

Aus dem Department für Diagnostische Labormedizin der
Universität Tübingen

Institut für Pathologie und Neuropathologie
Abteilung Allgemeine und Molekulare Pathologie und
Pathologische Anatomie

**Der Einsatz des „Next Generation Sequencing“ zum
Klonalitätsnachweis bei malignen Lymphoproliferationen**

**Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin**

**der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen**

vorgelegt von

Raisch, Philipp Georg

2021

Dekan: Professor Dr. B. Pichler
1. Berichterstatter: Professor Dr. F. Fend
2. Berichterstatter: Privatdozent Dr. F. Eßmann

Tag der Disputation: 26.07.2021

INHALTSVERZEICHNIS

1. Zusammenfassung	1
2. Einleitung.....	3
2.1 Maligne Lymphome: Epidemiologische und klinische Einordnung.....	3
2.1.1 Epidemiologie und Risikofaktoren	3
2.1.2 Klassifikation und Diagnostik	4
2.2 B-Zell-Reifung und B-Zell-Rezeptor-Repertoire.....	5
2.2.1 Bedeutung und Struktur von Immunglobulinmolekülen und -Genen	5
2.2.2 VDJ-Rekombination und somatische Hypermutation.....	7
2.2.3 B-Zell-Reifung im Kontext der malignen Entartung	9
2.3 Monoklonale BCR-Rearrangements als diagnostischer Marker maligner Lymphome	15
2.3.1 Monoklonalität in malignen Lymphomen.....	15
2.3.2 Prinzipien der molekularen Klonalitätsdiagnostik.....	15
2.3.3 Probleme der PCR-basierten Klonalitätsdiagnostik	18
2.4 Klonalitätsnachweis mittels „Next Generation Sequencing“	20
2.4.1 Grundlagen und -überlegungen	20
2.4.2 Zielsetzung der Arbeit	21
3. Material und Methoden	24
3.1 Liste der verwendeten Reagenzien und Geräte	24
3.1.1 Chemikalien und Reagenzien	24
3.1.2 Enzyme	24
3.1.3 Kits.....	24
3.1.4 Zelllinien	25
3.1.5 Primer (Sigma-Aldrich)	25
3.1.6 Geräte.....	26
3.1.7 Verbrauchsmaterialien.....	27
3.1.8 Programme und Tools	27
3.1.9 Hersteller	27
3.2 Extraktion und Aufreinigung von DNA	28
3.2.1 Extraktion und Aufreinigung von DNA aus FFPE-Gewebe	28
3.2.2 Extraktion und Aufreinigung von DNA aus Zellen einer B-NHL-Zelllinie	29
3.2.3 PCR zur Bestimmung der Amplifizierbarkeit der DNA	30
3.2.4 Herstellung von Verdünnungsproben von B-NHL-DNA in PBMC-DNA	32

3.3	Klonalitätsuntersuchung mittels Fragmentanalyse	32
3.3.1	Multiplex-PCR mit den Primern des BIOMED-2 Konsortiums.....	32
3.3.2	Klonalitätsanalyse am GenomeLab GeXP Analyzer.....	33
3.3.3	Kriterien zur Auswertung	34
3.4	Sequenzierung klonaler IGH-Rearrangements und Charakterisierung mittels IMGT/V-Quest	35
3.4.1	Ermittlung der VH-Segmentfamilie und Amplifikation klonaler IGH- Rearrangements.....	35
3.4.2	Überprüfung der Amplifikation mittels Agarose-Gelelektrophorese	35
3.4.3	DNA-Isolation aus Agarosegel zur Isolation der klonalen Bande	36
3.4.4	Sequenzierung von IGH-Rearrangements nach Sanger	37
3.4.5	Analyse des klonalen IGH-Rearrangements mittels IMGT/V-Quest	40
3.5	Next Generation Sequencing des amplifizierten IGH-Repertoires	41
3.5.1	Multiplex-PCR mit VH-Primern mit Fluidigm-Anhang.....	41
3.5.2	Barcode-PCR und Proben-Pooling.....	43
3.5.3	Sequenzierung auf der Illumina MiSeq-Plattform.....	44
3.5.4	Analyse des amplifizierten IGH-Repertoires mittels IMGT/HighV-Quest	45
4.	Ergebnisse	50
4.1	Beschreibung des Probenkollektivs	50
4.1.1	B-NHL-Zelllinien.....	50
4.1.2	Zusammenstellung der verwendeten Patientenproben	50
4.2	Amplifizierbarkeit der DNA	53
4.2.1	Amplifizierbarkeit der DNA aus Formalin-fixiertem Paraffin-eingebettetem Gewebe	53
4.2.2	Amplifizierbarkeit der DNA aus Glaskörperaspiraten	54
4.2.3	Amplifizierbarkeit von B-NHL-DNA in Multiplex-PCR-Ansätzen mit VH-Primern mit Fluidigm-Anhang.....	54
4.3	Parameter der Klonalitätsanalyse in NGS-Datensätzen von Verdünnungsproben von B-NHL-Zelllinien-DNA	56
4.3.1	Ermittelte IGH-Rearrangements der Lymphom-Zelllinien	56
4.3.2	Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen	58
4.3.3	Anteil des IGH-Rearrangements der Lymphom-Zelllinie an der Gesamtzahl der IGH-Sequenzen.....	63
4.3.4	IGH-Rearrangements der lymphozytären Hintergrundpopulation.....	79

