB-Subtyp sind²⁹⁶. Hingegen exprimieren DLBCL-ABC deutlich weniger CD10. Zudem konnten verschiedene Arbeiten zeigen, dass die Expression von CD10 mit einer besseren Prognose für die Patienten assoziiert ist²⁹⁷⁻²⁹⁹.

Ki67 ist ein Protein, das als Proliferationsmarker im menschlichen Organismus fungiert und bei allen sich teilenden Zellen im Nukleus nachweisbar ist. Aufgrund der stark erhöhten Proliferationsrate exprimieren Tumorzellen deutlich mehr Ki67 als normale somatische Zellen³⁰⁰Perry, 2012 #1124. Inzwischen dienen Ki67-Antikörper standardmäßig zur Diagnostik oder Vorhersage des Therapieansprechens bei verschiedenen malignen Erkrankungen^{301,302}. Verschiedene Studien veranschaulichten, dass eine hohe Expression von Ki67 für einen schlechteren klinischen Verlauf bzw. eine ungünstige Prognose der Patienten spricht^{303,304}.

Aus der aktuellen Literatur ist bereits bekannt, dass DLBCL-ABC verstärkt den Proliferations-Marker Ki67 exprimieren^{143,294}, wohingegen die Expression des Keimzentrums-Markers CD10^{142,145} im Gegensatz zu DLBCL-GCB deutlich geringer ist. In den vergangenen Jahren konnten Wissenschaftler zeigen, dass DLBCL vom GCB-Subtyp mit dem Ersterkrankungsalter der Patienten abnimmt und im hohen Alter mehr DLBCL-ABC bzw. nonGCB vorliegen^{165,171,305 144,165,171,172}.

Insgesamt ist die Prognose für DLBCL-Patienten vom ABC- bzw. nonGCB-Subtyp deutlich schlechter als für Patienten vom GCB-Subtyp. So liegt die 5 Jahres-Überlebensrate für Patienten vom Subtyp GCB bei 60% und vom ABC bzw. nonGCB bei 30%^{143,292-294}. Es gibt Studien, die inzwischen zeigen konnten, dass die Patienten vom ABC-Subtyp deutlich schlechter bis gar nicht auf eine R-CHOP-Therapie ansprechen^{306,307}. In der Klinik werden bislang DLBCL-Patienten mit der Standardtherapie R-CHOP unabhängig von dem entsprechenden Subtyp ABC bzw. GCB behandelt³⁰⁸. Jedoch gibt es mittlerweile verschiedene neue Therapeutika, die in den bisherigen Phase-II-Studien erfolgreiche neue Therapieansätze versprechen. So gelten für die Behandlung von DLBCL insbesondere der Bruton-Tyrosinkinase-Inhibitor für den ABC-³⁰⁹ und BCL2-Inhibitoren (z.B. Venetoclax) für den GCB-Subtypen³¹⁰ als vielversprechende und wirksame Therapieoptionen.

Zusätzlich gibt inzwischen molekulargenetische Hochdurchsatzverfahren wie den Nanostring, die es ermöglichen eine schnelle Differenzierung von DLBCL in die Subtypen ABC und GCB hinsichtlich ihres Expressionsmusters vorzunehmen^{311,312}. Es ist davon auszugehen, dass derartige Methoden zeitnahe in die Routine-Diagnostik von Lymphomen übernommen werden und daraufhin eine gezielte Therapie der DLBCL-Patienten hinsichtlich ABC- und GCB-Subtyp vorgenommen werden kann.

Neben genetischen Aberrationen in Genen die den NFκB-Signalweg aktivieren und/oder in Assoziation zum ABC-Subtyp stehen, wurden in dieser Arbeit weitere Gene in DLBCL identifiziert, die vorwiegend in sehr alten Patienten mutiert vorliegen und bisher in keinem Zusammenhang mit dem Ersterkrankungsalter des Patienten standen. Beispielsweise wurde das *MTOR*-Gen als Kandidatengen für genetische Aberrationen in senilen DLBCL-Patienten ermittelt. MTOR ist ein Enzym, das in allen Säugetieren vorkommt und von zentraler Bedeutung für das Zellwachstum, der -bewegung und des -überlebens ist^{313,314}. Zudem kann das Polypeptid MTOR verschiedene Makromoleküle phosphorylieren, wodurch diese aktiviert werden³¹⁴⁻³¹⁶. Eine Blockade des Enzyms ist zwangsläufig mit der Inhibierung der Signalkaskade der Immunantwort verbunden und führt zur Hemmung des Immunsystems^{317,318}. Bei vielen hämatologischen sowie Tumor-Erkrankungen (z.B. Brust-, Prostata- und Lungenkrebs) konnte bereits gezeigt werden, dass die Patienten genetische Aberrationen im *MTOR*-Gen aufweisen^{319,320}. Vor Kurzem konnten Okosun et al. Mutationen im *RRAGC*-Gen von FL nachweisen, die folglich zur Aktivierung des MTOR-Signalweges bei den mutierten Patienten führten³²¹. In den Ergebnissen dieser Arbeit konnten in 6% der analysierten DLBCL-Patienten Aberrationen in *MTOR* identifiziert werden, wobei alle mutierten Patienten dem Senium zugehörig waren (Median: 81 Jahre). Die Betrachtung der identifizierten Mutationen unter Verwendung des Tools „intOGen“ ermittelte *MTOR* als *Driver-Gen*³²². Es besteht daher die Vermutung, dass es sich um aktivierende Mutationen im *MTOR*-Gen der senilen DLBCL Patienten handelt, die zur verstärkten Proliferation der Tumorzellen und Lymphogenese führen. In der medizinischen Krankenversorgung wird Rapamycin bereits als Therapeutikum bei verschiedenen Erkrankungen eingesetzt, wodurch der MTOR-Signalweg gehemmt wird. Derartige Immunsuppressiva werden u.a. bei neurodegenerativen Erkrankungen (z.B. Parkinson, Alzheimer)³²³, fortgeschrittenen Nierenzellkarzinomen³²⁴ und bei der Organtransplantation³²⁵ verabreicht. Möglicherweise stellen MTOR-Hemmer auch eine wirksame Therapieoption für senile DLBCL-Patienten dar und unterbinden die unkontrollierte Proliferation der Lymphomzellen.

Ein zusätzlicher Faktor der in den Analysen dieser Arbeit bisher nicht genauer betrachtet wurde (außer für *CD58*) sind strukturelle Aberrationen in den Genen. Es ist bereits bekannt, dass neben Punktmutationen auch strukturelle Aberrationen zur Entstehung und Progression von Lymphomen führen. Beispielsweise sind Deletionen in *CDKN2A* charakteristisch für DLBCL³²⁶⁻³²⁸. Aufgrund der angewandten Methode der Target-Resequenzierung ist eine Analyse von strukturellen Aberrationen in dieser Arbeit nur bedingt möglich. Der Coverage-Plot der analysierten Lymphom-Panel-Proben dieser Arbeit zeigte für einzelne Proben für das *CDKN2A*-Gen eine sehr geringe Coverage von unter 10x. Da im Rahmen des MMML-Verbundprojektes für viele Proben Array CGH- oder SNP-Array-Analysen durchgeführt wurden, lagen für einzelne Lymphom-Panel-Proben die Daten von strukturellen Aberrationen vor.

Ein Vergleich dieser low Coverage MMML-Lymphom-Panel-Proben mit den Array CGH- oder SNP-Array-Daten des MMML-Verbundprojektes bestätigte, dass bei diesen MMML-Lymphom-Panel-Proben eine Deletion im *CDKN2A*-Gen vorliegt. Aufgrund dieser Erkenntnis sollen die altersassoziierten genetischen Analysen dieser Arbeit auf strukturelle Veränderungen ausgeweitet werden. Dazu sollen zuerst die verfügbaren Array-CGH, SNP-Array und WGS-Daten der Forschungsstudien MMML und ICGC MMML-Seq in den Genen des Lymphom-Panels hinsichtlich struktureller Aberrationen analysiert werden. Anschließend erfolgt für potenzielle Kandidatengene eine Korrelationsanalyse mit dem Ersterkrankungsalter der Patienten. Sollten weitere potenzielle Kandidatengene für senile DLBCL-Patienten identifiziert werden, kann die neu zusammengestellte *Jack*-Kohorte zur Validierung der altersassoziierten strukturellen Aberrationen in DLBCL verwendet werden (z.B. mittels FISH-Analysen).

Insbesondere *CD58* wurde in den Analysen der vorliegenden Arbeit als Top Kandidatengenen für sehr alte DLBCL-Patienten ermittelt (siehe Abschnitt 3.2.1.2), wobei dessen Funktion und Bedeutung im menschlichen Organismus im nachfolgenden Abschnitt genauer betrachtet wird.

4.2.3 Mutationen und strukturelle Veränderungen im Kandidatengen *CD58*

Die in der vorliegenden Arbeit erhobenen Daten zeigen, dass genetische Aberrationen in *CD58* signifikant häufiger bei Patienten mit einem höheren Ersterkrankungsalter vorliegen. In den bislang veröffentlichten Manuskripten zu Mutationsanalysen in DLBCL wurden keine altersassoziierten genetischen Analysen vorgenommen, sodass in den Analysen dieser Arbeit das *CD58*-Gen erstmals als molekulargenetischer Marker für senile DLBCL-Patienten identifiziert wurde.

CD58 ist ein transmembranständiges Glykoprotein, das insbesondere von Antigen-repräsentierenden Zellen (APC) wie Makrophagen exprimiert wird und als Ligand für den CD2-Rezeptor fungiert, welcher von T-Zellen und vielen natürlichen Killerzellen (NK) repräsentiert wird^{329,330}. Das CD2-Glycoprotein ist wiederum für die Antigen-unabhängige und -abhängige Interaktionen von T-Zellen mit deren Zielmolekülen verantwortlich ist^{252,253,331-333}.

Die *CD58*-Mutationsdaten der vorliegenden Arbeit veranschaulichen, dass die genetischen Aberrationen größtenteils im Exon 2 und 3 des Gens liegen und somit die CD2-bindende Domäne des Adhäsionsproteins *CD58* betroffen ist. Diese Ergebnisse stehen im Einklang mit dem Mutationsprofil von *CD58* aus bereits publizierten Daten anderer Arbeitsgruppen³³⁴⁻³³⁶. Es ist bekannt, dass die CD2-Domäne mit dem CD2-Glycoprotein interagiert, welches wiederum für Interaktionen von T-Zellen mit deren Zielmolekülen verantwortlich ist^{252,253}.

Es besteht daher die Hypothese, dass genetische Aberrationen im *CD58*-Gen in Assoziation zur Immundefizienz beim menschlichen Organismus stehen und möglicherweise zur Entstehung von Lymphomen (insbesondere DLBCL) bei Patienten im höheren Alter führen. Insgesamt konnte somit die Fragestellung erfolgreich beantwortet werden, dass es genetische Aberrationen bzw. rekurrent mutierte Gene gibt, die vorwiegend bei Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem hohen Ersterkrankungsalter auftreten.

In der aktuellen Literatur ist bereits beschrieben, dass *CD58* rekurrent mutiert bei verschiedenen Lymphomen vorliegt. Beispielweise zeigten Schneider et al., dass genetische Aberrationen im *CD58*-Gen bei klassischen Hodgkin-Lymphomen zur Inaktivierung des Proteins führen, was möglicherweise zu einer veränderten Immunantwort der Tumorzellen führt³³⁴.

Die Mutationsdaten der vorliegenden Arbeit zeigen, dass neben *CD58* auch *B2M* rekurrent mutiert in DLBCL auftritt, jedoch sind *B2M*-Mutationen nicht altersassoziiert. Challa-Malladi et al.³³⁵ konnten in ihrer Arbeit zu DLBCL-Patienten zeigen, dass 21% der Patienten eine Aberration im *CD58*-Gen und 29% der Patienten eine Aberration im *B2M*-Gen aufwiesen. Zudem publizierten vor Kurzem französische Wissenschaftler die Ergebnisse der LYSA-Studie, worin das Mutationsprofil von 215 DLBCL-Patienten hinsichtlich 34 lymphom-spezifischer Gene mittels Target-Resequenzierung analysiert wurde³³⁶. In den Analysen der LYSA-Studie wurden die Gene *CD58*, *B2M*, *CIITA* und *TNFRSF14* als Gen-Gruppe des Immunsystems zusammengefasst. Die Akkumulation der identifizierten Aberrationen in den Immunsystem-Genen zeigte keine Assoziation mit dem Ersterkrankungsalter der Patienten. Eine separate Analyse der einzelnen Gene hinsichtlich der altersassoziierten genetischen Aberrationen und dem Patientenalter wurde in der LYSA-Studie nicht durchgeführt³³⁶.

4.2.4 Epigenetische Aberrationen in diffus großzelligen B-Zell-Lymphomen

Zu Beginn der vorliegenden Arbeit waren keine veröffentlichten Daten bekannt in denen die DNA-Methylierung von DLBCL gegenüber FL verglichen wurden oder worin altersassoziierte Analysen des DNA-Methylierungsmusters von Patienten mit einem Keimzentrums-B-Zell-Lymphom durchgeführt wurden. Es bestand daher die Frage, ob es Unterschiede in der DNA-Methylierung von jüngeren gegenüber älteren Patienten mit einem Keimzentrums-B-Zell-Lymphom gibt. Die Ergebnisse der epigenetischen Analysen in der vorliegenden Arbeit zeigen, dass sich BL und FL bzw. BL und DLBCL hinsichtlich ihres DNA-Methylierungsmusters signifikant unterscheiden. Während der Bearbeitung der vorliegenden Arbeit wurden die Daten von Kretzmer et al. publiziert in denen gezeigt werden konnte, dass sich die Keimzentrums-B-Zell-Lymphome in ihrem Methylierungsmuster entsprechend der Lymphom-Entitäten unterscheiden lassen³³⁷.

In der Arbeit von Kretzmer et al. konnte anhand der Korrelation von Methylom-, Genom- und Transkriptomdaten von BL und FL gezeigt werden, dass es signifikante Unterschiede in den differenziell methylierten Regionen zwischen FL und BL gibt und bei BL häufig aberrante Regulationen des SWI/SNF-, TCF-ID3-, und $G\alpha$ -Signalweges vorliegen³³⁷.

In den Methylierungsanalysen der vorliegenden Arbeit konnte das Phänomen der Heterogenität der DLBCL bestätigt werden und es wurden keine signifikant differenziell methylierten Regionen in DLBCL ermittelt, die ausschließlich bei senilen Patienten auftreten (siehe Abschnitt 3.2.2). Dies steht im Einklang mit den Ergebnissen von Kretzmer et al., worin DLBCL aufgrund der Heterogenität dieser Entität aus dem finalen Datensatz der Veröffentlichung ausgeschlossen wurden.

Nachdem in den Methylierungsanalysen keine Assoziation zwischen differenziell methylierten Regionen in DLBCL-Patienten gegenüber dem Patientenalter ermittelt werden konnten, wurde das Altersvorhersage-Modell nach Horvath für die Kontroll- und Tumorproben angewandt. Hierzu wurde das epigenetische Alter der Proben nach dem Horvath-Modell bestimmt, das die Altersvorhersage der Individuen basierend auf den Methylierungswerten von 353 CpG-Loci, die auf dem Infinium HumanMethylation450 BeadChip (Illumina) lokalisiert sind, ermöglicht²². Die Ergebnisse dieser Arbeit zeigen, dass das biologische und epigenetische Alter der analysierten Normalkontrollen miteinander korrelieren, wohingegen bei den DLBCL-Patienten das epigenetische Alter signifikant höher ist als das biologische Alter (siehe Abbildung 49). Die von mir beschriebene Assoziation von Mutationen in den rekurrent mutierten Genen (z.B. *CD58*) mit dem biologischen Alter der Patienten konnte auch in der Assoziation von Mutationen in den rekurrent mutierten Genen mit dem epigenetischen Alter nach Horvath bestätigt werden. Als Top Kandidatengen aus dem Vergleich der Gene mit dem höchsten biologischen und epigenetischen Alter der mutierten Patienten wurde *CD58* identifiziert (siehe Abbildung 51). Analog zu den erhobenen Daten der vorliegenden Arbeit konnte Steve Horvath bereits zeigen, dass das epigenetische Alter nach seinem Algorithmus signifikant mit dem biologischen Alter der Individuen korreliert und unabhängig vom Zelltyp nur die altersassoziierten Aberrationen in der DNA-Methylierung berücksichtigt²². Hingegen liegt bei Tumoren das epigenetische Alter häufig deutlich über dem biologischen Alter der Proben³³⁸, was vermuten lässt, dass die Tumorzellen bereits künstlich „vorgealtert“ sind.

Zusätzlich wurde in der vorliegenden Arbeit die Differenz des biologischen und epigenetischen Alters von jedem DLBCL-Patienten berechnet und im Zusammenhang mit den 50 rekurrent mutierten Genen betrachtet. Hierbei konnten Gene wie *CREBBP*, *KMT2D*, *TET1* oder *TET2* identifiziert werden, wobei Mutationen in diesen Genen möglicherweise einen Einfluss auf das epigenetische Alter der DLBCL-Patienten haben.

CREBBP ist eine Histonacetyltransferase und von entscheidender Bedeutung in der Veränderung der Chromatinstruktur innerhalb des menschlichen Organismus. Verschiedene Arbeitsgruppen konnten bereits zeigen, dass Mutationen in *CREBBP* bei FL vorliegen. Zudem konnten Mutationen im *CREBBP*-Gen von DLBCL-Patienten nachgewiesen werden, wobei die Lymphome größtenteils dem GCB-Subtyp zugeordnet wurden. Die Assoziationsanalysen der vorliegenden Arbeit stehen im Einklang mit den publizierten *CREBBP*-Daten und zeigten, dass das Gen rekurrent mutiert in FL vorliegt (Abbildung 29) und Mutationen mit dem Subtyp DLBCL-GCB korrelieren (Abbildung 42). Zudem zeigten die erhobenen Daten der vorliegenden Arbeit, dass das Delta-Alter der *CREBBP*-mutierten DLBCL-Patienten deutlich höher ist als bei den nicht *CREBBP*-mutierten DLBCL-Patienten (Abbildung 52). Möglicherweise führen Mutationen in *CREBBP* zur Veränderung der Chromatinstruktur, wodurch die maligne Transformation der B-Zellen bei DLBCL-Patienten begünstigt wird.

Ten-eleven translocation (TET) sind Enzyme die eine wichtige Rolle in der Modifikation von DNA insbesondere der DNA-Methylierung spielen, wobei es drei verschiedene TET-Enzyme gibt: TET1, -2 und -3. Aus publizierten Arbeiten ist bereits bekannt, dass TET die Oxidation von 5-Methylcytosin zu 5-Hydroxymethylcytosin der DNA katalysiert¹⁹³⁻¹⁹⁵. Verschiedene Arbeitsgruppen konnten bereits zeigen, dass insbesondere *TET2* bei hämatologischen Erkrankungen wie myelodysplastischem Syndrom (MDS), akuter myeloischer Leukämie (AML) und myeloproliferativen Neoplasien rekurrent mutiert ist^{200,201}. Zudem konnten inzwischen auch genetische Aberrationen in *TET2* bei Patienten mit chronisch myelomonozytären Leukämien und malignen Lymphomen nachgewiesen werden²⁰²⁻²⁰⁴. Die erhobenen Daten der vorliegenden Arbeit konnten zeigen, dass Mutationen in *TET1* und -2 vorwiegend bei DLBCL-Patienten mit hohem Ersterkrankungsalter auftreten und bei diesen *TET*-mutierten Patienten das Delta-Alter deutlich geringer ist als bei den nicht *TET*-mutierten DLBCL-Patienten (Abbildung 52). Möglicherweise führen Mutationen in den TET-Enzymen insbesondere bei älteren Patienten zu Modifikationen der DNA-Methylierung, wodurch es folglich zur Lymphogenese bei den sehr alten DLBCL-Patienten kommt.

Insgesamt konnten sechs Gene identifiziert werden, die potenziell das epigenetische Alter der betroffenen Patienten beeinflussen (sogenannte „epigenetic modifiers“) und die bereits mit epigenetischen Modifikationen in Tumoren assoziiert wurden: *CREBBP*, *TET1*, *TET2*, *KMT2D*, *HIST1H2AC*, *HIST1H1C*. In verschiedenen Arbeiten konnte in der Vergangenheit gezeigt werden, Mutationen in Chromatin- oder Histonmodifizierenden Genen insbesondere bei FL aufzutreten^{128,129,132,272,273}. Inzwischen gibt es verschiedenen Arbeiten die darauf hinweisen, dass Aberrationen in Chromatin- oder Histonmodifizierenden Genen auch bei DLBCL eine Rolle spielen^{62,127,132,339}.