4.4	NGS- und GeneScan-basierte Klonalitätsanalyse im Vergleich bei Patientenproben mit hohem Lymphomzellgehalt	79
4.4.1	Sequenzierbarkeit nach Sanger und ermittelte IGH-Rearrangements	80
4.4.2	Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen	81
4.4.3	Anteile sequenzstarker IGH-Rearrangements an der Gesamtzahl der IGH-Sequenzen	83
4.4.4	Vergleich der NGS-Daten mit Befunden der GeneScan-Analyse	93
4.5	NGS-basierte Klonalitätsuntersuchung bei Verlaufspuben	95
4.5.1	Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen	95
4.5.2	Anteile der sequenzstärksten IGH-Rearrangements an der Gesamtzahl der IGH-Sequenzen und Detektierbarkeit im zeitlichen Verlauf	97
4.5.3	Varianten klonaler IGH-Rearrangements punktuell und im zeitlichen Verlauf	104
4.6	NGS-basierte Klonalitätsanalyse von Glaskörperaspiraten mit geringem Lymphomzellgehalt.....	115
4.6.1	Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen	115
4.6.2	Beurteilung des nachweisbaren IGH-Repertoires und Vergleich mit Ergebnissen der GeneScan-Analyse	116
5.	Diskussion	118
5.1	Anhaltspunkte für Klonalität in NGS-Datensätzen und Vergleich mit der GeneScan-Analyse.....	118
5.1.1	Charakteristika klonaler IGH-Rearrangements in NGS-Datensätzen	118
5.1.2	Charakteristika der amplifizierten Hintergrundpopulation	121
5.1.3	Vergleich der Klonalitätstestungen mittels NGS und GeneScan	122
5.1.4	Gründe für falsch negative GeneScan- und NGS-Befunde	124
5.2	Mutationsstatus und Varianten klonaler IGH-Rearrangements und Abgrenzbarkeit zu Artefakten.....	127
5.2.1	Varianten klonaler Rearrangements als Problem in der Klonalitätsdiagnostik	127
5.2.2	Bedeutung von Mutationen und Varianten in mittels NGS detektierten Lymphompopulationen.....	131
5.2.3	Methodische Schwächen in der Beurteilung von Varianten und klonaler Evolution	141
5.3	Möglichkeiten und Grenzen der NGS-Diagnostik zum Nachweis kleiner Zahlen klonaler Zellen	144

INHALTSVERZEICHNIS

5.3.1	Nachweisgrenze klonaler Populationen mittels NGS und Eignung zum MRD-Nachweis	144
5.3.2	Aussagekraft von NGS- und GeneScan-basierter Klonalitätsuntersuchung bei Proben mit geringem Lymphomzellgehalt.....	147
5.4	Schlussfolgerungen und Ausblick	149
6.	Verzeichnisse	152
6.1	Abkürzungsverzeichnis.....	152
6.2	Literaturverzeichnis	155
6.3	Abbildungsverzeichnis	167
6.4	Tabellenverzeichnis.....	169
7.	Erklärung zum Eigenanteil der Dissertationsschrift	171
8.	Danksagung	172

1. Zusammenfassung

Der Nachweis klonaler Immunglobulin-Umlagerungen (IGH-Rearrangements) mittels PCR und anschließender Fragmentanalyse ist ein Standardverfahren der molekularpathologischen Lymphomdiagnostik. Die etablierten Techniken sind jedoch für kleine extranodale Biopsien, sowie Liquor- und Glaskörperaspirate oft problematisch, da aufgrund des geringen Lymphozytengehalts und somit in der DNA-Probe nachweisbaren eingeschränkten Repertoires an IGH-Rearrangements hier oft pseudoklonale Amplifikate auftreten, die zu falsch positiven Befunden führen können. Ziel der vorliegenden Arbeit war die Etablierung eines Verfahrens, bei dem mittels PCR amplifizierte IGH-Rearrangements mittels „Next Generation Sequencing“ (NGS) zum B-Zell-Klonalitätsnachweis analysiert werden sollten. Dieses sollte insbesondere für Proben mit erhöhter Frequenz falsch positiver oder nicht interpretierbarer Ergebnisse bei Standardtechnik angewandt und damit eine Verbesserung der molekularen Lymphomdiagnostik und der Bestimmung von Minimal Residual Disease (MRD) erreicht werden.

Das hier erprobte Verfahren basierte auf einem zweistufigen PCR-Ansatz, in dem zunächst die Amplifikation des IGH-Repertoires einer Lymphomprobe mittels BIOMED-2 Framework Region 1 und 2 Primersets erfolgte und in einem zweiten Schritt Fluidigm Barcode Primer für die massiv parallele Sequenzierung mit der Illumina MiSeq-Plattform eingesetzt wurden. Die NGS-Daten wurden mit dem Online-Tool IMGT/HighV-Quest untersucht, das die Detektion und Analyse von IGH-Rearrangements in NGS-Datensätzen, sowie deren statistische Auswertung ermöglicht.