Möglicherweise haben Mutationen in den „epigenetic modifiers“-Genen verschiedenen Funktionen und Auswirkungen in der Pathogenese von DLBCL und FL. Beispielsweise besteht aus den Arbeiten von Jiang et al. (2016) die Hypothese, dass *KMT2D*-Mutationen zum Einen ein frühes Event in der Lymphogenese darstellen und zum Anderen bei Rezidiven auftreten, wobei derartige Mutationen mit einer Schwächungen der Immunantwort verbunden sind³⁴⁰. Die erhobenen Daten der vorliegenden Arbeit lassen daher vermuten, dass Mutationen in Chromatin- oder Histonmodifizierenden Genen bei DLBCL zu Veränderungen im Epigenom und dem epigenetischen Alter der Patienten führen, wodurch die Fragestellung erfolgreich beantwortet wurde.

5 Ausblick

In den Mutationsdaten dieser Arbeit wurde *CD58* als Top Kandidatengen für genetische Aberrationen in senilen DLBCL-Patienten identifiziert. Es ist bekannt, dass auch in Hodgkin-Lymphomen *CD58* rekurrent mutiert vorliegt³³⁴ und aktuelle Studien zeigen, dass die Wirkung von *immune checkpoint inhibitors* u.a. PD-L1 oder PD-1-Antikörper wie Nivolumab bei Hodgkin-Lymphomen vielversprechende neue Therapeutika sind^{341,342}. Um eine erste Einschätzung der Aktivität von PD-L1 und PD-1 in DLBCL zu bekommen, sollen die Expressionen der Proteine PD-L1 und PD-1 in DLBCL-Patienten verschiedener Altersgruppen mittels Immunhistochemie analysiert werden. Möglicherweise können derartige *immune checkpoint inhibitors* auch bei DLBCL eine neue und wirksame Therapieoption für senile Patienten darstellen.

Zudem besteht die Möglichkeit DLBCL-Zelllinien mittels FACS-Analyse in *CD58*-positive und negative Fraktionen zu trennen und anschließend die gleichen Zelllinien mit einem spezifischen Immun-Panel mittels Nanostring zu analysieren. Hierzu ist bereits ein kommerzielles Immun-Panel (*Pancer Immune Profiling Panel* der Firma Nanostring) verfügbar, das die Expressionsanalyse von 770 Genen ermöglicht, die bereits in Assoziation mit der Immunantwort des menschlichen Organismus stehen. Anschließend könnte eine multiple Datenanalyse der strukturellen Aberrationen, Mutations-, FACS- und Expressionsdaten vorgenommen werden, um mögliche Interaktionen zwischen genetischen Aberrationen in *CD58* und einer Inhibierung des Immunsystems bei DLBCL zu identifizieren.

Für eine längerfristige und detaillierte Bearbeitung der Thematik wäre es sinnvoll das Mutationsprofil von einem Set von senilen DLBCL-Patienten mittels *Whole Exome Sequencing* oder *Whole Genome Sequencing* zu analysieren. In dieser Arbeit wurde die Mutationsanalyse ausschließlich an bereits publizierten lymphom-spezifischen Target-Regionen bzw. Genen vorgenommen. Da es in der Vergangenheit keine Studie gab, die gezielt senile DLBCL-Patienten analysiert haben, sind möglicherweise die Top Kandidatengene für senile DLBCL-Patienten noch nicht identifiziert worden und können demnach mittels Target-Resequenzierung nicht identifiziert werden. Zudem sind Analysen der Methylome (WGBS³⁴³), Transkriptome (RNA-Seq³⁴⁴), Histonstruktur (ChIP-Seq³⁴⁵) und Chromatin-Struktur (ATAC-Seq³⁴⁶) der senilen DLBCL-Patienten interessante Alternativen, um möglicherweise molekulargenetische Marker zu identifizieren, die längerfristig als potenzielle Therapieoption für senile DLBCL-Patienten dienen können.

6 Zusammenfassung

Diffus großzellige B-Zell-Lymphome (DLBCL) und folliculäre Lymphome (FL) sind die weitaus häufigsten Lymphome im Erwachsenenalter in Zentraleuropa und umfassen etwa 70-80% aller B-Zell-Lymphome bei Patienten im hohen Alter. Standardmäßig werden hochmaligne FL und DLBCL mit einer Kombination aus Chemotherapie und dem monoklonalen Antikörper Rituximab (R-CHOP) behandelt. Jedoch kann diese Therapie häufig bei sehr alten Patienten nicht angewandt werden, da das Immunsystem die Nebenwirkungen einer derartigen Hochdosis-Chemotherapie nicht übersteht. Für sehr alte Patienten mit einem Keimzentrums-B-Zell-Lymphom und einem hohen Internationalen prognostischen Index (IPI) gibt es derzeit keine wirksame Therapie. Aufgrund des kontinuierlichen Anstiegs der Lebenserwartung der Bevölkerung der westlichen Länder nimmt auch die Bedeutung der Behandlung von sehr alten Patienten mit einem Keimzentrums-B-Zell-Lymphom in der onkologischen Versorgung zu. Das Ziel dieser Arbeit war es daher altersassoziierte genetische und epigenetische Aberrationen bei Keimzentrums-B-Zell-Lymphomen zu identifizieren und dabei potenzielle molekulargenetische Marker zu ermitteln, die vorzugsweise bei sehr alten Patienten aberrant vorliegen und als mögliche Therapieoption für die sehr alten Patienten fungieren können. In dieser Arbeit wurde der Begriff Senium für Patienten ab einem Ersterkrankungsalter von 80 Jahren verwendet.

Zur Bearbeitung der Ziele war eine Methoden-Optimierung der Hochdurchsatz-Sequenzierung inkl. der Austestung verschiedener kommerzieller und Custom Panel zur Probenvorbereitung von Frisch- und FFPE-Material erforderlich. Insgesamt konnte im Rahmen dieser Arbeit die Probenvorbereitung von jeweils zwei kommerziellen und zwei Custom Panel optimiert werden. Insbesondere konnte ein Custom Panel (Lymphom-Panel) erstellt und optimiert werden, dass die Mutationsanalyse an Frisch- und FFPE-Material in Lymphomen von häufig mutierten Target-Regionen mittels Hochdurchsatz-Sequenzierung ermöglicht.

Für die Analyse der Keimzentrums-B-Zell-Lymphome wurde in dieser Arbeit eine neue Kohorte von insgesamt 198 sehr alten Patienten zusammengestellt und molekulargenetisch mittels Fluoreszenz-in-situ-Hybridisierung (FISH: *BCL2*, *BCL6*, *IGH*, *MYC*), Immunhistochemie (IHC: *BCL2*, *BCL6*, *CD10*, *Ki67*, *MUM1*), Target-Resequenzierung und Infinium HumanMethylation450 Bead ChiP charakterisiert. Zusätzlich standen die molekulargenetischen Daten von insgesamt 1694 Patienten aus den klinischen Studien RICOVER und NHL-B sowie den Forschungsstudien MMML und ICGC MMML-Seq zur Verfügung, sodass insgesamt 1892 Patienten im Alter von 3 bis 97 Jahren hinsichtlich genetischer und/oder epigenetischer Aberrationen in Keimzentrums-B-Zell-Lymphomen analysiert werden konnte.

Die Ergebnisse dieser Analysen konnten zeigen, dass bei Patienten mit einem Keimzentrums-B-Zell-Lymphom im hohen Alter vorwiegend DLBCL und nur selten FL diagnostiziert wurden und DLBCL bei älteren Patienten vorwiegend vom ABC- bzw. nonGCB-Subtypen waren. Darüber hinaus konnte gezeigt werden, dass die Tumore mit zunehmendem Alter der Patienten verstärkt BCL2, Ki67 und weniger CD10 exprimieren. Insgesamt wurden 50 rekurrent mutierte Gene in der analysierten DLBCL-Kohorte identifiziert, wovon 11 Gene eine Mutationsfrequenz von über 15% aufwiesen. Zudem überlappten die Gene stark mit bereits bekannten rekurrent mutierten Genen in DLBCL. Für 12 der rekurrent mutierten Gene wurde ein medianes Alter der Patienten von über 80 Jahren berechnet. Als Top Kandidatengen für DLBCL-Patienten im hohen Alter wurde *CD58* ermittelt, wobei in einer altersassoziierten Auswertung der Punktmutationen und strukturellen Aberrationen in *CD58* gezeigt wurde, dass das Gen signifikant häufiger bei DLBCL-Patienten mit einem hohen Ersterkrankungsalter verändert vorlag (Median: 74 Jahre, P-Wert: 0,0035).

In den epigenetischen Analysen dieser Arbeit konnte bestätigt werden, dass das DNA-Methylierungsprofil von DLBCL deutlich heterogener ist als das Methylierungsprofil von FL. Es konnten keine CpG-Loci in DLBCL-Patienten identifiziert werden, die signifikant häufiger im hohen Alter differentiell methylierte auftreten. Die Analysen der DNA-Methylierungswerte der Tumorproben mit dem Horvath-Modell zeigten, dass das berechnete epigenetische Alter der DLBCL-Patienten im Median deutlich über dem biologischen Alter lag. In einem Vergleich der Top 20 rekurrent mutierten Genen mit dem höchsten medianen Alter nach dem biologischen bzw. epigenetischen Alter konnten 12 Gene identifiziert werden, die in beiden Altersberechnungen rekurrent mutiert auftraten. Die Auswertung der 50 rekurrent mutierten Gene in DLBCL in Assoziation zur Differenz des biologischen und epigenetischen Alters (Delta-Alter) identifizierte 33 Gene als potenzielle „epigenetic modifier“. Diese Gene können möglicherweise das epigenetische Alter der analysierten Patienten beeinflussen und können damit ggf. als „epigenetic modifier“ bezeichnet werden. Insgesamt wurden 6 der 33 „epigenetic modifier“ als Gene identifiziert, die bereits im Zusammenhang mit epigenetischen Modifikationen bei Tumoren publiziert wurden: *CREBBP*, *KMT2D (MLL2)*, *HIST1H2AC*, *HIST1H1C*, *TET1* und *TET2*.

Zusammenfassend haben die in dieser Arbeit durchgeführten Analysen zur Identifizierung von genetischen und epigenetischen Aberrationen geführt, die vorwiegend in sehr alten Patienten auftraten. Diese Ergebnisse bestätigen die Hypothese, dass es Unterschiede in der Molekulargenetik von DLBCL zwischen jüngeren und älteren Patienten gibt. Insbesondere wurde *CD58* als Top Kandidatengen für genetische Aberrationen in sehr alten DLBCL-Patienten ermittelt. Hierbei handelt es sich um ein transmembranständiges Glykoprotein von dem bereits bekannt ist, dass es eine zentrale Bedeutung in der Interaktion von B- und T-Zellen und der damit verbundenen Immunantwort besitzt. Möglicherweise könnte das *CD58*-Gen in Zukunft einen neuen molekularen Marker für die Diagnostik und altersangepasste Therapie bei DLBCL-Patienten im Senium darstellen.

7 Summary

Diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) are the most common lymphoma in patients with advanced age in central Europe and account for around 70-80 % of all B-cell lymphoma at the advanced age. The standard therapy for FL and DLBCL is a combination of chemotherapy and monoclonal antibody Rituximab (R-CHOP). However, this therapy is often not applicable to very elderly patients because the immune system does not get over the side effects of that chemotherapy. For very elderly patients with a germinal center derived B-cell lymphoma and a high International Prognostic Index (IPI) score no effective standard therapy is currently available. Based on the continuous increase of the life expectancy in western countries the need of potential treatment in oncological care for very elderly patients with a germinal center derived B-cell lymphoma increases. Therefore, the aim of this work was to identify age-associated genetic and epigenetic alterations in germinal center derived B-cell lymphoma which are more prone to be mutated in very elderly patients. These markers might be help to achieve new therapy options for those patients. In this thesis the term “senium” was used for patients with an age at diagnosis of 80 and more years.

To identify genetic aberrations especially mutation targeted resequencing should be used. Therefore, it was necessary to optimize the high-throughput sequencing method using commercial and custom panels for the sample preparation of DNA extracted from fresh and FFPE material. In summary two commercial and two custom panels could be successfully optimized for the sample preparation followed by high-throughput sequencing within this thesis. In particular a custom panel (Lymphoma-Panel) could be established that enable the analysis of recurrently mutated target regions in lymphomas analysis in DNA extracted from fresh and FFPE material.

For the analysis of the germinal center derived B-cell lymphoma a new cohort of 198 very old patients was identified and characterized using fluorescence in-situ hybridization (FISH: *BCL2*, *BCL6*, *IGH*, *MYC*), immunohistochemistry (IHC: *BCL2*, *BCL6*, *CD10*, *Ki67*, *MUM1*), targeted resequencing and Infinium HumanMethylation450 Bead chip.

In addition molecular genetic data of 1694 germinal center derived B-cell lymphoma patients from the clinical studies RICOVER and NHL-B as well as the research studies MMML and ICGC MMML-Seq were available. In total of 1892 patients with an age at diagnosis of 3 to 97 years could be analyzed concerning age-associated genetic and/or epigenetic alterations in germinal center derived B-cell lymphoma.

The results of this analysis showed that very elderly patients with a germinal center derived B-cell lymphoma were mainly diagnosed with DLBCL and only seldom with FL were diagnosed. Furthermore, DLBCL of elderly patients were mainly of the ABC or nonGCB-subtype. In addition the results showed that with increasing age of the patients the frequency of tumors expressing BCL2 and Ki67 increases and decreases for CD10. Altogether, 50 recurrent mutated genes were identified in the analyzed DLBCL cohort with 11 genes showing a mutation frequency of more than 15% and a general high overlap to genes which are already known to be recurrently mutated in DLBCL. 12 recurrently mutated genes had a median age at diagnosis of 80 or more years. The top candidate gene for age-associated genetic alterations in very elderly DLBCL patients was *CD58*. An age-associated analysis of point mutations and structural aberrations showed that *CD58* occur significantly more often in DLBCL patients with high age at diagnosis (median: 74 years, p-value= 0.0035).

In this thesis the epigenetic analysis could confirm that the DNA methylation pattern of DLBCLs is compared to FLs more heterogeneous. No age-associated differential methylated CpG-Loci could be identified in DLBCL patients. Using the Horvath model the DNA methylation of the tumor samples showed that the epigenetic age of the DLBCL patients was in the median higher compared to the biological age. The top 20 recurrently mutated genes with the highest median biological age were compared to the top 20 recurrently mutated genes with the highest median epigenetic age and identified 12 genes which were mutated in both age calculations. The analysis of the 50 recurrently mutated genes in DLBCL in association to the difference of the biological and epigenetic age (delta age) identified 33 genes as potential "epigenetic modifier". Those genes might possibly influence the epigenetic age of the analyzed patients and can be called "epigenetic modifier". Altogether 6 of 33 "epigenetic modifier" were previously published in association to epigenetic modifications in tumors: *CREBBP*, *KMT2D (MLL2)*, *HIST1H2AC*, *HIST1H1C*, *TET1* and *TET2*.

In summary, the analysis performed in the framework of this thesis lead to the identification of the genetic and epigenetic aberrations, which appeared mainly in very elderly DLBCL patients. These results confirm the hypothesis that there are differences in the molecular genetics between younger (adult) and older (very elderly) DLBCL patients. In particular the *CD58* gene was identified as a top candidate gene for genetic aberrations in very elderly DLBCL patients. *CD58* is a transmembranous glycoprotein and it's already known that this protein has a central role in the interaction of B and T cells and the associated immune system function. Maybe, the *CD58* gene could be a new molecular marker for the diagnostics and age-adapted therapy for very elderly DLBCL patients.

8 Referenzen

1. Schultz, K.T. & Grieder, F. Structure and function of the immune system. *Toxicol Pathol* **15**, 262-4 (1987).
2. Neumann. Immunbiologie. *Springer-Verlag* (2008).
3. Mak, T.W. The Immune Response Basic and Clinical Principles. *Elsevier Academic Press* (2006).
4. Gordon, M.Y. & Gordon-Smith, E.C. Lymphocytes and haemopoiesis. *Br J Haematol* **47**, 163-9 (1981).
5. Gordon-Smith, T. Haemopoiesis - the formation of blood cells. *Medicine* (2013).
6. Frans P. Nijkamp, M.J.P. *Principles of Immunopharmacology* (2005).
7. Simm, A. *et al.* Potential biomarkers of ageing. *Biol Chem* **389**, 257-65 (2008).
8. Garagnani, P. *et al.* Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* **11**, 1132-4 (2012).
9. B. Kark, H.W. *Krebs im Alter - Zur Onkologie und Immunologie im höheren Lebensalter*, (1988).
10. Schuhmacher, B. *Das Geheimnis des menschlichen Alterns*, (2015).
11. Blair, S.N. *et al.* Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* **262**, 2395-401 (1989).
12. Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* **49**, 359-67 (2013).
13. Christensen, B.C. *et al.* Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet* **5**, e1000602 (2009).
14. Schraml, E. & Grillari, J. From cellular senescence to age-associated diseases: the miRNA connection. *Longev Healthspan* **1**, 10 (2012).
15. Fernandez, A.F. *et al.* A DNA methylation fingerprint of 1628 human samples. *Genome Res* **22**, 407-19 (2012).
16. Teschendorff, A.E. *et al.* Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* **20**, 440-6 (2010).
17. Rakyan, V.K. *et al.* Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* **20**, 434-9 (2010).
18. Javierre, B.M. *et al.* Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* **20**, 170-9 (2010).
19. Hernandez, D.G. *et al.* Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum Mol Genet* **20**, 1164-72 (2011).
20. Alisch, R.S. *et al.* Age-associated DNA methylation in pediatric populations. *Genome Res* **22**, 623-32 (2012).
21. Heyn, H. *et al.* Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci U S A* **109**, 10522-7 (2012).
22. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* **14**, R115 (2013).
23. Johansson, A., Enroth, S. & Gyllensten, U. Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. *PLoS One* **8**, e67378 (2013).
24. Weidner, C.I. *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol* **15**, R24 (2014).
25. Bocklandt, S. *et al.* Epigenetic predictor of age. *PLoS One* **6**, e14821 (2011).
26. Koch, C.M. & Wagner, W. Epigenetic-aging-signature to determine age in different tissues. *Aging (Albany NY)* **3**, 1018-27 (2011).
27. Yi, S.H., Jia, Y.S., Mei, K., Yang, R.Z. & Huang, D.X. Age-related DNA methylation changes for forensic age-prediction. *Int J Legal Med* **129**, 237-44 (2015).
28. Martin-Subero, J.I. *et al.* New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling. *Blood* **113**, 2488-97 (2009).

29. Martin-Subero, J.I. *et al.* A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One* **4**, e6986 (2009).
30. Ammerpohl, O. *et al.* Array-based DNA methylation analysis in classical Hodgkin lymphoma reveals new insights into the mechanisms underlying silencing of B cell-specific genes. *Leukemia* **26**, 185-8 (2012).
31. Herrmann, A. *et al.* Pipeline for large-scale microdroplet bisulfite PCR-based sequencing allows the tracking of hepitype evolution in tumors. *PLoS One* **6**, e21332 (2011).
32. Wiestner, A. The role of B-cell receptor inhibitors in the treatment of patients with chronic lymphocytic leukemia. *Haematologica* **100**, 1495-507 (2015).
33. Jorritsma, A., Schotte, R., Coccoris, M., de Witte, M.A. & Schumacher, T.N. Prospects and limitations of T cell receptor gene therapy. *Curr Gene Ther* **11**, 276-87 (2011).
34. Kershaw, M.H., Westwood, J.A. & Darcy, P.K. Gene-engineered T cells for cancer therapy. *Nat Rev Cancer* **13**, 525-41 (2013).
35. Uttenthal, B.J., Chua, I., Morris, E.C. & Stauss, H.J. Challenges in T cell receptor gene therapy. *J Gene Med* **14**, 386-99 (2012).
36. Choi, I., Lee, S. & Hong, Y.K. The new era of the lymphatic system: no longer secondary to the blood vascular system. *Cold Spring Harb Perspect Med* **2**, a006445 (2012).
37. Mak, T.W. & Saunders, M.E. *The immune response basic and clinical principles ; [CD-ROM included]*, XX, 1216 S. (Elsevier Acad. Press, Amsterdam u.a., 2006).
38. Melchers, F. *et al.* Positive and negative selection events during B lymphopoiesis. *Curr Opin Immunol* **7**, 214-27 (1995).
39. Weill, J.C. & Reynaud, C.A. Rearrangement/hypermutation/gene conversion: when, where and why? *Immunol Today* **17**, 92-7 (1996).
40. MacLennan, I.C. Germinal centers. *Annu Rev Immunol* **12**, 117-39 (1994).
41. Stein, H., Hummel, M., Johrens, K. & Anagnostopoulos, I. [Classification and pathogenesis of malignant lymphomas]. *Internist (Berl)* **48**, 351-61 (2007).
42. Basso, K. & Dalla-Favera, R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol* **15**, 172-84 (2015).
43. Conacci-Sorrell, M., McFerrin, L. & Eisenman, R.N. An overview of MYC and its interactome. *Cold Spring Harb Perspect Med* **4**, a014357 (2014).
44. Calado, D.P. *et al.* The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol* **13**, 1092-100 (2012).
45. Dominguez-Sola, D. *et al.* The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol* **13**, 1083-91 (2012).
46. Klein, U. *et al.* Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A* **100**, 2639-44 (2003).
47. Shaffer, A.L. *et al.* Signatures of the immune response. *Immunity* **15**, 375-85 (2001).
48. Dalla-Favera, R. *et al.* Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A* **79**, 7824-7 (1982).
49. Taub, R. *et al.* Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A* **79**, 7837-41 (1982).
50. Basso, K. & Dalla-Favera, R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev* **247**, 172-83 (2012).
51. Ci, W. *et al.* The BCL6 transcriptional program features repression of multiple oncogenes in primary B cells and is deregulated in DLBCL. *Blood* **113**, 5536-48 (2009).
52. Basso, K. *et al.* Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood* **115**, 975-84 (2010).
53. Hatzi, K. *et al.* A hybrid mechanism of action for BCL6 in B cells defined by formation of functionally distinct complexes at enhancers and promoters. *Cell Rep* **4**, 578-88 (2013).
54. Hatzi, K. & Melnick, A. Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis. *Trends Mol Med* **20**, 343-52 (2014).
55. Huang, C. *et al.* The BCL6 RD2 domain governs commitment of activated B cells to form germinal centers. *Cell Rep* **8**, 1497-508 (2014).