Zunächst wurden anhand von Verdünnungsreihen von Lymphomzelllinien-DNA in DNA einer polyklonalen Lymphozytenpopulation die optimalen PCR-Bedingungen für BIOMED-2 Primersets und die Nachweisgrenze des NGS etabliert. Anschließend wurden verschiedene Patientenkollektive untersucht, um die Tauglichkeit der Methode anhand von 15 Proben mit hohem Lymphomzellgehalt sowie an 18 Verlaufspalten, jeweils drei, im Verlauf der

Erkrankung von sechs Patienten gewonnen, zu testen. Schließlich wurden drei Glaskörperaspirate bei intraokulärer Lymphommanifestation untersucht, um Vorteile der Anwendung von NGS bei Proben mit geringer Zellzahl zu ermitteln.

Anhand der Verdünnungsreihen ließ sich mittels NGS ein überproportionaler Anteil des IGH-Rearrangements der Lymphomzelllinie am Gesamtrepertoire der Proben darstellen, der mit sinkender Ausgangskonzentration der Lymphom-DNA abnahm. Auch in Verdünnungen bis 1:1.000 konnte das IGH-Rearrangement der Lymphomzelllinie durch NGS detektiert werden. In Patientenproben ließen sich primär mittels Sanger-Sequenzierung bestimmte IGH-Rearrangements eines malignen Klons durch NGS detektieren. Auch in einigen Fällen, in denen keine Sanger-Sequenzierung gelungen war, konnten durch NGS überrepräsentierte IGH-Rearrangements erkannt und als klonal interpretiert werden. In Verlaufsproben gelang mittels NGS die Detektion und Nachverfolgung einzelner Lymphozytenpopulationen anhand ihres jeweiligen IGH-Rearrangements. In schwer interpretierbaren Glaskörperaspiraten konnten mittels NGS zusätzliche Informationen über das zugrundeliegende IGH-Repertoire der Probe gewonnen werden. Als limitierende Faktoren der Sensitivität des in der vorliegenden Arbeit praktizierten Vorgehens zeigten sich insbesondere vorbeschriebene Schwächen der BIOMED-2 Assays an Proben aus Formalin-fixiertem Paraffin-eingebettetem Gewebe sowie die Schwierigkeit, Sequenzvarianten eindeutig als biologisch bedingt oder als Artefakt der Fixierung, des Amplifikations- oder Sequenzierprozesses einzuordnen.

Die Analyse amplifizierter IGH-Rearrangements mittels NGS liefert folglich eine Aussage über das Vorliegen von Klonalität. Darüber hinaus ermöglicht die Methode die spezifische Detektion und Nachverfolgung einzelner Lymphozytenpopulationen, beispielsweise zum Zweck der MRD-Untersuchung. Durch eine detaillierte Charakterisierung des IGH-Repertoires von Proben mit geringem Zellgehalt kann die Methode zur Interpretation problematischer Proben beitragen.

2. Einleitung

2.1 Maligne Lymphome: Epidemiologische und klinische Einordnung

2.1.1 Epidemiologie und Risikofaktoren

Lymphome sind von den Lymphozyten des Immunsystems ausgehende Neoplasien. Damit umfasst der Begriff eine Vielzahl maligner lymphoproliferativer Erkrankungen mit vielfältiger klinischer und histologischer Präsentation [1]. Vom Hodgkin Lymphom, das etwa 10 % aller Lymphome ausmacht, werden die sog. Non-Hodgkin-Lymphome (NHL) abgegrenzt, die wiederum zu etwa 85 – 90 % von B-Lymphozyten ausgehen (B-NHL) [2]. In Deutschland erkrankten im Jahr 2012 16.150 Menschen neu und 6.362 Menschen starben an einem NHL, wobei jeweils der Anteil der Männer etwas größer war. Die relative 5-Jahres-Überlebensrate betrug zu diesem Zeitpunkt etwa 60 – 70 %. Zwar betrug das mittlere Erkrankungsalter ca. 70 Jahre. Dennoch existieren beispielsweise Typen von NHL, die hauptsächlich Kinder betreffen. Die Inzidenz maligner Lymphome ist in den letzten Jahrzehnten in vielen Regionen der Welt stetig steigend [3-6].

In den meisten Fällen von NHL lässt sich keine konkrete Ursache eruieren [1]. Zu den Risikofaktoren zählt an erster Stelle die Immunschwäche, beispielsweise angeboren oder bei HIV-Infektion, aber auch medikamentös induziert nach Organ- oder Stammzelltransplantation. Des Weiteren sind Assoziationen bestimmter NHL mit spezifischen Virusinfektionen gesichert, und auch im Allgemeinen können chronisch entzündliche Prozesse zur Entstehung von Lymphomen beitragen [1-3, 7]. Die Pathomechanismen der Lymphomentstehung stehen in engem Zusammenhang zu den physiologischen Prozessen der Lymphozytenentwicklung und -reifung in den lymphatischen Organen, die mit extensiver Proliferation und riskanten Modifikationsprozessen der Keimbahn-DNA einhergehen [8, 9].