56. Huang, C., Hatzi, K. & Melnick, A. Lineage-specific functions of Bcl-6 in immunity and inflammation are mediated by distinct biochemical mechanisms. *Nat Immunol* **14**, 380-8 (2013).
57. Ochiai, K. *et al.* Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity* **38**, 918-29 (2013).
58. Pasqualucci, L. *et al.* Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood* **101**, 2914-23 (2003).
59. Kerfoot, S.M. *et al.* Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity* **34**, 947-60 (2011).
60. Lee, C.H. *et al.* Regulation of the germinal center gene program by interferon (IFN) regulatory factor 8/IFN consensus sequence-binding protein. *J Exp Med* **203**, 63-72 (2006).
61. Ying, C.Y. *et al.* MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat Immunol* **14**, 1084-92 (2013).
62. Morin, R.D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303 (2011).
63. Willis, S.N. *et al.* Transcription factor IRF4 regulates germinal center cell formation through a B cell-intrinsic mechanism. *J Immunol* **192**, 3200-6 (2014).
64. Falini, B. *et al.* A monoclonal antibody (MUM1p) detects expression of the MUM1/IRF4 protein in a subset of germinal center B cells, plasma cells, and activated T cells. *Blood* **95**, 2084-92 (2000).
65. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol* **1**, a000034 (2009).
66. Ghosh, G., Wang, V.Y., Huang, D.B. & Fusco, A. NF-kappaB regulation: lessons from structures. *Immunol Rev* **246**, 36-58 (2012).
67. Victora, G.D. *et al.* Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **143**, 592-605 (2010).
68. Basso, K. *et al.* Tracking CD40 signaling during germinal center development. *Blood* **104**, 4088-96 (2004).
69. Mullighan, C.G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-64 (2007).
70. Fazio, G., Palmi, C., Rolink, A., Biondi, A. & Cazzaniga, G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* **68**, 181-9 (2008).
71. Kawamata, N. *et al.* Molecular allelotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* **111**, 776-84 (2008).
72. Dimitrova, L. *et al.* PAX5 overexpression is not enough to reestablish the mature B-cell phenotype in classical Hodgkin lymphoma. *Leukemia* **28**, 213-6 (2014).
73. Minnich, M. *et al.* Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. *Nat Immunol* **17**, 331-43 (2016).
74. Tellier, J. *et al.* Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. *Nat Immunol* **17**, 323-30 (2016).
75. Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M. & Corcoran, L.M. The generation of antibody-secreting plasma cells. *Nat Rev Immunol* **15**, 160-71 (2015).
76. Hu, C.C., Dougan, S.K., McGehee, A.M., Love, J.C. & Ploegh, H.L. XBP-1 regulates signal transduction, transcription factors and bone marrow colonization in B cells. *EMBO J* **28**, 1624-36 (2009).
77. Reimold, A.M. *et al.* Plasma cell differentiation requires the transcription factor XBP-1. *Nature* **412**, 300-7 (2001).
78. Hahn, W.C. & Weinberg, R.A. Rules for making human tumor cells. *N Engl J Med* **347**, 1593-603 (2002).
79. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
80. Bayani, J. *et al.* Genomic mechanisms and measurement of structural and numerical instability in cancer cells. *Semin Cancer Biol* **17**, 5-18 (2007).

81. Balmain, A., Gray, J. & Ponder, B. The genetics and genomics of cancer. *Nat Genet* **33 Suppl**, 238-44 (2003).
82. Jones, P.A. & Baylin, S.B. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* **3**, 415-28 (2002).
83. Sadikovic, B., Al-Romaih, K., Squire, J.A. & Zielenska, M. Cause and consequences of genetic and epigenetic alterations in human cancer. *Curr Genomics* **9**, 394-408 (2008).
84. Kupperts, R., Klein, U., Hansmann, M.L. & Rajewsky, K. Cellular origin of human B-cell lymphomas. *N Engl J Med* **341**, 1520-9 (1999).
85. Kupperts, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* **5**, 251-62 (2005).
86. Stevenson, F.K. *et al.* The occurrence and significance of V gene mutations in B cell-derived human malignancy. *Adv Cancer Res* **83**, 81-116 (2001).
87. Greaves, M.F. Differentiation-linked leukemogenesis in lymphocytes. *Science* **234**, 697-704 (1986).
88. Shaffer, A.L., Rosenwald, A. & Staudt, L.M. Lymphoid malignancies: the dark side of B-cell differentiation. *Nat Rev Immunol* **2**, 920-32 (2002).
89. Swerdlow HC, E.H., N.L.; Jaffe, E.S.; Pileri, S.A.; Stein, H.; Thiele, J.; Vardiman, J.W. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. *International Agency for Research on Cancer 2008* (2008).
90. Vitolo, U., Ferreri, A.J. & Montoto, S. Follicular lymphomas. *Crit Rev Oncol Hematol* **66**, 248-61 (2008).
91. Buske, C. Indolente Lymphome: Pathophysiologie, Prognosefaktoren und aktuelle Therapieempfehlungen. *Deutscher Ärzte-Verlag* (2011).
92. Bende, R.J., Smit, L.A. & van Noesel, C.J. Molecular pathways in follicular lymphoma. *Leukemia* **21**, 18-29 (2007).
93. Freedman, A. Follicular lymphoma: 2011 update on diagnosis and management. *Am J Hematol* **86**, 768-75 (2011).
94. Fend, F. & Quintanilla-Martinez, L. Assessing the prognostic impact of immune cell infiltrates in follicular lymphoma. *Haematologica* **99**, 599-602 (2014).
95. Younes, S.F. *et al.* Immunoarchitectural patterns in follicular lymphoma: efficacy of HGAL and LMO2 in the detection of the interfollicular and diffuse components. *Am J Surg Pathol* **34**, 1266-76 (2010).
96. Bahler, D.W. & Levy, R. Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc Natl Acad Sci U S A* **89**, 6770-4 (1992).
97. Lenz, G. & Staudt, L.M. Aggressive lymphomas. *N Engl J Med* **362**, 1417-29 (2010).
98. Nogai, H., Dorken, B. & Lenz, G. Pathogenesis of non-Hodgkin's lymphoma. *J Clin Oncol* **29**, 1803-11 (2011).
99. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood* **89**, 3909-18 (1997).
100. Roman, E. & Smith, A.G. Epidemiology of lymphomas. *Histopathology* **58**, 4-14 (2011).
101. Smith, A., Howell, D., Patmore, R., Jack, A. & Roman, E. Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. *Br J Cancer* **105**, 1684-92 (2011).
102. Schneider, C., Pasqualucci, L. & Dalla-Favera, R. Molecular pathogenesis of diffuse large B-cell lymphoma. *Semin Diagn Pathol* **28**, 167-77 (2011).
103. Shaffer, A.L., 3rd, Young, R.M. & Staudt, L.M. Pathogenesis of human B cell lymphomas. *Annu Rev Immunol* **30**, 565-610 (2012).
104. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* **329**, 987-94 (1993).
105. Coiffier, B. *et al.* Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood* **116**, 2040-5 (2010).
106. Coiffier, B. *et al.* CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* **346**, 235-42 (2002).

107. Delarue, R. *et al.* Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. *Lancet Oncol* **14**, 525-33 (2013).
108. Cunningham, D. *et al.* Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet* **381**, 1817-26 (2013).
109. Pfreundschuh, M. *et al.* CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol* **7**, 379-91 (2006).
110. Philip, T. *et al.* Autologous bone marrow transplantation as compared with salvage chemotherapy in relapses of chemotherapy-sensitive non-Hodgkin's lymphoma. *N Engl J Med* **333**, 1540-5 (1995).
111. Peyrade, F. *et al.* Attenuated immunochemotherapy regimen (R-miniCHOP) in elderly patients older than 80 years with diffuse large B-cell lymphoma: a multicentre, single-arm, phase 2 trial. *Lancet Oncol* **12**, 460-8 (2011).
112. Pettengell, R. *et al.* Pixantrone dimaleate versus other chemotherapeutic agents as a single-agent salvage treatment in patients with relapsed or refractory aggressive non-Hodgkin lymphoma: a phase 3, multicentre, open-label, randomised trial. *Lancet Oncol* **13**, 696-706 (2012).
113. Fields, P.A. *et al.* De novo treatment of diffuse large B-cell lymphoma with rituximab, cyclophosphamide, vincristine, gemcitabine, and prednisolone in patients with cardiac comorbidity: a United Kingdom National Cancer Research Institute trial. *J Clin Oncol* **32**, 282-7 (2014).
114. Mounier, N. *et al.* Rituximab plus gemcitabine and oxaliplatin in patients with refractory/relapsed diffuse large B-cell lymphoma who are not candidates for high-dose therapy. A phase II Lymphoma Study Association trial. *Haematologica* **98**, 1726-31 (2013).
115. <http://www.dshnhl.org/studien/aktive-studien/b-r-enda-studie/>.
116. Pfreundschuh, M. *et al.* Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of young patients with good-prognosis (normal LDH) aggressive lymphomas: results of the NHL-B1 trial of the DSHNHL. *Blood* **104**, 626-33 (2004).
117. Pfreundschuh, M. *et al.* Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of elderly patients with aggressive lymphomas: results of the NHL-B2 trial of the DSHNHL. *Blood* **104**, 634-41 (2004).
118. Pfreundschuh, M. *et al.* Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). *Lancet Oncol* **9**, 105-16 (2008).
119. <http://www.lymphome.de/Projekte/MMML/>.
120. <http://icgc.org/>.
121. Tanaka, S., Louie, D.C., Kant, J.A. & Reed, J.C. Frequent incidence of somatic mutations in translocated BCL2 oncogenes of non-Hodgkin's lymphomas. *Blood* **79**, 229-37 (1992).
122. Horsman, D.E., Gascoyne, R.D., Coupland, R.W., Coldman, A.J. & Adomat, S.A. Comparison of cytogenetic analysis, southern analysis, and polymerase chain reaction for the detection of t(14; 18) in follicular lymphoma. *Am J Clin Pathol* **103**, 472-8 (1995).
123. Hockenbery, D., Nunez, G., Milliman, C., Schreiber, R.D. & Korsmeyer, S.J. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* **348**, 334-6 (1990).
124. Kridel, R., Sehn, L.H. & Gascoyne, R.D. Pathogenesis of follicular lymphoma. *J Clin Invest* **122**, 3424-31 (2012).
125. Huang, J.Z. *et al.* The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. *Blood* **99**, 2285-90 (2002).
126. Yang, J. *et al.* Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked. *Science* **275**, 1129-32 (1997).

127. Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* **43**, 830-7 (2011).
128. Okosun, J. *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet* **46**, 176-81 (2014).
129. Pasqualucci, L. *et al.* Genetics of follicular lymphoma transformation. *Cell Rep* **6**, 130-40 (2014).
130. Morin, R.D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* **42**, 181-5 (2010).
131. Bodor, C. *et al.* EZH2 mutations are frequent and represent an early event in follicular lymphoma. *Blood* **122**, 3165-8 (2013).
132. Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189-95 (2011).
133. Weiss, L.M., Warnke, R.A., Sklar, J. & Cleary, M.L. Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas. *N Engl J Med* **317**, 1185-9 (1987).
134. Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A* **105**, 13520-5 (2008).
135. Lohr, J.G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A* **109**, 3879-84 (2012).
136. Morin, R.D. *et al.* Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* **122**, 1256-65 (2013).
137. Zhang, J. *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* **110**, 1398-403 (2013).
138. Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-21 (2009).
139. de Miranda, N.F. *et al.* Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. *Blood* **124**, 2544-53 (2014).
140. Khodabakhshi, A.H. *et al.* Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget* **3**, 1308-19 (2012).
141. Hans, C.P. *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275-82 (2004).
142. Alizadeh, A.A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-11 (2000).
143. Rosenwald, A. *et al.* The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* **346**, 1937-47 (2002).
144. Lenz, G. *et al.* Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* **359**, 2313-23 (2008).
145. Wright, G. *et al.* A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **100**, 9991-6 (2003).
146. Iqbal, J. *et al.* BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am J Pathol* **165**, 159-66 (2004).
147. Iqbal, J. *et al.* Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia* **21**, 2332-43 (2007).
148. Bereshchenko, O.R., Gu, W. & Dalla-Favera, R. Acetylation inactivates the transcriptional repressor BCL6. *Nat Genet* **32**, 606-13 (2002).
149. Saito, M. *et al.* BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **106**, 11294-9 (2009).
150. Bakhshi, A. *et al.* Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. *Cell* **41**, 899-906 (1985).
151. Aukema, S.M. *et al.* Double-hit B-cell lymphomas. *Blood* **117**, 2319-31 (2011).
152. Li, S. *et al.* MYC/BCL2 double-hit high-grade B-cell lymphoma. *Adv Anat Pathol* **20**, 315-26 (2013).

153. Lenz, G. *et al.* Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J Exp Med* **204**, 633-43 (2007).
154. Tam, W. *et al.* Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. *Blood* **107**, 4090-100 (2006).
155. Pasqualucci, L. *et al.* Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J Exp Med* **203**, 311-7 (2006).
156. Saito, M. *et al.* A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell* **12**, 280-92 (2007).
157. Schmidlin, H. *et al.* Spi-B inhibits human plasma cell differentiation by repressing BLIMP1 and XBP-1 expression. *Blood* **112**, 1804-12 (2008).
158. Zainuddin, N. *et al.* TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. *Leuk Res* **33**, 60-6 (2009).
159. Young, K.H. *et al.* Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood* **112**, 3088-98 (2008).
160. Ngo, V.N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115-9 (2011).
161. Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676-9 (2008).
162. Davis, R.E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88-92 (2010).
163. Oschlies, I. *et al.* Pediatric follicular lymphoma--a clinico-pathological study of a population-based series of patients treated within the Non-Hodgkin's Lymphoma--Berlin-Frankfurt-Munster (NHL-BFM) multicenter trials. *Haematologica* **95**, 253-9 (2010).
164. Salaverria, I. *et al.* Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. *Blood* **118**, 139-47 (2011).
165. Klapper, W. *et al.* Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. *Blood* **119**, 1882-7 (2012).
166. <http://www.dshnhl.org/therapie/aggressives-b-zell-lymphom/>.
167. <http://www.who.int/mediacentre/factsheets/fs404/en/>.
168. Solal-Celigny, P. *et al.* Follicular lymphoma international prognostic index. *Blood* **104**, 1258-65 (2004).
169. Determann, O. *et al.* Ki-67 predicts outcome in advanced-stage mantle cell lymphoma patients treated with anti-CD20 immunochemotherapy: results from randomized trials of the European MCL Network and the German Low Grade Lymphoma Study Group. *Blood* **111**, 2385-7 (2008).
170. Morton, L.M. *et al.* Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood* **107**, 265-76 (2006).
171. Oschlies, I. *et al.* Diffuse large B-cell lymphoma in pediatric patients belongs predominantly to the germinal-center type B-cell lymphomas: a clinicopathologic analysis of cases included in the German BFM (Berlin-Frankfurt-Munster) Multicenter Trial. *Blood* **107**, 4047-52 (2006).
172. Klapper, W. *et al.* Molecular profiling of pediatric mature B-cell lymphoma treated in population-based prospective clinical trials. *Blood* **112**, 1374-81 (2008).
173. Oyama, T. *et al.* Senile EBV+ B-cell lymphoproliferative disorders: a clinicopathologic study of 22 patients. *Am J Surg Pathol* **27**, 16-26 (2003).
174. Hofscheier, A. *et al.* Geographic variation in the prevalence of Epstein-Barr virus-positive diffuse large B-cell lymphoma of the elderly: a comparative analysis of a Mexican and a German population. *Mod Pathol* **24**, 1046-54 (2011).
175. Hoeller, S., Tzankov, A., Pileri, S.A., Went, P. & Dirnhofer, S. Epstein-Barr virus-positive diffuse large B-cell lymphoma in elderly patients is rare in Western populations. *Hum Pathol* **41**, 352-7 (2010).

176. Stuhlmann-Laeisz, C., Szczepanowski, M., Borchert, A., Bruggemann, M. & Klapper, W. Epstein-Barr virus-negative diffuse large B-cell lymphoma hosts intra- and peritumoral B-cells with activated Epstein-Barr virus. *Virchows Arch* **466**, 85-92 (2015).
177. Jacobs, K.B. *et al.* Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet* **44**, 651-8 (2012).
178. http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf.
179. Mamanova, L. *et al.* Target-enrichment strategies for next-generation sequencing. *Nat Methods* **7**, 111-8 (2010).
180. <http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet-trusight-myeloid.pdf>.
181. Fischle, W. *et al.* Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* **17**, 1870-81 (2003).
182. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* **8**, 286-98 (2007).
183. Hergersberg, M. Biological aspects of cytosine methylation in eukaryotic cells. *Experientia* **47**, 1171-85 (1991).
184. Brena, R.M., Huang, T.H. & Plass, C. Toward a human epigenome. *Nat Genet* **38**, 1359-60 (2006).
185. Pinney, S.E. Mammalian Non-CpG Methylation: Stem Cells and Beyond. *Biology (Basel)* **3**, 739-51 (2014).
186. Suzuki, M.M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**, 465-76 (2008).
187. Lopez, J., Percharde, M., Coley, H.M., Webb, A. & Crook, T. The context and potential of epigenetics in oncology. *Br J Cancer* **100**, 571-7 (2009).
188. Kanwal, R. & Gupta, S. Epigenetic modifications in cancer. *Clin Genet* **81**, 303-11 (2012).
189. Choi, J.H. *et al.* Genome-wide DNA methylation maps in follicular lymphoma cells determined by methylation-enriched bisulfite sequencing. *PLoS One* **5**(2010).
190. Poetsch, A.R. & Plass, C. Transcriptional regulation by DNA methylation. *Cancer Treat Rev* **37 Suppl 1**, S8-12 (2011).
191. Tsai, H.C. & Baylin, S.B. Cancer epigenetics: linking basic biology to clinical medicine. *Cell Res* **21**, 502-17 (2011).
192. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-93 (2001).
193. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-5 (2009).
194. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929-30 (2009).
195. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-33 (2010).
196. Kohli, R.M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472-9 (2013).
197. Globisch, D. *et al.* Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* **5**, e15367 (2010).
198. Munzel, M. *et al.* Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl* **49**, 5375-7 (2010).
199. Wen, L. *et al.* Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol* **15**, R49 (2014).
200. Langemeijer, S.M. *et al.* Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet* **41**, 838-42 (2009).
201. Delhommeau, F. *et al.* Mutation in TET2 in myeloid cancers. *N Engl J Med* **360**, 2289-301 (2009).
202. Bacher, U. *et al.* Mutations of the TET2 and CBL genes: novel molecular markers in myeloid malignancies. *Ann Hematol* **89**, 643-52 (2010).