2.1.2 Klassifikation und Diagnostik

Als weiterer Ausdruck der Heterogenität der NHL, können diese sich prinzipiell in jedem Gewebe manifestieren und umfassen bezüglich des Krankheitsverlaufes ein breites Spektrum, von indolenten Entitäten wie dem folliculären Lymphom (FL) bis zu hoch aggressiven Formen wie dem Burkitt-Lymphom oder der Gruppe der diffus großzelligen B-Zell-Lymphome („diffuse large B-cell lymphoma“, DLBCL). Ein aktuelles und weit verbreitetes Klassifikationssystem ist die vierte Edition der „WHO-Klassifikation von Tumoren hämatopoetischer und lymphoider Gewebe“ aus dem Jahr 2008, überarbeitet im Jahr 2016 [10, 11]. Dieses Klassifikationssystem, das seinerseits auf der Version von 2001 basiert, teilt Lymphome anhand morphologischer, immunhistochemischer, molekularer, genetischer und klinischer Eigenschaften und diagnostischer Befunde in eine Vielzahl eigenständiger Erkrankungen ein [12]. Das verbreitetste System zur klinischen Stadienbestimmung von Lymphomen ist die Ann-Arbor-Klassifikation, die ursprünglich für das Hodgkin Lymphom entwickelt wurde [13, 14]. Sie inkorporiert das Stadium nach Ausbreitungsgrad, das Vorhandensein der Allgemeinsymptome Gewichtsverlust, Nachtschweiß und Fieber, sowie den Befall extralymphatischer Organe [15, 16]. Eine zuverlässige Kategorisierung von Entität und Krankheitsstadium ist die Basis für das richtige therapeutische Vorgehen. Zu den Therapiemodalitäten, die bei NHL Anwendung finden, zählen, abhängig von der Entität, neben Chemo- und Radiotherapie der Einsatz spezifischer monoklonaler Antikörper, sowie autologe und allogene Stammzelltransplantation [1, 2].

Grundprinzipien sowohl der Klassifikation maligner Lymphome, als auch ihrer Diagnostik, leiten sich aus der Entwicklung der Lymphozyten ab, in deren Zentrum die Generierung und Optimierung eines individuellen, antigen-bindenden Rezeptors steht [17]. Der Klonalitätsnachweis, ein molekularpathologisches Prinzip zur Diagnostik maligner Lymphome, basiert auf der Analyse des Repertoires antigen-bindender Rezeptoren innerhalb einer suspekten Lymphoproliferation [18].

2.2 B-Zell-Reifung und B-Zell-Rezeptor-Repertoire

2.2.1 Bedeutung und Struktur von Immunglobulinmolekülen und - Genen

B-Lymphozyten bilden, gemeinsam mit den T-Lymphozyten und natürlichen Killerzellen, eine Untergruppe der Leukozyten und besetzen Schlüsselpositionen des Immunsystems, insbesondere der adaptiven humoralen Immunantwort. Jeder reife B-Lymphozyt besitzt ein membrangebundenes Immunglobulin (Ig/Antikörper), das in Zusammenarbeit mit Molekülen der Signaltransduktion als individueller Antigenrezeptor mit eigener Spezifität fungiert und als B-Zell-Rezeptor (BCR) bezeichnet wird [7]. Der Vielzahl und Vielfalt der im B-Zell-Repertoire eines Menschen vertretenen BCRs und Immunglobuline kommt eine entscheidende Bedeutung hinsichtlich der Fähigkeit des erworbenen Immunsystems zu, unzählige verschiedene Krankheitserreger durch Antigenkontakt zu erkennen und im Rahmen der spezifischen Immunantwort zu bekämpfen [19].

Die Vielfalt der Antikörper ist das Resultat eines komplexen Rearrangement-Prozesses von Gensegmenten der Keimbahn-DNA und weiterer Modifikationsmechanismen, die während der B-Zell-Reifung durchlaufen werden [20]. Die Gene liegen auf Keimbahnebene zunächst nicht-funktional, gewissermaßen in Form einzelner „Bausteine“ vor und werden somatisch während der Entwicklung des einzelnen Lymphozyten zu Exons zusammengesetzt, die für den BCR bzw. das Immunglobulin kodieren [17, 20, 21].

Jedes Immunglobulin besteht aus vier Polypeptiden, zwei identischen miteinander kovalent verbundenen sogenannten schweren Ketten („heavy chain“, IgH) und zwei identischen leichten Ketten („light chain“, IgL), die jeweils mit einer schweren Kette kovalent verbunden sind [22, 23]. Es existieren zwei verschiedene Isotypen der leichten Kette, die als κ - und λ -Isotyp bezeichnet werden [17]. Die Ketten bestehen wiederum aus einer variablen Region („variable region“, V) die für die Antigenerkennung und -spezifität verantwortlich

ist und einer konstanten Region („constant region“, C), die strukturelle, Signal- und Effektorfunktionen wahrnimmt sowie im Bereich der schweren Kette, die Ig-Klasse bestimmt [17, 24, 25]. Die konstante Region wird von mehreren Exons kodiert, die durch konventionelles RNA-Spleißen zusammengesetzt werden. Eine C-Region kann durch eine andere ausgetauscht werden in einem DNA-Rearrangement-Prozess, der als Klassenwechsel („class switch recombination“, CSR) des Immunglobulins bezeichnet wird [26, 27].

Die V-Region hingegen wird kodiert von mehreren, in der Keimbahn-DNA nicht zusammenhängenden Gensegmenten, die in V- („variable“), D- („diversity“) und J- („joining“) Segmente eingeteilt werden [17, 25]. Um eine funktionelle schwere Kette zu generieren, müssen je ein V-, ein D- und ein J-Segment der Keimbahn-DNA in einem streng regulierten Rearrangement-Prozess zusammengefügt werden, der sogenannten VDJ-Rekombination. Es resultiert ein IGH-Rearrangement, ein neu zusammengesetzter DNA-Abschnitt, der für ein IGH-Molekül kodiert [21, 28]. Das aus den verschiedenen Kombinationen von Gensegmenten resultierende rekombinatorische Repertoire wird weiter diversifiziert durch den Prozess der somatischen Hypermutation, im Rahmen dessen Punktmutationen im Exon der V-Region generiert werden [29, 30].