203. Qivovoron, C. *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25-38 (2011).
204. Abdel-Wahab, O. *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114**, 144-7 (2009).
205. Nakajima, H. & Kunimoto, H. TET2 as an epigenetic master regulator for normal and malignant hematopoiesis. *Cancer Sci* **105**, 1093-9 (2014).
206. Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288-95 (2011).
207. Sandoval, J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **6**, 692-702 (2011).
208. Carew, J.A. & Feldberg, R.S. Recognition of a cytosine base lesion by a human damage-specific DNA binding protein. *Nucleic Acids Res* **13**, 303-15 (1985).
209. Bibikova, M. *et al.* High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* **16**, 383-93 (2006).
210. Clark, S.J., Harrison, J., Paul, C.L. & Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* **22**, 2990-7 (1994).
211. Bertrand, S. *et al.* Variant translocation in a non endemic case of Burkitt's lymphoma: t (8;22) in an Epstein-Barr virus negative tumour and in a derived cell line. *Eur J Cancer* **17**, 577-84 (1981).
212. Lenoir GM, V.M., Bonnardel C. The use of lymphomatous and lymphoblastoid cell lines in the study of Burkitt's lymphoma. *IARC Sci Publ.*, 309-318. (1985).
213. Burmeister, T. *et al.* A novel sporadic Burkitt lymphoma cell line (BLUE-1) with a unique t(6;20)(q15;q11.2) rearrangement. *Leuk Res* **30**, 1417-23 (2006).
214. Magrath, I.T. *et al.* Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative for Epstein-Barr virus nuclear antigen. I. Physical, cytogenetic, and growth characteristics. *J Natl Cancer Inst* **64**, 465-76 (1980).
215. Nadkarni, J.S. *et al.* Characteristics of new cell lines derived from Burkitt lymphomas. *Cancer* **23**, 64-79 (1969).
216. Ben-Bassat, H. *et al.* Establishment in continuous culture of a new type of lymphocyte from a "Burkitt like" malignant lymphoma (line D.G.-75). *Int J Cancer* **19**, 27-33 (1977).
217. Epstein MA, B.Y. CULTIVATION IN VITRO OF HUMAN LYMPHOBLASTS FROM BURKITT'S MALIGNANT LYMPHOMA. *Lancet*, 1: 252-253 (1964).
218. Pulvertaft, J.V. Cytology of Burkitt's Tumour (African Lymphoma). *Lancet* **1**, 238-40 (1964).
219. Epstein, A.L. *et al.* Biology of the human malignant lymphomas. IV. Functional characterization of ten diffuse histiocytic lymphoma cell lines. *Cancer* **42**, 2379-91 (1978).
220. Nilsson, K. & Sundstrom, C. Establishment and characteristics of two unique cell lines from patients with lymphosarcoma. *Int J Cancer* **13**, 808-23 (1974).
221. Beckwith, M., Longo, D.L., O'Connell, C.D., Moratz, C.M. & Urba, W.J. Phorbol ester-induced, cell-cycle-specific, growth inhibition of human B-lymphoma cell lines. *J Natl Cancer Inst* **82**, 501-9 (1990).
222. Dyer, M.J., Fischer, P., Nacheva, E., Labastide, W. & Karpas, A. A new human B-cell non-Hodgkin's lymphoma cell line (Karpas 422) exhibiting both t (14;18) and t(4;11) chromosomal translocations. *Blood* **75**, 709-14 (1990).
223. Schuetz, J.M. *et al.* BCL2 mutations in diffuse large B-cell lymphoma. *Leukemia* **26**, 1383-90 (2012).
224. Maria Murga Penas, E. *et al.* Comprehensive cytogenetic and molecular cytogenetic analysis of 44 Burkitt lymphoma cell lines: secondary chromosomal changes characterization, karyotypic evolution, and comparison with primary samples. *Genes Chromosomes Cancer* **53**, 497-515 (2014).
225. Hummel, M. *et al.* A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* **354**, 2419-30 (2006).
226. Horn, H. *et al.* MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. *Blood* **121**, 2253-63 (2013).

227. Klapper, W. *et al.* Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Leukemia* **22**, 2226-9 (2008).
228. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology* **24**, 104-8 (1992).
229. https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_cancer.pdf.
230. Richter, J. *et al.* Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat Genet* **44**, 1316-20 (2012).
231. Love, C. *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet* **44**, 1321-5 (2012).
232. Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490**, 116-20 (2012).
233. Jones, D.T. *et al.* Dissecting the genomic complexity underlying medulloblastoma. *Nature* **488**, 100-5 (2012).
234. Jones, D.T. *et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet* **45**, 927-32 (2013).
235. Rimmer, A. *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* **46**, 912-8 (2014).
236. Walker, D.L. *et al.* DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip. *Epigenomics* **7**, 1287-302 (2015).
237. Assenov, Y. *et al.* Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods* **11**, 1138-40 (2014).
238. Triche, T.J., Jr., Weisenberger, D.J., Van Den Berg, D., Laird, P.W. & Siegmund, K.D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res* **41**, e90 (2013).
239. Maksimovic, J., Gordon, L. & Oshlack, A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol* **13**, R44 (2012).
240. Dedeurwaerder, S. *et al.* Evaluation of the Infinium Methylation 450K technology. *Epigenomics* **3**, 771-84 (2011).
241. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **125**, 279-84 (2001).
242. Long, J.S. Regression Models for Categorical and Limited Dependent Variables *Sage Pubn*, 297 (1997).
243. Long, J.S. Regression models for categorical dependent variables using Stata. *Stata Press*, 527 (2014).
244. Hosmer, D.W. Applied Logistic Regression. *WILEY* (2013).
245. Bronstein, I.N. Taschenbuch der Mathematik. *Harri Deutsch* (2000).
246. Fagerland, M.W. & Sandvik, L. The Wilcoxon-Mann-Whitney test under scrutiny. *Stat Med* **28**, 1487-97 (2009).
247. Leonard A. Marascuilo, M.M. Nonparametric and Distribution-free Methods for the Social Sciences. *Social Sciences. Brooks/Cole Publishing Co* (1977).
248. Jürgen Bortz, G.A.L., Klaus Boehnke. Verteilungsfreie Methoden in der Biostatistik. 3. Auflage. . *Springer Verlag*, 259, 729 (2010).
249. McDonald, J.H. Handbook of Biological Statistics (3rd ed.). 157-164 (2014).
250. Kruskal, W.H. International Encyclopedia of Statistics. *Free Pr* (1979).
251. Oron, A.P. Kruskal-Wallis and Friedman type tests for nested effects in hierarchical designs. *Working Paper no. 68, Center for Statistics and the Social Sciences, University of Washington* (2006).
252. Sewell, W.A., Brown, M.H., Dunne, J., Owen, M.J. & Crumpton, M.J. Molecular cloning of the human T-lymphocyte surface CD2 (T11) antigen. *Proc Natl Acad Sci U S A* **83**, 8718-22 (1986).

253. Ikemizu, S. *et al.* Crystal structure of the CD2-binding domain of CD58 (lymphocyte function-associated antigen 3) at 1.8-Å resolution. *Proc Natl Acad Sci U S A* **96**, 4289-94 (1999).
254. Haber, D. & Harlow, E. Tumour-suppressor genes: evolving definitions in the genomic age. *Nat Genet* **16**, 320-2 (1997).
255. Weinberg, R.A. Tumor suppressor genes. *Science* **254**, 1138-46 (1991).
256. Do, H. & Dobrovic, A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin Chem* **61**, 64-71 (2015).
257. Spencer, D.H. *et al.* Comparison of clinical targeted next-generation sequence data from formalin-fixed and fresh-frozen tissue specimens. *J Mol Diagn* **15**, 623-33 (2013).
258. Chen, Y.C., Liu, T., Yu, C.H., Chiang, T.Y. & Hwang, C.C. Effects of GC bias in next-generation-sequencing data on de novo genome assembly. *PLoS One* **8**, e62856 (2013).
259. Yang, X., Chockalingam, S.P. & Aluru, S. A survey of error-correction methods for next-generation sequencing. *Brief Bioinform* **14**, 56-66 (2013).
260. Treangen, T.J. & Salzberg, S.L. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* **13**, 36-46 (2012).
261. <https://support.illumina.com/content/dam/illumina-marketing/documents/products/other/miseq-overclustering-primer-770-2014-038.pdf>.
262. Covaris. http://covarisinc.com/wp-content/uploads/pn_010200.pdf.
263. Diagenode. <https://www.diagenode.com/applications/ffpe-dna-extraction>.
264. Qiagen. <https://www.qiagen.com/de/shop/sample-technologies/dna/dna-preparation/generead-dna-ffpe-kit - orderinginformation>.
265. Al-Hamadani, M. *et al.* Non-Hodgkin lymphoma subtype distribution, geodemographic patterns, and survival in the US: A longitudinal analysis of the National Cancer Data Base from 1998 to 2011. *Am J Hematol* **90**, 790-5 (2015).
266. Institut, R.K. Gesundheit in Deutschland 2015: Kapitel 02. Wie steht es um unsere Gesundheit? (2015).
267. Leich, E., Ott, G. & Rosenwald, A. Pathology, pathogenesis and molecular genetics of follicular NHL. *Best Pract Res Clin Haematol* **24**, 95-109 (2011).
268. Williams, J.N. *et al.* Disease characteristics, patterns of care, and survival in very elderly patients with diffuse large B-cell lymphoma. *Cancer* **121**, 1800-8 (2015).
269. Chang, C.M. *et al.* Risk factors for non-Hodgkin lymphoma subtypes defined by histology and t(14;18) in a population-based case-control study. *Int J Cancer* **129**, 938-47 (2011).
270. Biagi, J.J. & Seymour, J.F. Insights into the molecular pathogenesis of follicular lymphoma arising from analysis of geographic variation. *Blood* **99**, 4265-75 (2002).
271. Pastore, A. *et al.* Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *Lancet Oncol* **16**, 1111-22 (2015).
272. Green, M.R. *et al.* Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* **121**, 1604-11 (2013).
273. Li, H. *et al.* Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood* **123**, 1487-98 (2014).
274. Davies, A.J. *et al.* Transformation of follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms. *Br J Haematol* **136**, 286-93 (2007).
275. Lossos, I.S. & Gascoyne, R.D. Transformation of follicular lymphoma. *Best Pract Res Clin Haematol* **24**, 147-63 (2011).
276. Vaidyanathan, G., Ngamphaiboon, N. & Hernandez-Ilizaliturri, F.J. Clinical spectrum and prognosis of follicular lymphoma with blastoid transformation: case series and a review of the literature. *Ann Hematol* **90**, 955-62 (2011).
277. Eide, M.B. *et al.* Genomic alterations reveal potential for higher grade transformation in follicular lymphoma and confirm parallel evolution of tumor cell clones. *Blood* **116**, 1489-97 (2010).

278. Wilson, W.H. *et al.* Relationship of p53, bcl-2, and tumor proliferation to clinical drug resistance in non-Hodgkin's lymphomas. *Blood* **89**, 601-9 (1997).
279. Gascoyne, R.D. Pathologic prognostic factors in diffuse aggressive non-Hodgkin's lymphoma. *Hematol Oncol Clin North Am* **11**, 847-62 (1997).
280. Ichikawa, A. *et al.* Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N Engl J Med* **337**, 529-34 (1997).
281. Leroy, K. *et al.* p53 gene mutations are associated with poor survival in low and low-intermediate risk diffuse large B-cell lymphomas. *Ann Oncol* **13**, 1108-15 (2002).
282. Starostik, P. *et al.* Genetic aberrations common in gastric high-grade large B-cell lymphoma. *Blood* **95**, 1180-7 (2000).
283. Lo Coco, F. *et al.* p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood* **82**, 2289-95 (1993).
284. Sander, C.A. *et al.* p53 mutation is associated with progression in follicular lymphomas. *Blood* **82**, 1994-2004 (1993).
285. Martinez-Climent, J.A. *et al.* Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood* **101**, 3109-17 (2003).
286. Xu-Monette, Z.Y. *et al.* Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood* **120**, 3986-96 (2012).
287. Turner, C.A., Jr., Mack, D.H. & Davis, M.M. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* **77**, 297-306 (1994).
288. Angelin-Duclos, C., Cattoretti, G., Lin, K.I. & Calame, K. Commitment of B lymphocytes to a plasma cell fate is associated with Blimp-1 expression in vivo. *J Immunol* **165**, 5462-71 (2000).
289. Cattoretti, G. *et al.* PRDM1/Blimp-1 is expressed in human B-lymphocytes committed to the plasma cell lineage. *J Pathol* **206**, 76-86 (2005).
290. Sciammas, R. *et al.* Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* **25**, 225-36 (2006).
291. Skunca, Z. & Planinc-Peraica, A. [PROGNOSTIC ROLE OF NF-kappaB EXPRESSION IN DIFFUSE LARGE B-CELL LYMPHOMA SUBGROUPS]. *Acta Med Croatica* **69**, 25-32 (2015).
292. Savage, K.J. *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* **102**, 3871-9 (2003).
293. Rosenwald, A. *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* **198**, 851-62 (2003).
294. Alizadeh, A. *et al.* The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harb Symp Quant Biol* **64**, 71-8 (1999).
295. Robertson, P.B., Neiman, R.S., Worapongpaiboon, S., John, K. & Orazi, A. 013 (CD99) positivity in hematologic proliferations correlates with TdT positivity. *Mod Pathol* **10**, 277-82 (1997).
296. Grogan, T.M., Warnke, R.A. & Kaplan, H.S. A comparative study of Burkitt's and non-Burkitt's "undifferentiated" malignant lymphoma: immunologic, cytochemical, ultrastructural, cytologic, histopathologic, clinical and cell culture features. *Cancer* **49**, 1817-28 (1982).
297. Bodoor, K., Matalka, I., Hayajneh, R., Haddad, Y. & Gharaibeh, W. Evaluation of BCL-6, CD10, CD138 and MUM-1 expression in diffuse large B-cell lymphoma patients: CD138 is a marker of poor prognosis. *Asian Pac J Cancer Prev* **13**, 3037-46 (2012).
298. Muris, J.J. *et al.* Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *J Pathol* **208**, 714-23 (2006).

299. Ohshima, K. *et al.* CD10 and Bcl10 expression in diffuse large B-cell lymphoma: CD10 is a marker of improved prognosis. *Histopathology* **39**, 156-62 (2001).
300. He, X. *et al.* Ki-67 is a valuable prognostic predictor of lymphoma but its utility varies in lymphoma subtypes: evidence from a systematic meta-analysis. *BMC Cancer* **14**, 153 (2014).
301. L, K. Ki67 zur Vorhersage des Therapieansprechens bei neoadjuvanten Mammakarzinomen. *Senologie- Zeitschrift für Mammadiagnostik und -therapie* (2011).
302. Kim, S.J. *et al.* Ki-67 expression is predictive of prognosis in patients with stage I/II extranodal NK/T-cell lymphoma, nasal type. *Ann Oncol* **18**, 1382-7 (2007).
303. Yoon, D.H. *et al.* Ki-67 expression as a prognostic factor in diffuse large B-cell lymphoma patients treated with rituximab plus CHOP. *Eur J Haematol* **85**, 149-57 (2010).
304. Broyde, A. *et al.* Role and prognostic significance of the Ki-67 index in non-Hodgkin's lymphoma. *Am J Hematol* **84**, 338-43 (2009).
305. Miles R, R.M., Mc Carthy K, et al. Diffuse large B-cell lymphomas in pediatric patients demonstrate a marked predominance of germinal center cell phenotype [abstract]. *Ann Oncol.* (2005).
306. Lenz, G. Insights into the Molecular Pathogenesis of Activated B-Cell-like Diffuse Large B-Cell Lymphoma and Its Therapeutic Implications. *Cancers (Basel)* **7**, 811-22 (2015).
307. Miyazaki, K. [Diffuse large B-cell lymphoma: therapeutic development based on clinical and biological heterogeneity]. *Rinsho Ketsueki* **56**, 2047-55 (2015).
308. Armitage, J.O. My treatment approach to patients with diffuse large B-cell lymphoma. *Mayo Clin Proc* **87**, 161-71 (2012).
309. Wilson, W.H. The Bruton's tyrosine kinase inhibitor ibrutinib (PCI-32765) has preferential activity in the ABC subtype of relapsed/refractory de novo diffuse large B-cell lymphoma: Interim results of a multicenter, open-label phase 2 study. *Blood* **120:686 [abstract]**(2012).
310. Divids. The single agent BCL-2 inhibitor ABT 199 (GDC 0199) in patients with relapsed/refractory Non-Hodgkin lymphoma (NHL): Responses observed in all mantle cell lymphoma patients. *Blood* **122:1789 [abstract]**(2013).
311. Scott, D.W. *et al.* Prognostic Significance of Diffuse Large B-Cell Lymphoma Cell of Origin Determined by Digital Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies. *J Clin Oncol* **33**, 2848-56 (2015).
312. Masque-Soler, N. *et al.* Clinical and pathological features of Burkitt lymphoma showing expression of BCL2--an analysis including gene expression in formalin-fixed paraffin-embedded tissue. *Br J Haematol* **171**, 501-8 (2015).
313. Brown, E.J. *et al.* A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature* **369**, 756-8 (1994).
314. Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev* **18**, 1926-45 (2004).
315. Chiang, G.G. & Abraham, R.T. Phosphorylation of mammalian target of rapamycin (mTOR) at Ser-2448 is mediated by p70S6 kinase. *J Biol Chem* **280**, 25485-90 (2005).
316. Copp, J., Manning, G. & Hunter, T. TORC-specific phosphorylation of mammalian target of rapamycin (mTOR): phospho-Ser2481 is a marker for intact mTOR signaling complex 2. *Cancer Res* **69**, 1821-7 (2009).
317. Powell, J.D., Pollizzi, K.N., Heikamp, E.B. & Horton, M.R. Regulation of immune responses by mTOR. *Annu Rev Immunol* **30**, 39-68 (2012).
318. Soliman, G.A. The role of mechanistic target of rapamycin (mTOR) complexes signaling in the immune responses. *Nutrients* **5**, 2231-57 (2013).
319. Lee, J.S., Vo, T.T. & Fruman, D.A. Targeting mTOR for the treatment of B cell malignancies. *Br J Clin Pharmacol* (2016).
320. Xu, K., Liu, P. & Wei, W. mTOR signaling in tumorigenesis. *Biochim Biophys Acta* **1846**, 638-54 (2014).
321. Okosun, J. *et al.* Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma. *Nat Genet* **48**, 183-8 (2016).
322. Gundem, G. *et al.* IntOGen: integration and data mining of multidimensional oncogenomic data. *Nat Methods* **7**, 92-3 (2010).

323. Bove, J., Martinez-Vicente, M. & Vila, M. Fighting neurodegeneration with rapamycin: mechanistic insights. *Nat Rev Neurosci* **12**, 437-52 (2011).
324. Battelli, C. & Cho, D.C. mTOR inhibitors in renal cell carcinoma. *Therapy* **8**, 359-367 (2011).
325. Macdonald, A.S. Use of mTOR inhibitors in human organ transplantation. *Expert Rev Clin Immunol* **3**, 423-36 (2007).
326. Garcia-Sanz, R. *et al.* Deletions and rearrangements of cyclin-dependent kinase 4 inhibitor gene p16 are associated with poor prognosis in B cell non-Hodgkin's lymphomas. *Leukemia* **11**, 1915-20 (1997).
327. Guney, S. *et al.* Several mechanisms lead to the inactivation of the CDKN2A (P16), P14ARF, or CDKN2B (P15) genes in the GCB and ABC molecular DLBCL subtypes. *Genes Chromosomes Cancer* **51**, 858-67 (2012).
328. Jardin, F. *et al.* Diffuse large B-cell lymphomas with CDKN2A deletion have a distinct gene expression signature and a poor prognosis under R-CHOP treatment: a GELA study. *Blood* **116**, 1092-104 (2010).
329. Dustin, M.L., Selvaraj, P., Mattaliano, R.J. & Springer, T.A. Anchoring mechanisms for LFA-3 cell adhesion glycoprotein at membrane surface. *Nature* **329**, 846-8 (1987).
330. Springer, T.A., Dustin, M.L., Kishimoto, T.K. & Marlin, S.D. The lymphocyte function-associated LFA-1, CD2, and LFA-3 molecules: cell adhesion receptors of the immune system. *Annu Rev Immunol* **5**, 223-52 (1987).
331. Bolhuis, R.L., Roozmond, R.C. & van de Griend, R.J. Induction and blocking of cytolysis in CD2+, CD3- NK and CD2+, CD3+ cytotoxic T lymphocytes via CD2 50 KD sheep erythrocyte receptor. *J Immunol* **136**, 3939-44 (1986).
332. Kanner, S.B., Damle, N.K., Blake, J., Aruffo, A. & Ledbetter, J.A. CD2/LFA-3 ligation induces phospholipase-C gamma 1 tyrosine phosphorylation and regulates CD3 signaling. *J Immunol* **148**, 2023-9 (1992).
333. Wang, J.H. *et al.* Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors. *Cell* **97**, 791-803 (1999).
334. Schneider, M. *et al.* Alterations of the CD58 gene in classical Hodgkin lymphoma. *Genes Chromosomes Cancer* **54**, 638-45 (2015).
335. Challa-Malladi, M. *et al.* Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell* **20**, 728-40 (2011).
336. Dubois, S. *et al.* Next Generation Sequencing in Diffuse Large B Cell Lymphoma Highlights Molecular Divergence and Therapeutic Opportunities: a LYSA Study. *Clin Cancer Res* (2016).
337. Kretzmer, H. *et al.* DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. *Nat Genet* **47**, 1316-25 (2015).
338. Horvath, S. Erratum to: DNA methylation age of human tissues and cell types. *Genome Biol* **16**, 96 (2015).
339. Jiang, Y., Dominguez, P.M. & Melnick, A. The many layers of epigenetic dysfunction in B-cell lymphomas. *Curr Opin Hematol* (2016).
340. Jiang, Y. *et al.* Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas. *Genome Biol* **15**, 432 (2014).
341. Ansell, S.M. Where Do Programmed Death-1 Inhibitors Fit in the Management of Malignant Lymphoma? *J Oncol Pract* **12**, 101-6 (2016).
342. Ansell, S.M. *et al.* PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med* **372**, 311-9 (2015).
343. Bibikova, M. & Fan, J.B. Genome-wide DNA methylation profiling. *Wiley Interdiscip Rev Syst Biol Med* **2**, 210-23 (2010).
344. Nagalakshmi, U., Waern, K. & Snyder, M. RNA-Seq: a method for comprehensive transcriptome analysis. *Curr Protoc Mol Biol* **Chapter 4**, Unit 4 11 1-13 (2010).
345. Sheaffer, K.L. & Schug, J. ChIP-Seq: Library Preparation and Sequencing. *Methods Mol Biol* **1402**, 101-17 (2016).