Innerhalb der V-Region eines individuellen IGH-Rearrangements lassen sich Regionen besonders hoher Variabilität abgrenzen, die auf Polypeptidebene den Antigenkontakt vermitteln, sogenannte „complementary determining regions“ (CDR1, CDR2 und CDR3). Sie werden flankiert von weniger variablen Regionen, die die Struktur der V-Region vermitteln, und „framework regions“ (FR1, FR2 und FR3) genannt werden [17, 25] (Abbildung 1). Die CDR1 und 2 sind im jeweiligen V-Gensegment kodiert. Die CDR3 erstreckt sich von der Verbindung („junction“) zwischen V- und D-Segment, über das gesamte D-Segment bis zur Verbindung von D- und J-Segment. Ihre Sequenz ist damit abhängig von den Mechanismen der VDJ-Rekombination und ist daher, als Region mit der höchsten Variabilität, die wichtigste Region für die Antigenaffinität und Diversität des Immunglobulins [31, 32]. Untersuchungen mittels „Next Generation Sequencing“ (NGS) konnten zeigen, dass allein im Blut eines erwachsenen Menschen ca. drei bis neun Millionen verschiedener IGH-CDR3s nachweisbar sind [33]. Das gesamte

Repertoire eines erwachsenen Menschen wird auf mehrere Milliarden verschiedener Immunglobuline geschätzt [34].



Abbildung 1: Schematische Darstellung eines IGH-Rearrangements mit „framework regions“ (FRs) und „complementary determining regions“ (CDRs, schraffiert) der V-Region. Jedes Rearrangement besteht aus je einem V- („variable“, in blau), D- („diversity“, in Gelb) und J- („joining“, in Rot) Segment. Die drei wenig variablen FRs vermitteln auf Polypeptidebene die räumliche Struktur der V-Region, die hochvariablen CDRs den Antigenkontakt. Während die CDR1 und 2 Bestandteil des jeweiligen V-Segments sind, umfasst die CDR3 Abschnitte von V- und J- sowie das gesamte D-Segment inklusive der Verbindungsregionen der Segmente (mit N bezeichnet). Sie ist die Region der höchsten Variabilität des IGH-Rearrangements.

2.2.2 VDJ-Rekombination und somatische Hypermutation

Für den IGH-Locus des Menschen auf Chromosom 14 werden etwa 50 verschiedene V-Segmente, 27 D-Segmente und sechs J-Segmente katalogisiert [35, 36] (Abbildung 2). Diese V, D- und J-Segmente liegen beim Menschen in der sog. Translokon-Konfiguration jeweils in Gruppen upstream neun verschiedener C-Gene für die konstante Region [29]. Die Rekombination-aktivierenden Gene (RAG) RAG1 und RAG2 kodieren gemeinsam für die RAG-Endonuklease, die eine Schlüsselposition im Rekombinationsprozess einnimmt [37, 38]. Sie erzeugt DNA-Doppelstrangbrüche an den Rändern zwischen zwei kodierenden Gensegmenten und jeweils flankierenden speziellen Signalsequenzen und bringt zwei Gensegmente in räumliche Nähe, die anschließend in Zusammenarbeit mit DNA-Reparaturproteinen miteinander verbunden werden [25, 28]. Diese flankierenden Signalsequenzen, sogenannte „recombination signal sequences“ (RSS), setzen sich zusammen aus einem konservierten Heptamer- und einem Nonamer-Element, separiert durch einen relativ variablen „Spacer“ von entweder zwölf oder 23 bp Länge. Jedes V- und jedes J-Gensegment wird von einer 23-RSS flankiert, jedes D-Gensegment von zwei 12-RSS, je einer an jedem Ende [25, 39]. Um die Doppelstrangbruchformation zu erzeugen, muss RAG eine RSS

mit einem zwölf und eine mit einem 23 bp langen Spacer erkennen, eine Einschränkung, die auch als 12/23-Regel bezeichnet wird [40]. Dies führt dazu, dass in einem IGH-Rearrangement regelhaft ein D-Gensegment zwischen einem V- und einem J-Segment zu liegen kommt. Prinzipiell kann jedes beliebige D-Segment mit jedem beliebigen J-Segment und jedes resultierende DJ-Segment mit jedem V-Segment verbunden werden, was zu einer großen Zahl möglicher verschiedener VDJ-Rearrangements und damit verschiedener Polypeptide mit eigenen Antigenbindungseigenschaften führt. Dies wird als kombinatorische Diversität bezeichnet [28].