346. Buenrostro, J.D., Wu, B., Chang, H.Y. & Greenleaf, W.J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 1-9 (2015).

9 Anhang

Tabelle A 1: Genliste der abgedeckten Gene und Regionen im Lymphom Panel

| Gen | Target-Region | Transkript | Gen | Target-Region | Transkript |
|----------------------|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------|-------------------------------------------------------------------|
| <i>ARID1A</i> | CDS/Exon | NM-006015, NM-139135 | <i>KDM2B</i> | CDS/Exon | NM-001005366, NM-032590 |
| <i>B2M</i> | UTR,CDS/Exon | NM-004048 | <i>LLT1</i> | SHM | NM-001004419, NM-001197317, NM-001197318, NM-001197319, NM-013269 |
| <i>BACH2</i> | SHM | NM-001170794, NM-021813 | <i>LPHN2</i> | CDS/Exon | NM-012302 |
| <i>BCL2</i> | SHM,CDS/Exon | NM-000633, NM-000657 | <i>LRMP</i> | SHM | NM-001204126, NM-001204127, NM-006152 |
| <i>BCL6</i> | SHM,CDS/Exon | NM-001130845, NM-001134738, NM-001706 | <i>LRP1B</i> | CDS/Exon | NM-018557 |
| <i>BCL7A</i> | SHM | NM-001024808, NM-020993 | <i>LTB</i> | SHM | NM-002341, NM-009588 |
| <i>BIRC3</i> | SHM | NM-001165, NM-182962 | <i>MALAT1</i> | SHM | |
| <i>BRAF</i> | hotspot p.V600E | NM-004333 | <i>MAP3K7</i> | CDS/Exon | NM-003188, NM-145331, NM-145332, NM-145333 |
| <i>BTG1</i> | SHM,CDS/Exon | NM-001731 | <i>MEF2B</i> | UTR,CDS/Exon | NM-001145785 |
| <i>BTG2</i> | SHM | NM-006763 | <i>KMT2D</i> | UTR,CDS/Exon | NM-003482 |
| <i>CARD11</i> | UTR,CDS/Exon | NM-032415 | <i>MLL3</i> | CDS/Exon | NM-170606 |
| <i>CCND3</i> | CDS/Exon | NM-001136017, NM-001136125, NM-001136126, NM-001760 | <i>MS4A1</i> | SHM | NM-021950, NM-152866 |
| <i>CD58</i> | SHM,CDS/Exon | NM-001144822, NM-001779 | <i>MTOR</i> | UTR,CDS/Exon | NM-004958 |
| <i>CD74</i> | SHM | NM-001025158, NM-001025159, NM-004355 | <i>MUC16</i> | CDS/Exon | NM-024690 |
| <i>CD79A</i> | SHM,CDS/Exon | NM-001783, NM-021601 | <i>MYC</i> | SHM,CDS/Exon | NM-002467 |
| <i>CD79B</i> | SHM,CDS/Exon | NM-000626, NM-001039933, NM-021602 | <i>MYD88</i> | Hotspot p.L265 | NM-001172566, NM-001172567, NM-001172568, NM-001172569, NM-002468 |
| <i>CD83</i> | SHM | NM-001040280, NM-001251901, NM-004233 | <i>MYOM2</i> | CDS/Exon | NM-003970 |
| <i>CDC123</i> | CDS/Exon | NM-006023 | <i>NCOA3</i> | SHM | NM-001174087, NM-001174088, NM-006534, NM-181659 |
| <i>CDC73</i> | CDS/Exon | NM-024529 | <i>NOTCH1</i> | CDS/Exon | NM-017617 |
| <i>CDKN2A</i> | CDS/Exon | NM-000077, NM-001195132, NM-058195, NM-058197 | <i>P2RY8</i> | SHM | NM-178129 |
| <i>CIITA</i> | SHM,UTR,CDS/Exon | NM-000246 | <i>PAX5</i> | SHM | NM-016734 |
| <i>CREBBP</i> | CDS/Exon | NM-001079846, NM-004380 | <i>PCBP1</i> | CDS/Exon | NM-006196 |
| <i>CXCR4</i> | SHM | NM-001008540, NM-003467 | <i>PCLO</i> | CDS/Exon | NM-014510, NM-033026 |
| <i>DDX3X</i> | CDS/Exon | NM-001193416, NM-001193417, NM-001356 | <i>PIM1</i> | SHM,CDS/Exon | NM-001243186, NM-002648 |
| <i>DMD</i> | SHM | NM-000109, NM-004006, NM-004007, NM-004009, NM-004010, NM-004011, NM-004012, NM-004013, NM-004014, NM-004015, NM-004016, NM-004017, NM-004018, NM-004019, NM-004020, NM-004021, NM-004022, NM-004023 | <i>POU2AF1</i> | SHM | NM-006235 |
| <i>DMXL1</i> | CDS/Exon | NM-005509 | <i>PRDM1</i> | CDS/Exon | NM-001198, NM-182907 |
| <i>DNMT1</i> | CDS/Exon | NM-001130823, NM-001379 | <i>PTPRT</i> | CDS/Exon | NM-007050, NM-133170 |
| <i>DNMT2/TRDM T1</i> | CDS/Exon | NM-004412 | <i>REL</i> | SHM | NM-002908 |
| <i>DNMT3a</i> | CDS/Exon | NM-022552, NM-153759, NM-175629, NM-175630 | <i>RELN</i> | CDS/Exon | NM-005045, NM-173054 |

Anhang

| Gen | Target-Region | Transkript | Gen | Target-Region | Transkript |
|------------------|----------------|------------------------------------------------------------------------|---------------------|---------------|-------------------------------------------------------------------------------------------------------------|
| <i>DNMT3b</i> | CDS/Exon | NM-001207055, NM-001207056, NM-006892, NM-175848, NM-175849, NM-175850 | <i>RHOA</i> | CDS/Exon | NM-001664 |
| <i>DSP</i> | CDS/Exon | NM-001008844, NM-004415 | <i>RHOH</i> | SHM | NM-004310 |
| <i>DTX1</i> | SHM,CDS/Exon | NM-004416 | <i>SIPR2</i> | SHM | NM-004230 |
| <i>DUSP2</i> | SHM | NM-004418 | <i>SERPINA9</i> | SHM | NM-001042518, NM-175739 |
| <i>EGFR</i> | CDS/Exon | NM-005228, NM-201282, NM-201283, NM-201284 | <i>SF3B1</i> | CDS/Exon | NM-001005526, NM-012433 |
| <i>EP300</i> | CDS/Exon | NM-001429 | <i>SGK1</i> | SHM,CDS/Exon | NM-001143676, NM-001143677, NM-001143678, NM-005627 |
| <i>EP300</i> | UTR,CDS/Exon | NM-001429 | <i>SMARCA4</i> | CDS/Exon | NM-001128844, NM-001128845, NM-001128846, NM-001128847, NM-001128848, NM-001128849, NM-003072 |
| <i>ETS1</i> | SHM | NM-001143820, NM-001162422, NM-005238 | <i>SOCS1</i> | SHM,CDS/Exon | NM-003745 |
| <i>EZH2</i> | HotSpot p,Y461 | NM-001203247, NM-001203248, NM-001203249, NM-004456, NM-152998 | <i>SON</i> | CDS/Exon | NM-032195, NM-138927 |
| <i>FAS</i> | SHM,CDS/Exon | NM-000043, NM-152871, NM-152872 | <i>SPRED2</i> | SHM | NM-001128210, NM-181784 |
| <i>FAT1</i> | CDS/Exon | NM-005245 | <i>ST6GAL1</i> | SHM | NM-003032, NM-173216, NM-173217 |
| <i>FAT4</i> | CDS/Exon | NM-024582 | <i>STAT3</i> | CDS/Exon | NM-003150, NM-139276, NM-213662 |
| <i>FBXO11</i> | CDS/Exon | NM-001190274, NM-025133 | <i>STAT5/STAT5b</i> | CDS/Exon | NM-012448 |
| <i>FN1</i> | CDS/Exon | NM-002026, NM-054034, NM-212474, NM-212476, NM-212478, NM-212482 | <i>TCF3</i> | CDS/Exon | NM-001136139, NM-003200 |
| <i>FOXO1</i> | CDS/Exon | NM-002015 | <i>TCL1A</i> | SHM | NM-001098725, NM-021966 |
| <i>GADD45B</i> | SHM | NM-015675 | <i>TET1</i> | CDS/Exon | NM-030625 |
| <i>GNA13</i> | CDS/Exon | NM-006572 | <i>TET2</i> | CDS/Exon | NM-001127208, NM-017628 |
| <i>GRHR</i> | SHM | NM-012203 | <i>TFAP4</i> | CDS/Exon | NM-003223 |
| <i>HIST1H1C</i> | CDS/Exon | NM-005319 | <i>TMEM30A</i> | CDS/Exon | NM-001143958, NM-018247 |
| <i>HIST1H1E</i> | CDS/Exon | NM-005321 | <i>TMSB4X</i> | SHM | NM-021109 |
| <i>HIST1H2AC</i> | SHM | NM-003512 | <i>TNFAIP3</i> | CDS/Exon | NM-001270507, NM-001270508, NM-006290 |
| <i>HMCN1</i> | CDS/Exon | NM-031935 | <i>TNFRSF11A</i> | CDS/Exon | NM-001270949, NM-001270950, NM-001270951, NM-003839 |
| <i>ID3</i> | UTR,CDS/Exon | NM-002167 | <i>TNFRSF14</i> | CDS/Exon | NM-003820 |
| <i>IDH1</i> | CDS/Exon | NM-005896 | <i>TP53</i> | CDS/Exon | NM-000546, NM-001126112, NM-001126113, NM-001126114, NM-001126115, NM-001126116, NM-001126117, NM-001126118 |
| <i>IDH2</i> | CDS/Exon | NM-002168 | <i>TRAF2</i> | CDS/Exon | NM-021138 |
| <i>IRF4</i> | SHM,CDS/Exon | NM-001195286, NM-002460 | <i>TRAF3</i> | CDS/Exon | NM-001199427, NM-003300, NM-145725, NM-145726, |
| <i>IRF8</i> | SHM | NM-002163 | <i>TRAF5</i> | CDS/Exon | NM-001033910, NM-004619, NM-145759 |
| <i>JAK1</i> | CDS/Exon | NM-002227 | <i>UBE2J1</i> | SHM | NM-016021 |
| <i>JAK2</i> | CDS/Exon | NM-004972 | <i>USH2A</i> | CDS/Exon | NM-007123, NM-206933 |
| <i>JAK3</i> | CDS/Exon | NM-000215 | <i>ZFP36L1</i> | SHM | NM-001244698, NM-001244701, NM-004926 |

Abkürzungen und Verweise: CDS: (engl. coding sequence) kodierende Region im Genom. UTR: untranslatierte Region im Genom. SHM: somatisch hypermutierte Region im Genom. HotSpot: Häufig rekurrent mutierte Region im Genom.

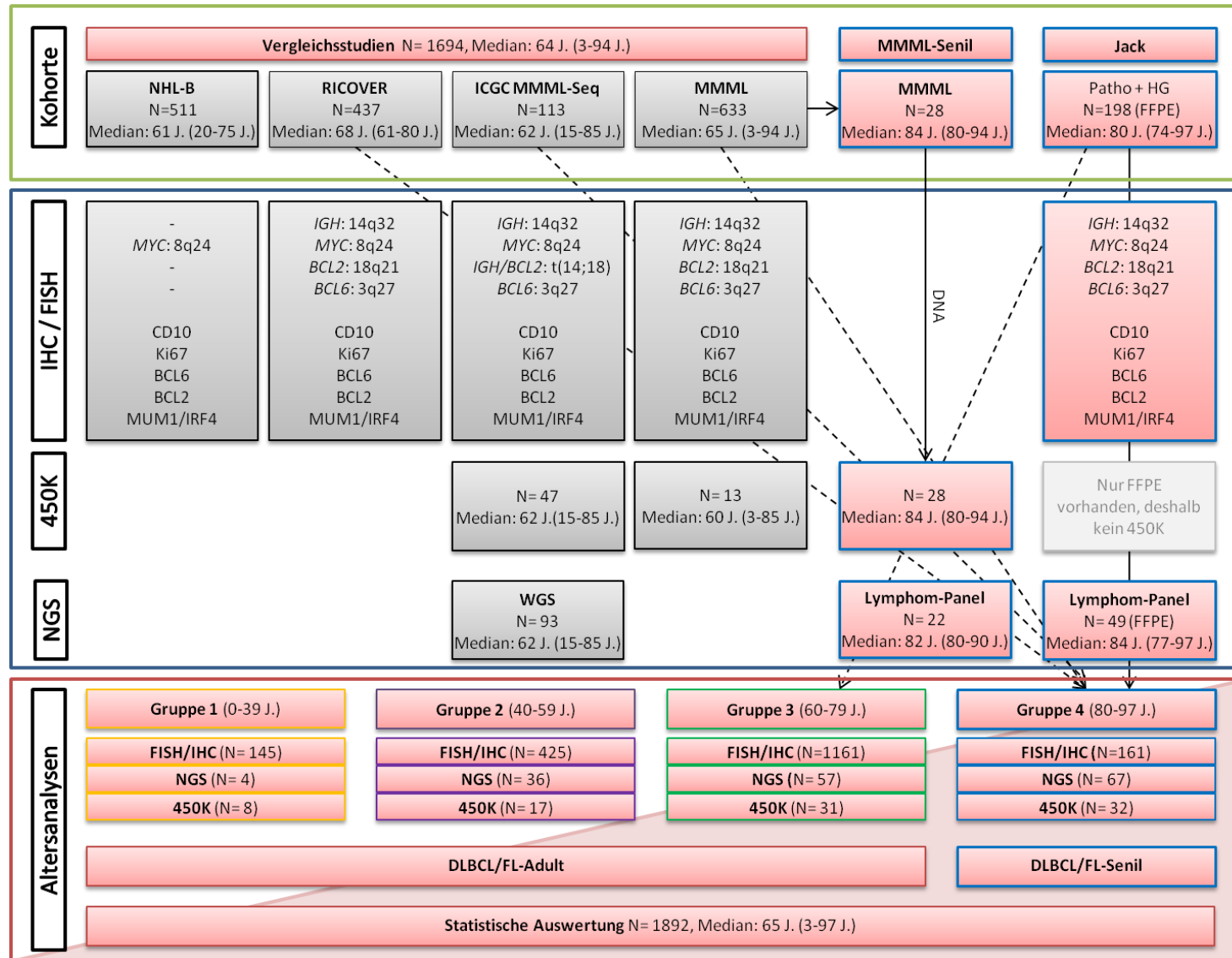


Abbildung A 1: Schematische Darstellung der analysierten Tumor-Probe Im Rahmen der vorliegenden Arbeit wurden Analysen zu altersassoziierten genetischen und epigenetischen Keimzentrums-B-Zell-Lymphomen durchgeführt. Hierzu wurde eine Kohorte von 198 Patienten (FFPE-Material) mit einem Ersterkrankungsalter ≥ 74 Jahre aus dem Untersuchungsmaterial des Instituts für Humangenetik und dem Institut für Pathologie zusammengestellt (*Jack*). Zusätzlich wurden 28 Tumor-Proben (Frischmaterial) von senilen Patienten aus dem MMML-Verbundprojekt detaillierter charakterisiert (*MMML-Senil*). Zum Vergleich der Altersassoziation wurden existierende Daten von 1694 Patienten mit einem Keimzentrums-B-Zell-Lymphom aus den klinischen (NHL-B, RICOVER) sowie Forschungsstudien (MMML, ICGC MMML-Seq) genutzt. Das Gesamtkollektiv aller nicht senilen Patienten dieser 4 Studien wurde als *DLBCL/FL-Adult* bezeichnet. Die senilen Patienten aus allen analysierten Studien (≥ 80 Jahre) wurden in den Analysen als eine Kohorte zusammengefasst und als *DLBCL/FL-Senil* bezeichnet (N= 161). Insgesamt wurden in den Altersanalysen FISH-, IHC-, NGS- und 450K-Daten verglichen und statistisch ausgewertet. Hierbei wurden die analysierten Proben in vier Altersgruppen nach dem Ersterkrankungsalter der Patienten eingeteilt: 1.) 0-39 Jahre (gelber Rahmen), 2.) 40-59 Jahre (violetter Rahmen), 3.) 60-79 Jahre (grüner Rahmen) und 4.) 80-97 Jahre (blauer Rahmen).

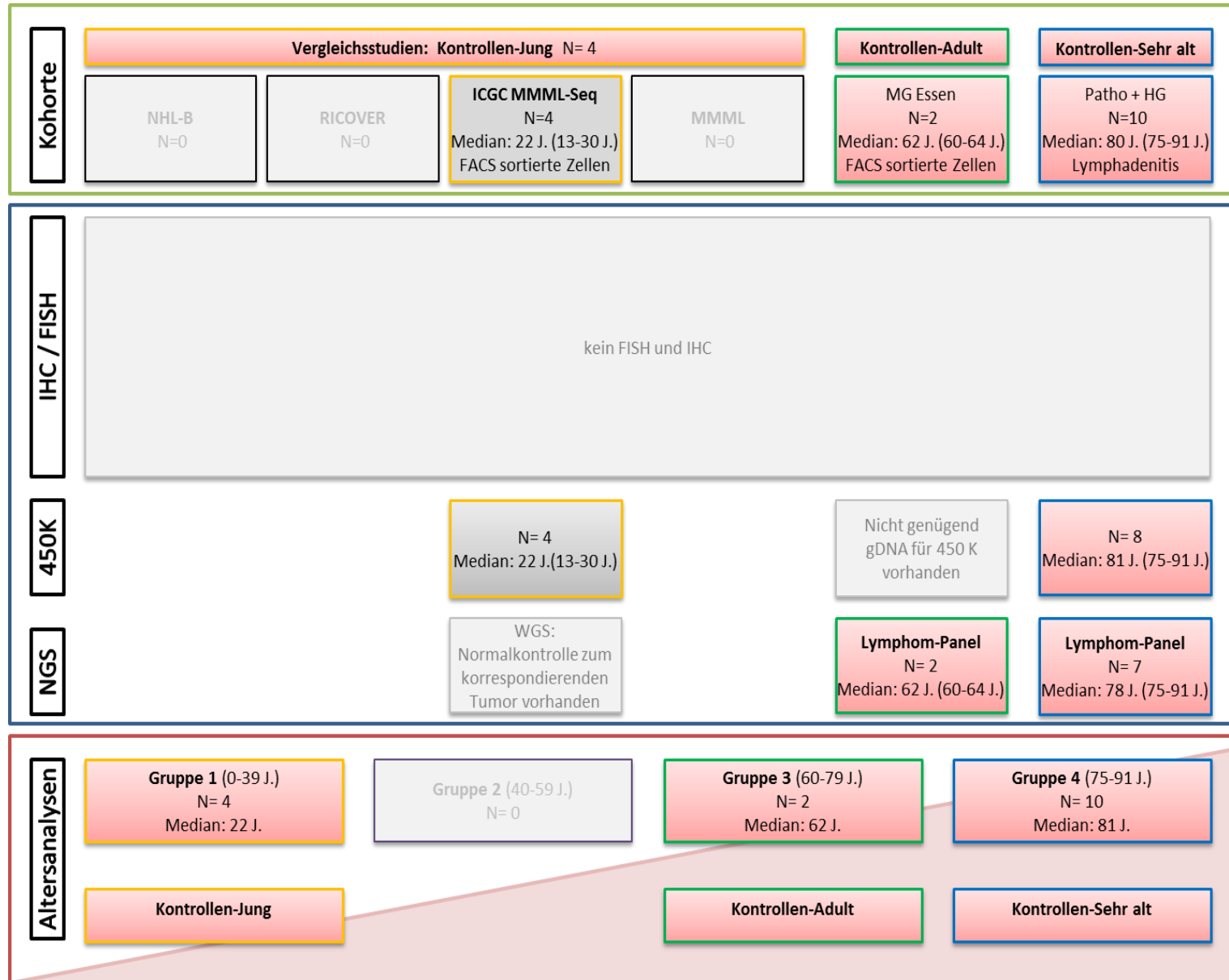


Abbildung A 2: Schematische Darstellung der analysierten Normalkontrollen In den Analysen zu altersassoziierten genetischen und epigenetischen Aberrationen in Keimzentrums-B-Zell-Lymphomen wurden zu Vergleichszwecken zusätzlich Untersuchungen an gesunden Probanden (Normalkontrollen) durchgeführt. Hierzu wurden drei Kollektive von Probanden verwendet: 1.) Kontrollen-Jung: FACS sortierte B- und T-Zellen von Probanden im Alter zwischen 0 und 39 Jahre (gelber Rahmen), 2.) Kontrollen-Adult: FACS sortierte B-Zellen von Probanden im Alter zwischen 40 und 59 Jahre (grüner Rahmen) und 3.) Kontrollen-Sehr alt: Lymphadenitis-Proben von Patienten im Alter zwischen 75 und 91 Jahre (blauer Rahmen). Je nach Verfügbarkeit der DNA-Menge und des Ausgangsmaterials wurden die Proben der Normalkontroll-Kollektive in der Auswertung verschiedener Analysen genutzt. Die Kontrollen-Sehr alt wurden mittels 450K und NGS analysiert. Hingegen wurden für die Kontrollen-Adult nur Untersuchungen mit NGS und für die Kontrollen-Jung nur mit 450K durchgeführt. FISH- und IHC-Daten wurden in keinem der Normalkontroll-Kollektive erhoben.