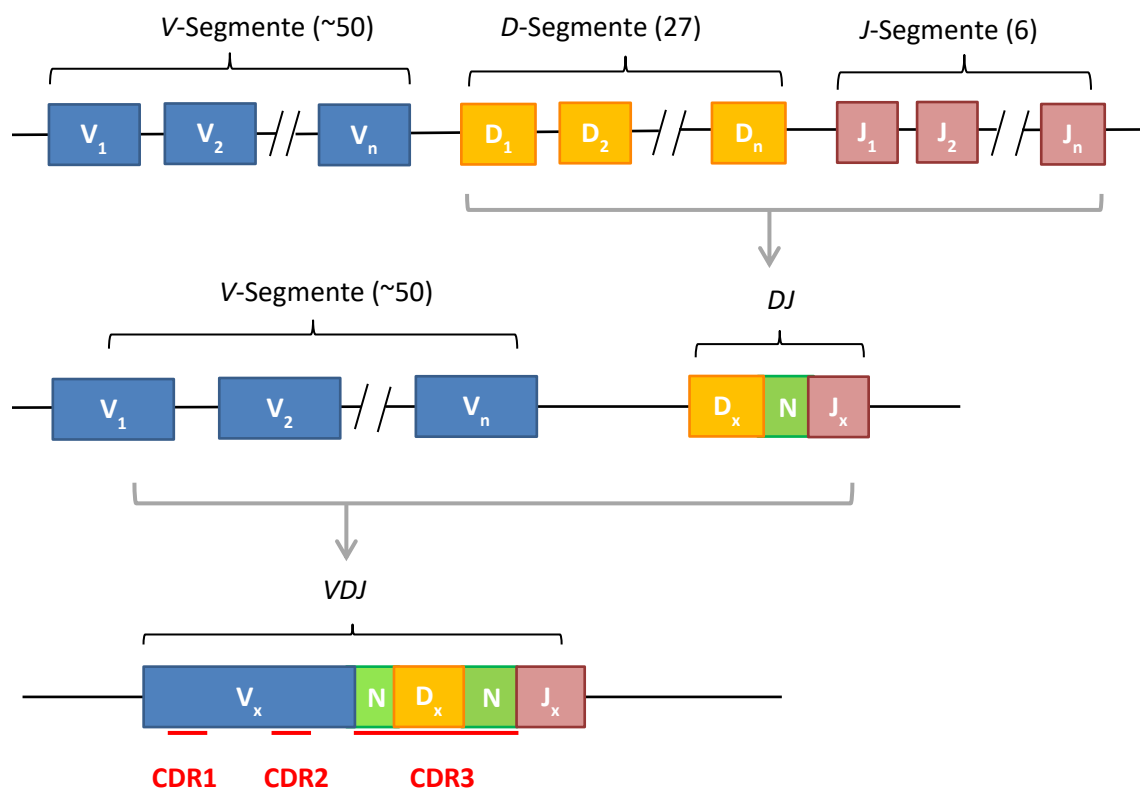


Abbildung 2: VDJ-Rekombination des IGH-Locus. Während der Entwicklung zu reifen B-Zellen im Knochenmark, wird zunächst ein D-Segment mit einem J-Segment zu einem DJ-Rearrangement verbunden. Anschließend erfolgt die Verbindung mit einem V-Segment zu einem VDJ-Rearrangement. In den Verbindungsregionen an beiden Enden des D-Segments können N-Insertionen und -Deletionen erfolgen. Die drei CDRs sind bevorzugte Ziele der SHM in der späteren Keimzentrumsreaktion. Die CDR3 erstreckt sich über die Verbindungsregionen und beinhaltet das gesamte D-Segment. Abbildung erstellt in Anlehnung an Kiyoi et al. [41]

Diese wird ergänzt durch die sog. junktionale, also die Verbindungsregionen der kombinierten Gensegmente betreffende Diversität („junctional diversity“). Sie resultiert aus der Tatsache, dass, wenn zwei Gensegmente verbunden werden, an ihren Enden häufig Nukleotide entfernt und neue hinzugefügt werden. An diesem Prozess ist unter anderem die Terminale Deoxynukleotidyl-Transferase (TdT) beteiligt, eine Template-unabhängige Polymerase, deren einzige bekannte Funktion es ist, „non-templated“ (N) Nukleotide, meist weniger als zehn, an die Enden von V-, D- und J-Segmenten vor deren Verbindung miteinander einzufügen [25, 42, 43].

Die V(D)J-Rekombination erfolgt in der frühen Entwicklung der B-Zellen im Knochenmark [25, 31]. (Anm.: Da nur in IGH-Rearrangements D-Gensegmente existieren, in den Antigenrezeptor-Gen-Rearrangements für die leichten Ketten jedoch nur V- und J-Segmente, wird zur allgemeinen Bezeichnung die Schreibweise „V(D)J-Rearrangement“ verwandt.) B-Lymphozyten ist ein weiterer Mechanismus der Affinitätssteigerungen ihrer Rezeptoren und Immunglobuline vorbehalten, die somatische Hypermutation (SHM), die in reifen B-Zellen nach antigenabhängiger Aktivierung während des Durchlaufens der Keimzentrumsreaktion erfolgt [44]. SHM erzeugt Punktmutationen im Exon der V-Region, bevorzugt im Bereich der CDRs, sowie mit geringer Frequenz Deletionen und Insertionen. In den Keimzentren werden B-Zellen mit erhöhter Antigenaffinität ihres BCR positiv selektiert, sodass SHM zur Affinitätsreifung der Antikörperantwort führt [45, 46]. Die aus SHM resultierenden Abweichungen in der Sequenz eines IGH-Rearrangements von den ursprünglich rekombinierten Gensegmenten der Keimbahn-DNA stellt des Weiteren ein Problem in der molekularen Klonalitätsdiagnostik dar (2.3.3).