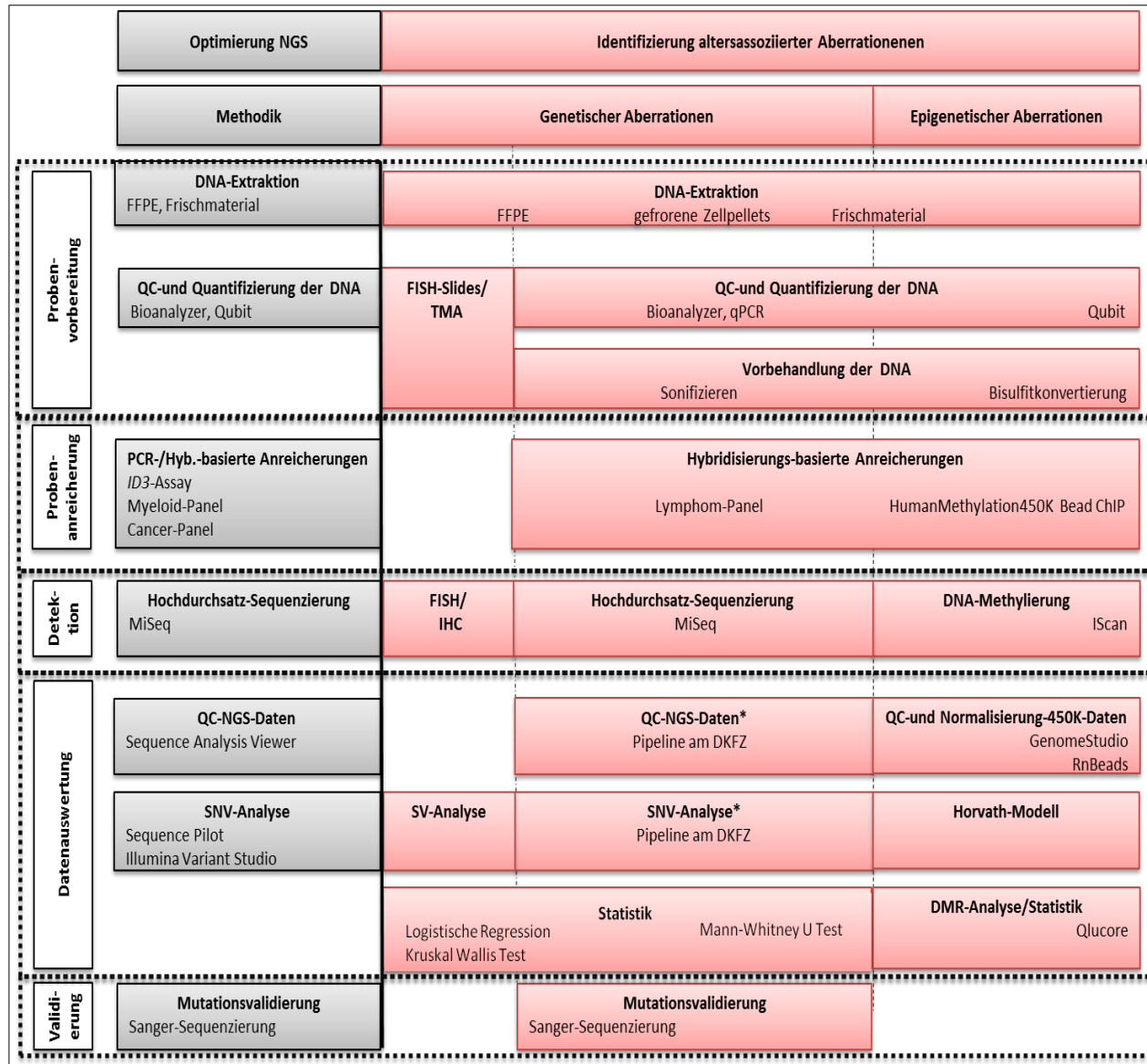


Abbildung A 3: Schematische Darstellung der verwendeten Methoden und Analysen In der vorliegenden Arbeit wurden zwei komplexe Aufgabenstellungen bearbeitet : 1.) Optimierung der Hochdurchsatz-Sequenzierung (grau markiert) und 2.) Identifizierung altersassoziierter Aberrationen in Keimzentrums-B-Zell-Lymphomen (rot markiert). Zur Bearbeitung der Themengebiete wurden verschiedene Verfahren der Probenvorbereitung, Anreicherung, Detektion, Datenauswertung und Validierung verwendet. Die Optimierung der Hochdurchsatz-Sequenzierung erfolgte an Frisch- und FFPE-Material., wobei die Qualitätskontrolle und Quantifizierung der DNA mittels Bioanalyzer und Qubit durchgeführt wurde. Es wurden drei verschiedene Panel/Assay optimiert und mit dem MiSeq Gerät sequenziert. Die Auswertung und Qualitätskontrolle der Sequenzdaten erfolgte mit unterschiedlichen Software-Programmen. Zur Validierung der detektierten Mutationen wurde als weitere Methode die Sanger-Sequenzierung genutzt. Zur Identifizierung der altersassozierten Aberrationen wurden genetische und epigenetische Veränderungen analysiert. Als Ausgangsmaterial für die Analysen der genetischen Aberrationen wurden gefrorene Zellpellets, FFPE- und Frischmaterial verwendet und die Qualitätskontrolle und Quantifizierung der DNA erfolgte mit dem Bioanalyzer und qPCR. Die gDNA wurde mittels Sonicator fragmentiert und eine Anreicherung der Target-spezifischen Regionen mit dem Lymphom Panel durchgeführt. Die Mutationsanalyse der Sequenzdaten wurden am DKFZ vorgenommen. Zur statistischen Auswertung wurden die logistische Regression, der Kruskal Wallis und Mann-Whitney U Test verwendet. Zur Analyse von strukturellen Veränderungen wurden FISH-Slides und TMAs von Tumoren aus FFPE erstellt und einzelne Marker mittels FISH und IHC analysiert. Die epigenetischen Aberrationen wurden nur an Frischmaterial analysiert, wobei die Quantifizierung mit dem Qubit erfolgt. Nach der Bisulfitkonvertierung wurden die Proben mit dem HumanMethylation450 Bead ChIP analysiert. Die Auswertung erfolgte unter Anwendung verschiedener Software-Programme. Parallel dazu wurde das epigenetische Alter der analysierten Proben mit dem Horvath-Modell berechnet.

Tabelle A 2: Übersicht der Parameter und Charakteristika von den analysierten Panel und Assays, die im Rahmen der vorliegenden Arbeit optimiert wurden

| Methodenetablierung: Hochdurchsatz-Sequenzierung | | | | | | | | | | |
|--------------------------------------------------|-----------------------|-----------------|---------------------------|-------|-------------------------|------------------------|----------------------------|------------------------------|--------------------|----|
| Verfahren | PCR-basiert | | | | Hybridisierungs-basiert | | | | | |
| Custom/Hersteller | Illumina | | Custom/ <i>inhouse</i> | | | Illumina | | Custom/ Agilent Technologies | | |
| Panel/Assay | TruSight Amplicon | | Two-Step-Assay | | | TruSight Rapid Capture | | SureSelect XY | | |
| Gen/Gen-Sets | Myeloid | | ID3 | | | Cancer | | Lymphom | | |
| Amplikons/Oligos | 573 | | 4 | | | ~4000 | | 13456 | | |
| Analysierte Proben (N) | 39 | | 41 | | | 28 | | 80 | | |
| Ausgangsmaterial | Frischmaterial | | Frisch- und FFPE-Material | | | Frischmaterial | | Frisch- und FFPE-Material* | | |
| Input gDNA [ng] | 50 | | 50 pro PCR | | | 50 | | 200 | | |
| Polymerase | TruSight Amplicon Kit | | AccuPrime | Roche | Q5 | Phusion | TruSight Rapid Capture Kit | Herkulase | | |
| Barcodes (N) | 2 | | 2 | | | 2 | | 1 | | |
| Flow cell | Standard | | Nano | | | Standard und Nano | | Standard | | |
| Sequenzierlänge [bp] | 2x 150 | | 2x 250 | | | 2x 150 | | 2x 100 | | |
| Sequenzier-Reagenzien | v2 (300 Zyklen) | v3 (600 Zyklen) | v2 (500 Zyklen) | | | v2 (300 Zyklen) | | v2 (300 Zyklen) | | |
| Normalisierung der Library | Standard | | Standard | | | Standard | <i>Small-Scale</i> | Standard | <i>Small-Scale</i> | |
| MiSeq-Läufe gesamt (N) | 1 | 4 | 1 | | | 2 | 2 | 5 | 5 | |
| Wiederholung Sequenzierung (N) | 0 | 1 | 0 | | | 0 | 0 | 3 | 0 | |
| Total Reads | 4918393 | 19985139 | 980178 | | | 16206406 | 2973491 | 2249129 | 2811712 | |
| Ø On-Target-Rate [%] | 99,5 | 98,6 | 62,5 | 25 | 0 | 87,5 | 100 | 99,9 | 47,6 | 42 |
| Ø On-Target-Coverage | 4358x | 4045x | 4179x [#] | | | 544x | 272x | 40x | 33x | |
| Ø PCR-Duplikate (Ø) | NA | | NA | | | 4 | 6 | 21 | 30 | |
| Ø Q30 [%] | 98 | 85 | 90 | | | 93 | 94 | 94 | 92 | |
| Ø Cluster PF [%] | 96 | 91 | 85 | | | 87 | 86 | 86 | 83 | |
| Ø Fehlerrate PhiX [%] | kein PhiX verwendet | | 1 | | | 0,41 | 0,32 | 0,24 | 0,29 | |

Anhang

| Methodenetablierung: Hochdurchsatz-Sequenzierung | | | | | | |
|--------------------------------------------------|---------------------------------------------------------------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-----|-------|
| Ø Fragmentlänge [bp] | 246 | 388 | 213 | 213 | 167 | 142 |
| Validierung (Sanger-Sequenzierung) | 17/17 | 8/8 | 13/13 | keine Mutation | NA | 11/11 |
| Auswertungssoftware/ | Illumina Variant Studio 2,0 | Sequence Pilot Modul SeqNext | Sequence Pilot Modul SeqNext | DKFZ Pipeline | | |
| Vorteil | kommerziell erhältlich, variabel (1-12 Proben/ Lauf) | FFPE-geeignet, Tiefensequenzierung (>1000x) von bis zu 96 Proben in einem Lauf | kommerziell erhältlich, Probenanzahl variabel (1-12 Proben/ Lauf) | FFPE-geeignet, Probenanzahl variabel (1-16 Proben/ Lauf) | | |
| Nachteil | nicht FFPE-geeignet | Mutationen in der Splice-Region werden mit einer geringen MAF <10% detektiert | nicht FFPE-geeignet, Coverage geringer bei der Sequenzierung von wenigen Proben (Nano Kit) | Niedrige On-Target-Rate (<50%) | | |
| Varierte Parameter | Sequenzier-Reagenzien (v2 und v3) | Polymerasen, Ausgangsmaterial | Probenanzahl pro Lauf, Normalisierung der Library | Ausgangsmaterial (FFPE nur mit Small-Scale-Protokoll) getestet, Normalisierung der Library | | |
| Schlussfolgerung für weitere Analysen | <4 Proben werden mit v2 und >4 Proben mit v3 Reagenzien sequenziert | Library Präparation mit der AccuPrime Polymerase durchführen | erzielte Coverage ausreichend für die <i>inhouse</i> Anforderung in der Routine Diagnostik (≥10 x) | Optimierung des Panel-Designs und Protokolls erforderlich, um die On-Target-Rate zu erhöhen | | |
| Paper/ Anwendung | Mutationsanalyse T-PLL Lopez et al., 2015 | Mutationsanalyse Burkitt-Lymphome | Mutationsanalyse verschiedener Tumore, u.a. Routine-Diagnostik <i>BRCA1/2</i> | Altersassoziierte Analysen von Keimzentrums-B-Zell-Lymphomen | | |

Abkürzungen und Verweise: * FFPE-Material wurde nur mit dem Small-Scale-Protokoll analysiert, #: On-Target-Coverage (Ø) basiert auf den sequenzierten Proben, die mittels AccuPrime vorbereitet wurden

Tabelle A 3: Qualitätsparameter der analysierten Hochdurchsatz-Sequenzierung mit dem Lymphom-Panel

| PID | Gruppe | Normalisierung | TOTAL-READS | MAPPING-RATE [%] | TOTAL-MAPPED-READS | DUPLICATE-RATE [%] | NON-DUPLICATED-READS | ON-TARGET-RATE [%] | ON-TARGET-READS | Ø TARGET-COVERAGE | Ø Target-Größe |
|--------------|-----------|----------------|-------------|------------------|--------------------|--------------------|----------------------|--------------------|-----------------|-------------------|----------------|
| control | Kontrolle | 2 | 11631754 | 99,9 | 11615062 | 27,3 | 8445541 | 37,1 | 4314233 | 105,9 | NA |
| Jack-36 | Tumor | 2 | 4509430 | 99,9 | 4502716 | 26,2 | 3324478 | 53,2 | 2394327 | 83,8 | 152 |
| Jack-MPI-859 | Tumor | 1 | 3313580 | 99,9 | 3309250 | 20,5 | 2630378 | 56,4 | 1866757 | 73,5 | 161 |
| Jack-NK-3 | Tumor | 2 | 3679885 | 100 | 3678845 | 11,5 | 3254206 | 42,4 | 1560734 | 68,2 | 160 |
| Jack-56 | Tumor | 2 | 4843160 | 100 | 4841661 | 25 | 3632370 | 43 | 2084110 | 63,6 | 138 |
| Jack-MPI-883 | Tumor | 1 | 2982540 | 100 | 2981342 | 11,3 | 2643372 | 48,2 | 1437456 | 63,6 | 168 |
| Jack-39 | Tumor | 2 | 2617269 | 99,8 | 2613209 | 9,6 | 2362190 | 51,3 | 1340931 | 60,9 | 158 |
| Jack-MPI-119 | Tumor | 1 | 2925259 | 99,8 | 2920731 | 7 | 2715111 | 42,7 | 1246140 | 60,5 | 165 |
| Jack-NK-10 | Tumor | 2 | 2953470 | 100 | 2952373 | 7,7 | 2725985 | 40,8 | 1206030 | 55,8 | 153 |
| Jack-23 | Tumor | 2 | 3064620 | 99,8 | 3059928 | 19,9 | 2449849 | 45,9 | 1405344 | 52,9 | 146 |
| Jack-5 | Tumor | 2 | 2613640 | 99,8 | 2608993 | 19,1 | 2111618 | 49,1 | 1280762 | 49,5 | 155 |
| Jack-35 | Tumor | 2 | 2465279 | 99,9 | 2461662 | 13 | 2140768 | 47,5 | 1170271 | 48 | 156 |
| Jack-20 | Tumor | 2 | 2732065 | 99,9 | 2728399 | 21,1 | 2152158 | 49,1 | 1338750 | 47,9 | 157 |
| Jack-18 | Tumor | 2 | 2765243 | 99,9 | 2761272 | 26 | 2044486 | 49 | 1352515 | 47,7 | 156 |
| Jack-59 | Tumor | 2 | 3051117 | 99,8 | 3045970 | 24,3 | 2305959 | 45,8 | 1396559 | 47,6 | 146 |
| Jack-MPI-768 | Tumor | 1 | 3052855 | 99,9 | 3049429 | 19,1 | 2465715 | 38,1 | 1162714 | 47,3 | 159 |
| Jack-MPI-243 | Tumor | 1 | 2117473 | 99,9 | 2115645 | 30,7 | 1465694 | 67,4 | 1426602 | 47,1 | 175 |
| Jack-1 | Tumor | 2 | 2069115 | 99,9 | 2066804 | 12,4 | 1810177 | 49,5 | 1022072 | 44,7 | 152 |
| Jack-MPI-756 | Tumor | 1 | 2507718 | 99,8 | 2503912 | 16,4 | 2092844 | 41,8 | 1047147 | 43,5 | 166 |
| Jack-MPI-118 | Tumor | 1 | 2170342 | 99,9 | 2167865 | 7,3 | 2009136 | 41,7 | 904468 | 43,1 | 170 |
| Jack-24 | Tumor | 2 | 2689544 | 99,8 | 2685473 | 26,9 | 1962526 | 46,1 | 1237285 | 42,9 | 150 |
| Jack-MPI-201 | Tumor | 1 | 2057153 | 99,9 | 2054626 | 31,2 | 1412950 | 64,9 | 1334135 | 42,6 | 170 |
| Jack-MPI-820 | Tumor | 1 | 2340373 | 99,9 | 2337797 | 17,9 | 1919029 | 43,3 | 1012940 | 42,5 | 172 |
| Jack-4 | Tumor | 2 | 2331992 | 100 | 2331185 | 12,2 | 2047023 | 43,5 | 1015069 | 40,7 | 147 |
| Jack-NK-9 | Tumor | 2 | 2362586 | 100 | 2361705 | 22,3 | 1833961 | 43 | 1014466 | 40,1 | 161 |
| Jack-33 | Tumor | 2 | 3671037 | 99,9 | 3668618 | 30,2 | 2560119 | 43,9 | 1609861 | 39,8 | 126 |
| Jack-47 | Tumor | 2 | 1858454 | 99,9 | 1856117 | 7,3 | 1721066 | 44 | 816829 | 39,3 | 159 |
| Jack-MPI-099 | Tumor | 1 | 2718929 | 99,9 | 2716004 | 3,6 | 2618138 | 27,8 | 755757 | 38,4 | 160 |
| Jack-MPI-857 | Tumor | 1 | 2031997 | 99,9 | 2030216 | 17 | 1684140 | 44,6 | 904937 | 37,3 | 165 |
| Jack-10 | Tumor | 2 | 1912873 | 100 | 1912223 | 1,5 | 1882778 | 37,6 | 718464 | 37,1 | 142 |
| Jack-43 | Tumor | 2 | 2262627 | 100 | 2261629 | 6,2 | 2121330 | 35,3 | 798536 | 36,8 | 147 |
| Jack-MPI-047 | Tumor | 1 | 2249615 | 99,9 | 2247926 | 23,9 | 1709576 | 43,1 | 968622 | 36,5 | 163 |
| Jack-MPI-331 | Tumor | 1 | 2291603 | 99,8 | 2287703 | 49,1 | 1163447 | 68,3 | 1561365 | 36,1 | 171 |
| Jack-57 | Tumor | 2 | 2241287 | 99,9 | 2238185 | 21,8 | 1749541 | 46,4 | 1039068 | 36 | 148 |
| Jack-MPI-841 | Tumor | 1 | 2470291 | 99,9 | 2468183 | 33,6 | 1638607 | 48,9 | 1207877 | 35,9 | 168 |
| Jack-NK-5 | Tumor | 2 | 1914643 | 100 | 1913922 | 12,5 | 1674869 | 43 | 822871 | 35,4 | 154 |