2.2.3 B-Zell-Reifung im Kontext der malignen Entartung

In jedem Schritt der komplexen Reifungs- und DNA-Modifikationsprozesse der B-Zell-Entwicklung, besteht die Möglichkeit der malignen Entartung. Tatsächlich weisen die meisten reifen B-NHL einen Keimzentrums- (GC-) oder Post-

Keimzentrumsphänotyp auf, was darauf hinweist, dass die Keimzentrumsreaktion, durch starke Proliferation kombiniert mit extensiven DNA-Modifikationsprozessen, ein großes Risiko der malignen Entartung birgt [47, 48]. Abhängig vom Entwicklungsstadium, in dem die Ursprungszelle sich zum Zeitpunkt der Entartung befand, resultieren distinkte Lymphomtypen, die häufig spezifische Charakteristika mit ihrer Ursprungszelle teilen. Die Kenntnis der Stadien und Schritte der B-Zell-Entwicklung ist also notwendig, um Pathomechanismen der Entstehung und des klinischen Verlaufes verschiedener Lymphomtypen, sowie die diagnostische Bedeutung der Modifikationen ihrer Antigenrezeptor-Gene zu verstehen.

Physiologie der B-Zell-Reifung

Die primären lymphatischen Organe des Menschen sind der Thymus und das Knochenmark. In Letzterem vollzieht sich der erste Entwicklungsabschnitt zu reifen, naiven B-Zellen, indem zunächst aus multipotenten hämatopoetischen Stammzellen lymphoide Vorläuferzellen entstehen [49, 50]. Der zentrale Prozess ihrer Reifung im Knochenmark ist die Generierung eines funktionstüchtigen BCR durch V(D)J-Rekombination, der einerseits ein Antigen erkennen und binden können muss, andererseits nicht auf körpereigene Epitope autoreagieren darf [17, 51].

Der erste Schritt in der B-Zell-Entwicklung ist die Bildung des DJ-Rearrangements wie oben beschrieben, das in beiden Allelen des IGH-Locus erfolgen kann [17]. Zellen mit erfolgtem DJ-Rearrangement werden als Pro-B-Zellen bezeichnet [7, 49]. Anschließend erfolgt die Verknüpfung mit einem V-Gensegment und es resultiert ein VDJ-Rearrangement. Ist dieses funktional, so werden Rearrangement-Prozesse des anderen Allels unterdrückt, was als allelische Exklusion bezeichnet wird [52]. Ist dies nicht der Fall, weil beispielsweise eines der funktionsuntüchtigen V-Segmente verwendet wurde oder Nukleotidinsertionen zu einer Leserasterverschiebung des IGH-Gens führen, kann entweder ein VH-Replacement, d.h. ein Austausch des ursprünglich rekombinierten V-Segments gegen ein anderes erfolgen, oder das andere Allel

des IGH-Locus rearrangiert werden, um auf diesem Weg eine funktionale schwere Kette zu generieren [21, 53]. Wird schließlich eine funktionale schwere Kette exprimiert, erreicht die Zelle das Stadium der Prä-B-Zelle [7]. Vorläuferzellen, denen dies trotz Ausweichmechanismen nicht gelingt, gehen in Apoptose [54].

In Prä-B-Zellen erfolgt nun zunächst die VJ-Rekombination des κ -Locus zur Generierung einer leichten Kette. Nur falls durch Rearrangement beider κ -Loci keine funktionierende leichte Kette generiert werden kann, erfolgt die VJ-Rekombination eines oder beider λ -Loci [17]. Lediglich einer der beiden Isotypen wird schließlich exprimiert, was als Isotyp-Exklusion bezeichnet wird [52]. Bilden schwere und leichte Kette einen funktionierenden BCR, erreicht die Zelle das Stadium der unreifen B-Zelle. Ihr BCR wird schließlich darauf getestet, ob körpereigene Strukturen als Antigen gebunden werden; diese Eigenschaft, die als Autoreaktivität bezeichnet wird, kann Ausgang autoimmuner Prozesse sein und führt daher physiologischerweise zum Untergang der unreifen B-Zelle. Liegt keine Autoreaktivität vor, entsteht eine reife, naive B-Zelle, die IgM und IgD koexprimiert, das Knochenmark verlässt und im peripheren Blut und in den sekundären lymphatischen Organen zirkuliert [49, 51].

Bei Kontakt mit ihrem spezifischen Antigen, werden naive, reife B-Zellen aktiviert, migrieren in die sekundären lymphatischen Organe und initiieren, gemeinsam mit T-Zellen und folliculären dendritischen Zellen (FDCs), die Keimzentrumsreaktion [7, 55]. Komplexe Interaktionsmuster der proliferierenden B-Zellen mit den übrigen beteiligten Zellen bedingen die Ausbildung verschiedener charakteristischer Zonen des Keimzentrums. Die sogenannte dunkle Zone besteht aus sich schnell teilenden, als Zentroblasten bezeichneten B-Lymphozyten, in denen der oben beschriebene Prozess der SHM stattfindet [55]. Die Einführung von Mutationen in den Exons der V-Regionen wirkt sich auf die Antigenaffinität der BCRs aus und kann diese steigern oder verringern. Während Zentroblasten mit gesteigerter Antigenaffinität Überlebenssignale von T-Zellen und FDCs erhalten, folglich proliferieren und weitere Zyklen somatischer Hypermutationen durchlaufen, gehen Zentroblasten mit verringerter Affinität ihres BCR in Apoptose [51]. Dieser Prozess führt Schritt für Schritt zu immer