Anhang

| PID | Gruppe | Normalisierung | TOTAL-READS | MAPPING-RATE [%] | TOTAL-MAPPED-READS | DUPLICATE-RATE [%] | NON-DUPLICATED-READS | ON-TARGET-RATE [%] | ON-TARGET-READS | Ø TARGET-COVERAGE | Ø Target-Größe |
|--------------|--------|----------------|-------------|------------------|--------------------|--------------------|----------------------|--------------------|-----------------|-------------------|----------------|
| Jack-MPI-309 | Tumor | 1 | 2501911 | 99,9 | 2499682 | 12,2 | 2194358 | 30 | 749173 | 35,3 | 167 |
| Jack-MPI-206 | Tumor | 1 | 2253251 | 99,9 | 2250150 | 40,4 | 1340129 | 64 | 1439683 | 35,2 | 169 |
| Jack-NK-4 | Tumor | 2 | 1732309 | 100 | 1731619 | 3,7 | 1668178 | 40,5 | 701090 | 35,2 | 159 |
| Jack-55 | Tumor | 2 | 3601720 | 100 | 3600125 | 22,9 | 2776469 | 31,2 | 1124387 | 34,3 | 128 |
| Jack-MPI-191 | Tumor | 1 | 1542696 | 99,9 | 1541271 | 29 | 1094109 | 65,3 | 1005995 | 33,6 | 173 |
| Jack-MPI-762 | Tumor | 1 | 1391348 | 99,9 | 1389706 | 21,7 | 1087665 | 62,1 | 863512 | 33 | 164 |
| Jack-34 | Tumor | 2 | 2685490 | 100 | 2684637 | 17,9 | 2204738 | 35,4 | 951048 | 32,8 | 147 |
| Jack-MPI-849 | Tumor | 1 | 1289356 | 99,9 | 1287735 | 17,2 | 1066473 | 62,9 | 810104 | 32,5 | 171 |
| Jack-25 | Tumor | 2 | 2442186 | 99,9 | 2438530 | 32,5 | 1646631 | 44 | 1073555 | 32,1 | 138 |
| Jack-31 | Tumor | 2 | 1503835 | 99,9 | 1501934 | 16 | 1262135 | 51,2 | 769067 | 32,1 | 148 |
| Jack-MPI-411 | Tumor | 1 | 1996970 | 99,9 | 1994059 | 11,9 | 1756799 | 34,3 | 684799 | 31,6 | 170 |
| Jack-NK-8 | Tumor | 2 | 1522900 | 100 | 1522366 | 3,2 | 1473902 | 40,2 | 611347 | 31,1 | 162 |
| Jack-27 | Tumor | 2 | 3270213 | 99,9 | 3268531 | 41,4 | 1914732 | 45,6 | 1490005 | 29,9 | 156 |
| Jack-MPI-851 | Tumor | 1 | 2758735 | 99,9 | 2755089 | 20,1 | 2202299 | 23,2 | 639202 | 29,8 | 160 |
| Jack-200 | Tumor | 2 | 1641567 | 100 | 1640967 | 8,2 | 1506277 | 37,5 | 614955 | 27,7 | 147 |
| Jack-NK-6 | Tumor | 2 | 1452280 | 100 | 1451792 | 3,1 | 1406486 | 37,9 | 549615 | 27,5 | 154 |
| Jack-38 | Tumor | 2 | 2228879 | 99,9 | 2227381 | 35,5 | 1436148 | 49,6 | 1104887 | 26,8 | 140 |
| Jack-48 | Tumor | 2 | 4971364 | 99,9 | 4968826 | 52 | 2383918 | 44,3 | 2200771 | 26,8 | 122 |
| Jack-58 | Tumor | 2 | 2242635 | 99,9 | 2241271 | 32,9 | 1504824 | 45,1 | 1011443 | 25,5 | 160 |
| Jack-29 | Tumor | 2 | 2270234 | 99,9 | 2268790 | 32,9 | 1523252 | 40,3 | 913356 | 25 | 133 |
| Jack-44 | Tumor | 2 | 3361291 | 99,9 | 3359345 | 49,2 | 1705307 | 46,7 | 1567398 | 25 | 146 |
| Jack-32 | Tumor | 2 | 3514897 | 100 | 3513325 | 55,9 | 1549515 | 51,7 | 1816199 | 24,5 | 139 |
| Jack-41 | Tumor | 2 | 2672503 | 100 | 2671428 | 44,9 | 1470800 | 50,3 | 1342441 | 23,9 | 160 |
| Jack-61 | Tumor | 2 | 3549196 | 99,9 | 3546984 | 45,8 | 1922996 | 42,5 | 1507806 | 23,1 | 117 |
| Jack-3 | Tumor | 2 | 1586340 | 99,9 | 1585317 | 10,2 | 1423137 | 31,2 | 493922 | 21 | 157 |
| Jack-22 | Tumor | 2 | 4051503 | 99,9 | 4049222 | 48,8 | 2072474 | 38,8 | 1572489 | 20,9 | 126 |
| Jack-26 | Tumor | 2 | 3597865 | 99,9 | 3595739 | 41,2 | 2114341 | 34,2 | 1228657 | 20,5 | 127 |
| Jack-37 | Tumor | 2 | 2043909 | 100 | 2043001 | 44,3 | 1138631 | 44,4 | 907246 | 18,1 | 126 |
| Jack-60 | Tumor | 2 | 1990240 | 99,9 | 1989088 | 43,2 | 1130664 | 46,9 | 933635 | 18 | 148 |
| Jack-19 | Tumor | 2 | 2524886 | 99,9 | 2523372 | 51,1 | 1233250 | 44,7 | 1128192 | 17,9 | 127 |
| Jack-6 | Tumor | 2 | 1990953 | 99,9 | 1989916 | 31,5 | 1362237 | 36,6 | 729007 | 16,7 | 149 |
| Jack-28 | Tumor | 2 | 1603064 | 99,9 | 1601911 | 38,6 | 983418 | 37,3 | 597120 | 16 | 144 |
| Jack-50 | Tumor | 2 | 2659468 | 99,9 | 2657763 | 53,3 | 1241842 | 38,3 | 1017347 | 13,4 | 114 |
| Jack-51 | Tumor | 2 | 4871561 | 99,9 | 4869102 | 53,3 | 2271771 | 28,5 | 1386813 | 12,6 | 104 |
| Jack-54 | Tumor | 2 | 5542332 | 99,9 | 5539403 | 70,4 | 1640147 | 44,5 | 2464956 | 12 | 137 |
| Jack-52 | Tumor | 2 | 2888354 | 99,9 | 2886220 | 57,6 | 1223598 | 40,2 | 1161338 | 9,9 | 126 |
| Jack-21 | Tumor | 2 | 2293419 | 100 | 2292614 | 37,2 | 1440305 | 23,8 | 546378 | 9,6 | 117 |
| Jack-42 | Tumor | 2 | 1284286 | 99,9 | 1282896 | 39,8 | 772210 | 35,4 | 453894 | 7,9 | 108 |

Anhang

| PID | Gruppe | Normalisierung | TOTAL-READS | MAPPING-RATE [%] | TOTAL-MAPPED-READS | DUPLICATE-RATE [%] | NON-DUPLICATED-READS | ON-TARGET-RATE [%] | ON-TARGET-READS | Ø TARGET-COVERAGE | Ø Target-Größe |
|--------------|--------|----------------|-------------|------------------|--------------------|--------------------|----------------------|--------------------|-----------------|-------------------|----------------|
| Jack-53 | Tumor | 2 | 1343992 | 100 | 1343344 | 46,4 | 720582 | 38 | 510718 | 7,9 | 129 |
| Jack-40 | Tumor | 2 | 2496959 | 100 | 2495887 | 66 | 849601 | 48,2 | 1202628 | 7,7 | 129 |
| Jack-45 | Tumor | 2 | 1609536 | 99,9 | 1608454 | 48,6 | 826201 | 34,1 | 548761 | 6,5 | 126 |
| Jack-MPI-749 | Tumor | 1 | 516841 | 99,8 | 515952 | 28,8 | 367125 | 28,8 | 148713 | 6 | 167 |
| Jack-30 | Tumor | 2 | 984315 | 99,9 | 983786 | 49,3 | 498604 | 41,1 | 404671 | 5,8 | 128 |

Abkürzungen und Verweise: PID: Probenidentifikationsname. control: Pool aus den sequenzierten Proben der Kohorte Kontrollen-Adult. Normalisierung: Verwendete Protokoll zur Normalisierung der Library, 1: Standard-Protokoll, 2: Small-Scale-Protokoll. grau markiert: Alle Proben mit Ø-OnTarget-Rate < 30x. Diese Proben wurden aus der weiteren Mutationsanalyse ausgeschlossen.

Tabelle A 4: Charakteristika des Gesamtkollektives von analysierten Patienten mit einem FL

| | Gruppe 1 | | Gruppe 2 | | Gruppe 3 | | Gruppe 4 | | Total | | LR | KW |
|------------------------------------|------------|------|-------------|------|-------------|------|------------|------|-------------|------|-------------|-------------|
| Altersgruppen | 0-39 Jahre | | 40-59 Jahre | | 60-79 Jahre | | >80 Jahre | | 0-100 Jahre | | | |
| Patienten | 27 | 100% | 112 | 100% | 120 | 100% | 24 | 100% | 283 | 100% | | |
| Median Alter (min- max) | 36 (10-39) | | 51 (40-59) | | 70 (60-79) | | 83 (80-95) | | 60 (10-95) | | | |
| Geschlecht | | | | | | | | | | | | |
| Männlich | 6 | 22 | 56 | 50 | 69 | 57,5 | 14 | 58,3 | 145 | 51,2 | <0,01 | <0,01 |
| Weiblich | 21 | 78 | 56 | 50 | 51 | 42,5 | 10 | 41,7 | 138 | 48,8 | <0,01 | <0,01 |
| Fluoreszenz-in-situ-Hybridisierung | | | | | | | | | | | | |
| MYC Bruch | 0/23 | 0 | 13/108 | 12 | 10/107 | 9 | 0/15 | 0 | 23/253 | 9 | 0,07 | 0,04 |
| BCL2 Bruch* | 17/23 | 74 | 85/109 | 78 | 75/113 | 66 | 11/17 | 65 | 188/263 | 71 | 0,15 | 0,05 |
| BCL6 Bruch | 5/24 | 21 | 30/110 | 27 | 17/110 | 15 | 6/16 | 38 | 58/260 | 22 | 0,64 | 0,53 |
| IGH Bruch | 20/24 | 83 | 86/110 | 78 | 83/112 | 74 | 11/17 | 65 | 199/262 | 76 | 0,86 | 0,77 |
| Immunhistochemie | | | | | | | | | | | | |
| CD10 Expression | 22/25 | 88 | 93/105 | 89 | 100/115 | 87 | 23/24 | 96 | 238/269 | 88 | 0,78 | 0,64 |
| BCL2 Expression | 25/26 | 96 | 103/108 | 95 | 113/117 | 97 | 20/22 | 91 | 261/273 | 96 | 0,70 | 0,64 |
| BCL6 Expression | 16/16 | 100 | 83/85 | 98 | 77/81 | 95 | 12/14 | 86 | 188/196 | 96 | 0,03 | 0,05 |
| Ki67 Expression | 5/25 | 20 | 7/103 | 7 | 15/113 | 13 | 5/22 | 23 | 32/263 | 12 | 0,48 | 0,26 |
| MUM1 Expression | 8/14 | 57 | 37/68 | 54 | 43/77 | 56 | 3/6 | 50 | 91/165 | 55 | 0,62 | 0,65 |

Abkürzungen und Verweise: *: Ergebnisse für BCL2 wurden in den FISH-Analyse mittels t(14,18) oder BCL2-Break-Apart-Probe (FISH-Sonde) erhoben. neg.: negativ. pos.: positiv. NA: nicht verfügbar oder nicht auswertbar. Alter: in Jahren angegeben. IHC: Immunhistochemie.

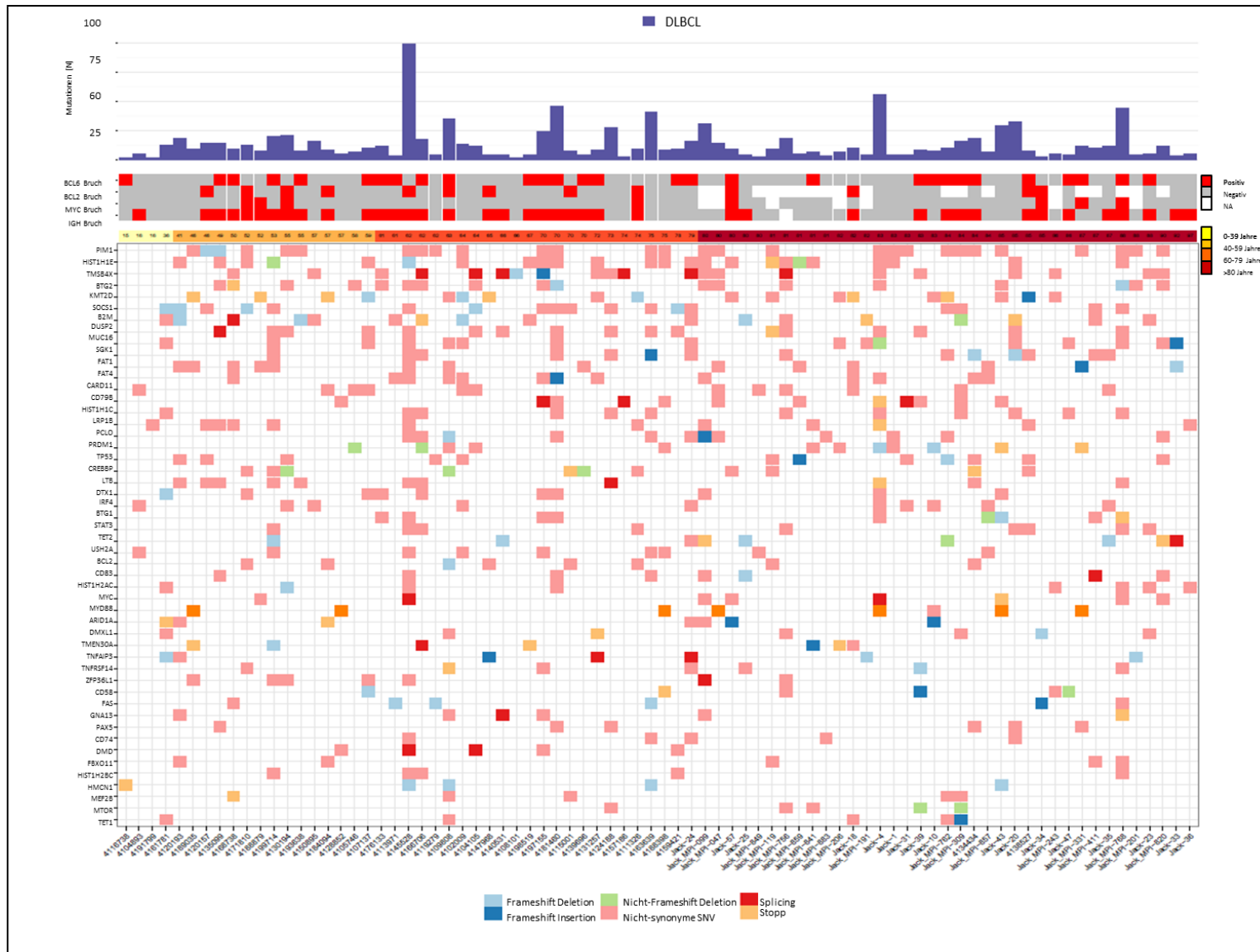


Abbildung A 4: Oncoprint der Ergebnisse aus den molekulargenetischen Untersuchungen in den analysierten DLBCL-Patienten Die erhobenen FISH- und Mutationsdaten der 84 analysierten DLBCL-Patienten (farbig nach Altersgruppen markiert) wurden in einem Oncoprint grafisch dargestellt. Die Rechtecke horizontal über der farbigen Matrix repräsentieren die analysierten Altersgruppen in einem Farbcode. Jede Spalte der farbigen Matrix stellt einen Patienten und jede Zeile ein Gen dar. Jedes farbige Rechteck innerhalb der Matrix spiegelt den Mutationstypen in einem Farbcode wider. Zusätzlich wurden in der Grafik die Ergebnisse der FISH-Analysen und die Gesamtanzahl an detektierten potenziell proteinverändernden Mutationen für jeden Patienten über der farbigen Matrix dargestellt. Insgesamt wurden 50 Gene identifiziert, bei denen in mind. 5 DLBCL-Patienten eine potenziell proteinverändernde Mutation detektiert wurde. Es war kein Zusammenhang zwischen dem Ersterkrankungsalter und der Anzahl an detektierten Mutationen der Patienten zu erkennen.

Tabelle A 5: Statistische Auswertung der rekurrent mutierten Gene in der analysierten DLBCL-Kohorte

| Gen | Min. Alter [Jahre] | Max. Alter [Jahre] | Median Alter [Jahre] | Interquartil-Abstand [Jahre] | U Test P-Wert | LR P-Wert | MF [%] |
|------------------|-----------------------|-----------------------|-------------------------|---------------------------------|------------------|--------------|-----------|
| <i>BTGI</i> | 70 | 85 | 83 | 15 | 0,52 | 0,25 | 11 |
| <i>CD58</i> | 76 | 85 | 82 | 9 | 0,28 | 0,24 | 7 |
| <i>MYC</i> | 76 | 86 | 82 | 10 | 0,08 | 0,22 | 10 |
| <i>MYD88</i> | 70 | 85 | 82 | 14 | 0,31 | 0,48 | 9 |
| <i>PRDMI</i> | 64 | 83 | 82 | 19 | 0,33 | 0,27 | 14 |
| <i>TET1</i> | 63 | 85 | 82 | 21 | 0,5 | 0,93 | 6 |
| <i>MTOR</i> | 81 | 83 | 81 | 2 | 0,15 | 0,12 | 6 |
| <i>DML1</i> | 68 | 85 | 81 | 16 | 0,52 | 0,71 | 8 |
| <i>STAT3</i> | 62 | 85 | 81 | 23 | 0,52 | 0,36 | 11 |
| <i>TP53</i> | 60 | 85 | 81 | 24 | 0,33 | 0,85 | 14 |
| <i>IRF4</i> | 59 | 83 | 81 | 24 | 0,75 | 0,79 | 12 |
| <i>FBO11</i> | 57 | 87 | 81 | 30 | 0,5 | 0,96 | 6 |
| <i>TET2</i> | 79 | 87 | 80 | 8 | 0,18 | 0,12 | 9 |
| <i>CD79B</i> | 73 | 83 | 80 | 10 | 0,33 | 0,15 | 11 |
| <i>PCLO</i> | 68 | 82 | 80 | 14 | 0,73 | 0,25 | 14 |
| <i>TNFAIP3</i> | 53 | 72 | 80 | 19 | 0,36 | 0,49 | 6 |
| <i>MUC16</i> | 63 | 85 | 80 | 22 | 0,68 | 0,45 | 19 |
| <i>CD74</i> | 75 | 81 | 79 | 6 | 0,82 | 0,42 | 6 |
| <i>TNFRSF14</i> | 66 | 82 | 79 | 16 | 0,9 | 0,61 | 8 |
| <i>SGK1</i> | 66 | 85 | 79 | 18 | 0,68 | 0,29 | 18 |
| <i>ARID1A</i> | 49 | 80 | 79 | 31 | 0,9 | 0,39 | 8 |
| <i>LTB</i> | 51 | 62 | 78 | 11 | 0,2 | 0,14 | 12 |
| <i>PA5</i> | 71 | 85 | 78 | 14 | 0,82 | 0,53 | 7 |
| <i>CD83</i> | 68 | 82 | 78 | 14 | 0,78 | 0,52 | 8 |
| <i>HIST1H1C</i> | 68 | 85 | 78 | 16 | 0,73 | 0,47 | 14 |
| <i>PIM1</i> | 62 | 85 | 78 | 22 | 0,54 | 0,49 | 28 |
| <i>HIST1H1E</i> | 62 | 85 | 78 | 22 | 0,54 | 0,49 | 28 |
| <i>HIST1H2AC</i> | 60 | 88 | 78 | 28 | 0,55 | 0,52 | 9 |
| <i>TMSB4</i> | 66 | 83 | 74 | 17 | 0,81 | 0,30 | 18 |
| <i>KMT2D</i> | 63 | 82 | 74 | 19 | 0,65 | 0,62 | 21 |
| <i>DUSP2</i> | 62 | 81 | 73 | 19 | 0,36 | 0,91 | 19 |
| <i>CARD11</i> | 59 | 81 | 72 | 22 | 0,73 | 0,54 | 14 |
| <i>USH2A</i> | 62 | 75 | 70 | 13 | 0,15 | 0,26 | 11 |
| <i>FAT4</i> | 63 | 82 | 70 | 19 | 0,8 | 0,89 | 15 |
| <i>MEF2B</i> | 63 | 85 | 70 | 21 | 0,65 | 0,85 | 7 |
| <i>BTG2</i> | 61 | 82 | 70 | 21 | 0,94 | 0,90 | 21 |
| <i>CREBBP</i> | 59 | 80 | 70 | 21 | 0,53 | 0,73 | 13 |
| <i>SOCS1</i> | 55 | 79 | 70 | 24 | 0,03 | 0,38 | 21 |
| <i>GNAI3</i> | 64 | 78 | 68 | 14 | 0,55 | 0,71 | 6 |
| <i>BCL2</i> | 62 | 76 | 68 | 14 | 0,23 | 0,7 | 9 |
| <i>FAS</i> | 61 | 82 | 68 | 21 | 0,55 | 0,97 | 7 |
| <i>LRP1B</i> | 50 | 82 | 68 | 32 | 0,8 | 0,21 | 14 |
| <i>TMEM30A</i> | 58 | 82 | 67 | 24 | 0,9 | 0,64 | 7 |
| <i>B2M</i> | 57 | 82 | 67 | 25 | 0,5 | 0,36 | 19 |
| <i>DT1</i> | 60 | 78 | 66 | 18 | 0,13 | 0,29 | 12 |
| <i>FAT1</i> | 52 | 82 | 66 | 30 | 0,85 | 0,46 | 16 |
| <i>DMD</i> | 62 | 70 | 64 | 8 | 0,04 | 0,57 | 6 |
| <i>HMCN1</i> | 62 | 75 | 63 | 13 | 0,25 | 0,17 | 6 |
| <i>HIST1H2BC</i> | 62 | 78 | 62 | 16 | 0,25 | 0,80 | 6 |
| <i>ZFP36LI</i> | 54 | 71 | 59 | 17 | 0,36 | 0,20 | 7 |

Abkürzungen und Verweise: Min. : Minimales Ersterkrankungsalter des Patienten in der analysierten Kohorte. Max. : Maimales Ersterkrankungsalter des Patienten in der analysierten Kohorte. Median Alter: Medianes Alter der analysierten DLBCL-Kohorte mit identifizierter Mutation. Interquartil-Abstand: Alterspektrum der Patienten mit identifizierter Mutation in der analysierten DLBCL-Kohorte (Differenz aus Max. Alter und M.in. Alter). U Test: Mann-Whitney U Test. LR: Logistische Regression. MF: Mutationsfrequenz, grau markiert: Nach der Betrachtung des Median und Interquartil-Abstands (82 bzw, 9 Jahre) wurden Mutationen im *CD58*-Gen vorzugsweise bei Patienten im sehr hohen Alter identifiziert (Neues Kandidatengen für DLBCL-Patienten im hohen Alter).

Tabelle A 6: Biologisches und epigenetisches Alter der analysierten DLBCL-Proben

| Probe | Biologisches Alter [Jahre] | Epigenetisches Alter [Jahre] | Delta-Alter [Jahre] |
|---------|-------------------------------|---------------------------------|------------------------|
| 4124188 | 73 | 36 | -37 |
| 4144131 | 75 | 48 | -27 |
| MPI-128 | 85 | 59 | -26 |
| MPI-164 | 88 | 70 | -18 |
| 4161781 | 36 | 21 | -15 |
| 4163639 | 75 | 64 | -11 |
| MPI-243 | 86 | 76 | -10 |
| MPI-105 | 60 | 52 | -8 |
| MPI-849 | 80 | 72 | -8 |
| MPI-802 | 94 | 86 | -8 |
| 4135278 | 43 | 36 | -7 |
| 4141476 | 79 | 73 | -6 |
| 4173863 | 62 | 57 | -5 |
| MPI-119 | 81 | 76 | -5 |
| MPI-191 | 82 | 78 | -4 |
| MPI-331 | 87 | 83 | -4 |
| MPI-857 | 84 | 81 | -3 |
| MPI-762 | 84 | 83 | -1 |
| MPI-851 | 84 | 83 | -1 |
| MPI-749 | 82 | 82 | 0 |
| MPI-053 | 66 | 70 | 4 |
| MPI-099 | 80 | 84 | 4 |
| 4179894 | 73 | 78 | 5 |
| MPI-210 | 89 | 95 | 6 |
| 4135099 | 49 | 56 | 7 |
| 4115001 | 70 | 77 | 7 |
| 4111326 | 74 | 82 | 8 |
| 4120157 | 46 | 55 | 9 |
| MPI-309 | 84 | 94 | 10 |
| MPI-188 | 93 | 103 | 10 |
| MPI-841 | 81 | 93 | 12 |
| MPI-411 | 87 | 99 | 12 |
| MPI-768 | 88 | 100 | 12 |
| MPI-845 | 85 | 98 | 13 |
| 4138527 | 85 | 98 | 13 |
| 4181460 | 70 | 84 | 14 |
| 4166151 | 74 | 89 | 15 |
| MPI-883 | 81 | 97 | 16 |
| 4145528 | 62 | 82 | 20 |
| MPI-820 | 90 | 110 | 20 |
| 4193638 | 55 | 76 | 21 |
| MPI-229 | 63 | 84 | 21 |
| 4131257 | 72 | 95 | 23 |
| 4176133 | 61 | 85 | 24 |
| MPI-756 | 81 | 105 | 24 |
| 4168738 | 50 | 75 | 25 |
| 4119279 | 62 | 87 | 25 |
| MPI-859 | 81 | 106 | 25 |
| 4134434 | 84 | 109 | 25 |
| 4151471 | 29 | 55 | 26 |
| 4108101 | 66 | 92 | 26 |
| MPI-127 | 69 | 95 | 26 |
| MPI-203 | 71 | 100 | 29 |
| 4157186 | 74 | 103 | 29 |
| MPI-078 | 8 | 40 | 32 |
| 4138464 | 47 | 80 | 33 |
| MPI-206 | 82 | 115 | 33 |

| Probe | Biologisches Alter [Jahre] | Epigenetisches Alter [Jahre] | Delta-Alter [Jahre] |
|---------|-------------------------------|---------------------------------|------------------------|
| MPI-233 | 8 | 42 | 34 |
| 4102009 | 64 | 101 | 37 |
| MPI-076 | 82 | 119 | 37 |
| MPI-117 | 11 | 49 | 38 |
| 4189035 | 46 | 85 | 39 |
| MPI-165 | 86 | 125 | 39 |
| 4184094 | 57 | 97 | 40 |
| MPI-142 | 60 | 100 | 40 |
| MPI-086 | 14 | 55 | 41 |
| 4171810 | 64 | 105 | 41 |
| 4147968 | 65 | 107 | 42 |
| 4116738 | 15 | 60 | 45 |
| 4188879 | 52 | 100 | 48 |
| MPI-315 | 49 | 101 | 52 |
| 4122063 | 47 | 100 | 53 |
| 4107559 | 72 | 125 | 53 |
| 4130051 | 75 | 128 | 53 |
| MPI-118 | 81 | 136 | 55 |
| MPI-112 | 3 | 60 | 57 |
| 4107137 | 59 | 116 | 57 |
| MPI-201 | 88 | 145 | 57 |
| 4104893 | 16 | 74 | 58 |
| 4130194 | 55 | 114 | 59 |
| MPI-698 | 68 | 128 | 60 |
| 4133263 | 40 | 104 | 64 |
| MPI-171 | 24 | 93 | 69 |
| 4183136 | 61 | 132 | 71 |
| MPI-047 | 80 | 151 | 71 |
| 4199714 | 53 | 131 | 78 |

Abkürzungen und Verweise: Epigenetisches Alter: Ermitteltes Alter aus den AVG-Beta-Werten von 353 altersassoziierten CpG-Loci des Infinium HumanMethylation450 Bead Chips nach dem Modell von Steve Horvath, Delta-Alter: Differenz aus dem biologischen und epigenetischen Alter,

10 Danksagung

An dieser Stelle möchte ich mich bei **allen Personen bedanken, ohne deren Hilfe, Zeit, Unterstützung und Zusammenarbeit diese Arbeit nicht möglich gewesen wäre.** Folgenden Personen möchte ich meinen **besonderen Dank** aussprechen:

Meiner 1.Gutachterin **Frau Prof. Dr. Manuela Dittmar** für die Unterstützung und Betreuung bei dieser Arbeit seitens der Mathematisch Naturwissenschaftlichen Fakultät der Christian-Albrecht-Universität zu Kiel.

Herrn Prof. Dr. Wolfram Klapper (Institut für Pathologie/ Sektion Hämatopathologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel) für das große Interesse, die anregenden Diskussionen über die erzielten Ergebnisse und die fachkundigen Hinweise beim Zusammenschreiben dieser Arbeit.

Frau Dr. Julia Richter für die Bereitstellung der Thematik und das Vertrauen bzw. die Möglichkeit meine Promotion im Rahmen der von Ihr geleiteten *Junior Excellence Research Group* durchzuführen. In diesem Zusammenhang möchte ich auch der Dr. Werner Jackstädt Stiftung für die Förderung meiner Arbeiten in der *Junior Excellence Research Group* danken. Insbesondere gilt mein Dank Frau Dr. Julia Richter für die sehr gute Betreuung über den gesamten Zeitraum meiner Promotion sowie die uneingeschränkte Unterstützung und kompetenten Anregungen bei der Erstellung meiner schriftlichen Ausführungen.

Der **gesamten Arbeitsgruppe der Sektion Hämatopathologie** für die angenehme Atmosphäre und die ständige Hilfsbereitschaft. Insbesondere geht mein Dank an **Frau Dr. Christiane Stuhlmann-Laeisz** und **Frau Charlotte Botz-von Drahten** für die Unterstützung bei der Erstellung der neuen Patienten-Kohorte sowie der Bereitstellung der immunhistochemischen Färbungen und der Einzeltumorschnitte bzw. TMAs.

Allen Mitarbeitern des Instituts für Humangenetik des Universitätsklinikum Schleswig-Holsteins, Campus Kiel für die Unterstützung und Hilfsbereitschaft über den gesamten Zeitraum meiner Arbeit. Insbesondere danke ich **Herrn Prof. Dr. Reiner Siebert** für die Möglichkeit diese Fragestellung am Institut für Humangenetik bearbeiten zu können und die Ergebnisse auf nationalen und internationalen Konferenzen präsentieren zu können. Weiterhin möchte ich mich an dieser Stelle bei **Frau Dr. Inga Nagel**, **Frau Dr. Cristina Lopez Gonzalez** und **Frau Reina Zühlke-Jenisch** für die Durchführung und Auswertung der FISH-Analysen sowie bei **Frau Lorena Valles** für die technische Ausübung der Methylierungsarrays bedanken.

Zudem geht mein Dank an **Frau Dr. Nadine Hornig, Frau Claudia Becher und Frau Dorit Schuster** für die freundliche und sehr gute Zusammenarbeit im Bereich der Methodenetablierung der Hochdurchsatz-Sequenzierung am Institut für Humangenetik.

Herrn Dr. Matthias Schlesner und **Herrn Dr. Bingding Huang** vom deutschen Krebsforschungszentrum (DKFZ) in Heidelberg für die bioinformatische Auswertung der erhobenen Hochdurchsatz-Daten.

Herrn Dr. Markus Kreuz vom Institut für medizinische Informatik, Statistik und Epidemiologie (IMISE) in Leipzig für die Unterstützung und Hilfsbereitschaft bei den statistischen Auswertungen im Rahmen der vorliegenden Arbeit.

Herrn Prof. Dr. Ralf Küppers, Frau Dr. Patricia Johansson und **Frau M.Sc. Martina Przekopowitz** vom Institut für Zellbiologie (Tumorforschung) des Universitätsklinikums Essen für die Bereitstellung der FACS-sortierten B-Zellpopulationen.

Allen beteiligten Arbeitsgruppen und Mitwirkenden der Studien NHL-B, RICOVER-60, Molekulare Mechanismen der malignen Lymphome (MMML) und ICGC MMML-Seq für die Bereitstellung der Materialien und/oder der bereits existierenden molekulargenetischen Daten der Patienten.

Zu guter Letzt gilt mein besonderer Dank **meiner Familie**, die mich über den kompletten Zeitraum meiner Promotion immer liebevoll und moralisch unterstützt hat. Besonders sind in diesem Zusammenhang **meine Eltern** zu nennen, ohne deren unermüdliche Mithilfe dieser akademische Werdegang gar nicht möglich gewesen wäre.

11 Lebenslauf

Persönliche Daten

| | |
|---------------|---------------|
| Name | Ulrike Paul |
| Geburtsdatum | 05.April 1988 |
| Geburtsort | Schwedt/Oder |
| Familienstand | Ledig |

Schulbildung

| | |
|---------|-------------------------------------------------------------|
| 07/2007 | Allgemeine Hochschulreife, Einstein Gymnasium Angermünde |
|---------|-------------------------------------------------------------|

Studium

| | |
|-----------------|-------------------------------------------------------------------------------------|
| 11/2012 | Master of Science |
| 09/2010-11/2012 | Masterstudium Biosystemtechnik/Bioinformatik an der Technischen Hochschule Wildau |
| 07/2010 | Bachelor of Science |
| 09/2007-07/2010 | Bachelorstudium Biosystemtechnik/Bioinformatik an der Technischen Hochschule Wildau |

Wissenschaftliche Arbeiten

| | |
|---------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 02/2016 bis heute | Wissenschaftliche Mitarbeiterin mit dem Ziel zur Promotion im Rahmen der Förderung einer <i>Junior Excellence Research Group</i> durch die Dr. Werner Jackstädt Stiftung (Leitung: Frau Dr. Julia Richter), Institut für Pathologie/ Sektion Hämatopathologie (Universitätsklinikum Schleswig-Holstein, Campus Kiel) |
| 03/2013 bis 02/2016 | Wissenschaftliche Mitarbeiterin mit dem Ziel zur Promotion im Rahmen der Förderung einer <i>Junior Excellence Research Group</i> durch die Dr. Werner Jackstädt Stiftung (Leitung: Frau Dr. Julia Richter), Institut für Humangenetik (Universitätsklinikum Schleswig-Holstein, Campus Kiel) |
| 04/2012- 11/2012 | Masterarbeit im Studiengang Biosystemtechnik/Bioinformatik Thema: „Analyse des T-Zell-Rezeptor-Repertoires mittels Hochdurchsatz-Sequenzierung“, Kooperation Technischen Hochschule Wildau (Labor für Molekulare Biotechnologie und funktionelle Genomik, Prof. Dr. Marcus Frohme) und Charité Berlin (Institut für Pathologie, Campus Benjamin Franklin, Prof. Dr. Michael Hummel) |
| 03/2010-05/2010 | Bachelorarbeit im Studiengang Biosystemtechnik/Bioinformatik Thema: „Bestimmung des genomweiten Acetylierungsprofils von Plasmazytomen“, Charité Berlin (Institut für Pathologie, Campus Benjamin Franklin, Prof. Dr. Michael Hummel) |

12 Publikationsliste

Orginalarbeiten, die Teil der vorliegenden Arbeit sind

López C, Bergmann AK, **Paul U**, Murga Penas EM, Nagel I, Betts MJ, Johansson P, Ritgen M, Baumann T, Aymerich M, Jayne S, Russell RB, Campo E, Dyer M JS, Dürig J, Siebert R, Genes encoding members of the JAK-STAT pathway or epigenetic regulators are recurrently mutated in T-PLL, *Br. J. Haematol.*, 2015, DOI: 10.1111/bjh.1395

Orginalarbeiten, die nicht Teil der vorliegenden Arbeit sind

Bens S, Kolarova J, Beygo J, Buiting K, Caliebe A, Eggermann T, Gillessen-Kaesbach G, Prawitt D, Thiele-Schmitz S, Begemann M, Enklaar T, Gutwein J, Haake A, **Paul U**, Richter J, Soellner L, Vater I, Monk D, Horsthemke B, Ammerpohl O, Siebert R, Phenotypic spectrum and extent of DNA methylation defects associated with multi-locus imprinting disturbances, *Epigenomics*, accepted und wird voraussichtlich in der Juni-Ausgabe 2016 veröffentlicht

Veröffentlichte Tagungsbeiträge und Präsentationen

*geteilte Erstautorenschaft

Richter J*, **Paul U***, Huang B, Stuhlmann-Laiesz C, Nagel I, Lopez C, Kreuz M, Horn H, Hummel M, Ott G, Spang R, Küppers R, Rosenwald A, Pfreundschuh M, Schmitz N, Löffler M, Trümper L, Schlesner M, Klapper W, Siebert R, Molecular characterization of diffuse large B-cell lymphoma in senile patients, Symposium Hematological Malignancies – From Mechanisms to Therapy, 09.-12.03.2016, Milan, Italien

Paul U, Huang B, López C, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Molecular characterization of diffuse large B-cell lymphoma in senile patients, EMBL Conference: Cancer Genomics 2015, 01. – 04.10.2015, Heidelberg, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Molecular characterization of diffuse large B-cell lymphoma in senile patients, Annual Meeting of the AGD 2015, 23. – 24.10.2015, Potsdam, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Genetische Charakterisierung von Diffus großzelligen B-Zell-Lymphomen bei senilen Patienten, Treffen Norddeutscher Humangenetiker, 12.09.2015, Hamburg, Deutschland

Murga Penas EM, **Paul U**, Pomplum C, Nagel I, Becher C, Gramatzki M, Siebert R, Günther A, Siebert R, Mutations in the *TP53* gene are recurrent in secondary myelodysplastic syndromes in patients with refractory multiple myeloma, Jahrestagung der Deutschen Gesellschaft für Hämatologie und Medizinische Onkologie, 9. – 13.10.2015, Basel, Schweiz

López C, Bergmann AK, **Paul U**, Baumann T, Johansson P, Ritgen M, Salaverria I, Aymerich M, Dürig J, Dührsen U, Dyer M, Küppers R, Campo E, Siebert R, Characterization of the molecular pathogenesis of T-prolymphocytic leukemia (T-PLL): new insights into the biology using genomic and epigenomic profiling, Annual Conference of the German Genetics Society, 28. – 30.9.2015, Kiel, Deutschland

Murga Penas EM, **Paul U**, Pomplum C, Nagel I, Becher C, Gramatzki M, Siebert R, Günther A, Siebert R, Mutations in the *TP53* gene are recurrent in secondary myelodysplastic syndromes in patients with refractory multiple myeloma, Annual Conference of the German Genetics Society, 28. – 30.9.2015, Kiel, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Genetic characterization of diffuse large B-cell lymphoma in senile patients, Annual Conference of the German Genetics Society, 28. – 30.9.2015, Kiel, Deutschland

Paul U, Pomplum C, Nagel I, Becher C, Gramatzki M, Siebert R, Günther A, Murga Penas EM, Molecular genetic and cytogenetic characteristics of secondary myelodysplastic syndromes in refractory multiple myeloma patients, 1. Tumorgenetische / 28. Tumorzytogenetische Arbeitstagung, 28. – 30.5.2015, Göttingen, Deutschland

13 Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation eigenständig und nur mit den von mir angegebenen Quellen und Hilfsmitteln bzw. mit der wissenschaftlichen Beratung meiner akademischen Betreuer angefertigt habe. Die vorliegende Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden und wurde bisher in gleicher oder ähnlicher Form an keiner anderen Universität zur Begutachtung im Rahmen eines Prüfungsverfahrens vorgelegt. Weiterhin bestätige ich, dass ich zuvor noch keine Promotionsversuche unternommen habe.

Folgende Teile, die im Rahmen der Methoden-Optimierung der Hochdurchsatz-Sequenzierung der vorliegenden Arbeit erzielt wurden, sind bereits in Form eines Artikels in einer internationalen Fachzeitschrift oder als Tagungsbeitrag veröffentlicht worden:

López C, Bergmann AK, **Paul U**, Murga Penas EM, Nagel I, Betts MJ, Johansson P, Ritgen M, Baumann T, Aymerich M, Jayne S, Russell RB, Campo E, Dyer M JS, Dürig J, Siebert R, Genes encoding members of the JAK-STAT pathway or epigenetic regulators are recurrently mutated in T-PLL, *Br. J. Haematol.*, 2015, DOI: 10.1111/bjh.1395

López C, Bergmann AK, **Paul U**, Baumann T, Johansson P, Ritgen M, Salaverria I, Aymerich M, Dürig J, Dührsen U, Dyer M, Küppers R, Campo E, Siebert R, Characterization of the molecular pathogenesis of T-prolymphocytic leukemia (T-PLL): new insights into the biology using genomic and epigenomic profiling, Annual Conference of the German Genetics Society, 28. – 30.9.2015, Kiel, Deutschland

Im Rahmen der genetischen Analysen der vorliegenden Arbeit wurde die praktische Laborarbeit unter Verwendung des Lymphom-Panels selbstständig von mir am Institut für Humangenetik in Kiel durchgeführt. Die Auswertung der Hochdurchsatz-Sequenzierungs-Daten des Lymphom-Panels und der Vergleich mit den bereits existierenden Sequenz-Daten aus dem ICGC MMML-Seq Projekt erfolgte am Deutschenkrebsforschungszentrum in Heidelberg in der Arbeitsgruppe zur Theoretischen Bioinformatik durch Herrn Dr. Matthias Schlesner und Herrn Dr. Bingding Huang. Die mir im Juni 2015 übermittelten ausgewerteten Hochdurchsatz-Sequenzierungs-Daten dienten als Basis für die Ergebnisse der Mutationsanalysen dieser Arbeit und wurden in Form von Tagungsbeiträgen bei folgenden internationalen Konferenzen veröffentlicht:

Richter J*, **Paul U***, Huang B, Stuhlmann-Laiasz C, Nagel I, Lopez C, Kreuz M, Horn H, Hummel M, Ott G, Spang R, Küppers R, Rosenwald A, Pfreundschuh M, Schmitz N, Löffler M, Trümper L, Schlesner M, Klapper W, Siebert R, Molecular characterization of diffuse large B-cell lymphoma in senile patients, Symposium Hematological Malignancies – From Mechanisms to Therapy, 09.-12.03.2016, Milan, Italien *geteilte Erstautorenschaft

Paul U, Huang B, López C, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Molecular characterization of diffuse large B-cell lymphoma in senile patients, EMBL Conference: Cancer Genomics 2015, 01. – 04.10.2015, Heidelberg, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Molecular characterization of diffuse large B-cell lymphoma in senile patients, Annual Meeting of the AGD 2015, 23. – 24.10.2015, Potsdam, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Genetische Charakterisierung von Diffus großzelligen B-Zell-Lymphomen bei senilen Patienten, Treffen Norddeutscher Humangenetiker, 12.09.2015, Hamburg, Deutschland

Paul U, Huang B, Stuhlmann-Laiesz C, Nagel I, Schlesner M, Klapper W, Siebert R, Richter, Genetic characterization of diffuse large B-cell lymphoma in senile patients, Annual Conference of the German Genetics Society, 28. – 30.9.2015, Kiel, Deutschland

Kiel, den 26.04.2016

Ulrike Paul