spezifischeren BCR der GC-B-Zellen [30, 56]. Auch der DNA-Rekombinationsprozess zum Klassenwechsel, in dem die initial durch alternatives Spleißen verwendeten C μ - und C δ -Gensegmente zur Expression von IgM und IgD gegen eines der Gensegmente C α , C ϵ oder eines der vier C γ -Gensegmente ausgetauscht werden, kann optional in den Zentrozyten stattfinden. Als Resultat ändert sich die Konfiguration des konstanten Teils der schweren Kette und Immunglobuline vom Typ IgA, IgE oder IgG werden exprimiert [7, 26, 45]. Nach Durchlaufen der Keimzentrumsreaktion differenzieren die B-Zellen sich entweder zu Plasma- oder zu B-Gedächtnis-Zellen [7, 49].

Ursprung häufiger B-NHL

Das folliculäre Lymphom ist das häufigste niedriggradige B-NHL und betrifft vor allem Erwachsene mit einem Gipfel in der fünften bis sechsten Lebensdekade [48, 57]. Charakteristisch ist die Translokation t(14;18), deren Häufigkeit im Allgemeinen als etwa 90 % der Fälle betreffend angegeben wird und die zu einer Überexpression des anti-apoptischen BCL2-Proteins führt, indem dessen transloziertes Gen von Chromosom 18 unter Kontrolle des IGH-Locus auf Chromosom 14 gerät [58-61]. Zytologie und Architektur betreffende Ähnlichkeiten von FL legen nahe, dass ihr Ursprung in Zentrozyten und Zentroblasten der Keimzentrumsreaktion liegt [48]. In der Mehrzahl der FL-Fälle können mutierte V-Gensegmente und eine intraklonale Variation der V-Gensegmente auf Nukleotidebene beobachtet werden, was ebenfalls mit einem Ursprung in der Keimzentrumsreaktion vereinbar ist [62, 63]. Es wird diskutiert, ob diese Beobachtung auf „ongoing somatic mutation“ („anhaltende somatische Mutation“), also die weitergehende Aktivität des SHM-Apparates in den proliferierenden Lymphomzellen zurückzuführen ist [48].

Die chronisch lymphatische Leukämie (CLL) ist die häufigste B-Zell-Neoplasie bei älteren Patienten [64]. Sie lässt sich in eine Form mit mutiertem und eine mit unmutiertem V-Gensegment unterteilen [65, 66]. Es wird vermutet, dass die unmutierte CLL ihren Ursprung in unmutierten CD5⁺ reifen B-Zellen hat, während mutierte CLL von einer Gruppe CD5⁺CD27⁺ B-Zellen abstammen, die bereits die

Keimzentrumsreaktion durchlaufen hat [7, 67]. Der Unterteilung nach Mutationsstatus kommt prognostische Bedeutung zu [65].

Als DLBCL klassifizierte Lymphome sind die häufigsten Lymphome insgesamt betrachtet und machen etwa 40 % aller NHL aus [57]. Es handelt sich um eine heterogene Gruppe hinsichtlich Morphologie, klinischer Präsentation und Therapierbarkeit, die sich in mehrere Untergruppen aufteilen lässt [7, 68, 69]. Die beiden Hauptgruppen leiten sich vermutlich von jeweils unterschiedlichen Ursprungszellen ab. Eine Gruppe, die hohe Ähnlichkeit zu GC-B-Zellen aufweist, wird als GC-DLBCL bezeichnet [68, 69]. Diese weisen häufig „ongoing somatic mutations“ ihrer V-Gensegmente auf [70]. Die andere Gruppe hingegen ähnelt stark in-vitro-aktivierten B-Zellen, weist mutierte V-Gensegmenten ohne anhaltende SHM auf und gleicht am stärksten Immunoblasten, die die Keimzentrumsreaktion durchlaufen haben. Sie bilden die Gruppe der ABC-DLBCL („activated B-cell like“) [7].

Mantelzelllymphome (MCL) sind aggressive Lymphome, die etwa 6 % aller NHL ausmachen und im medianen Alter von 60 bis 70 Jahren auftreten [71, 72]. Fast alle MCL weisen die Translokation t(11;14) auf, die das BCL-1-Gen auf Chromosom 11 neben dem IGH-Locus auf Chromosom 14 platziert und zur Überexpression von Cyclin D1 führt [73, 74]. Die resultierende Hochregulation des Zellzyklus trägt maßgeblich zur Pathogenese von MCL bei [71, 75]. Die vermutete Ursprungszelle des MCL ist eine B-Zelle der Mantelzone, die nicht die Keimzentrumsreaktion durchlaufen hat und daher keine SHM aufweist [76-78]. Dieser Einschätzung steht entgegen, dass in Studien konstant eine Gruppe von MCL mit SHM ihrer V-Gensegmente beobachtet wurden [77-80]. Einen Überblick über die Ursprünge ausgewählter B-NHL in Relation zur Keimzentrumsreaktion bietet Abbildung 3.

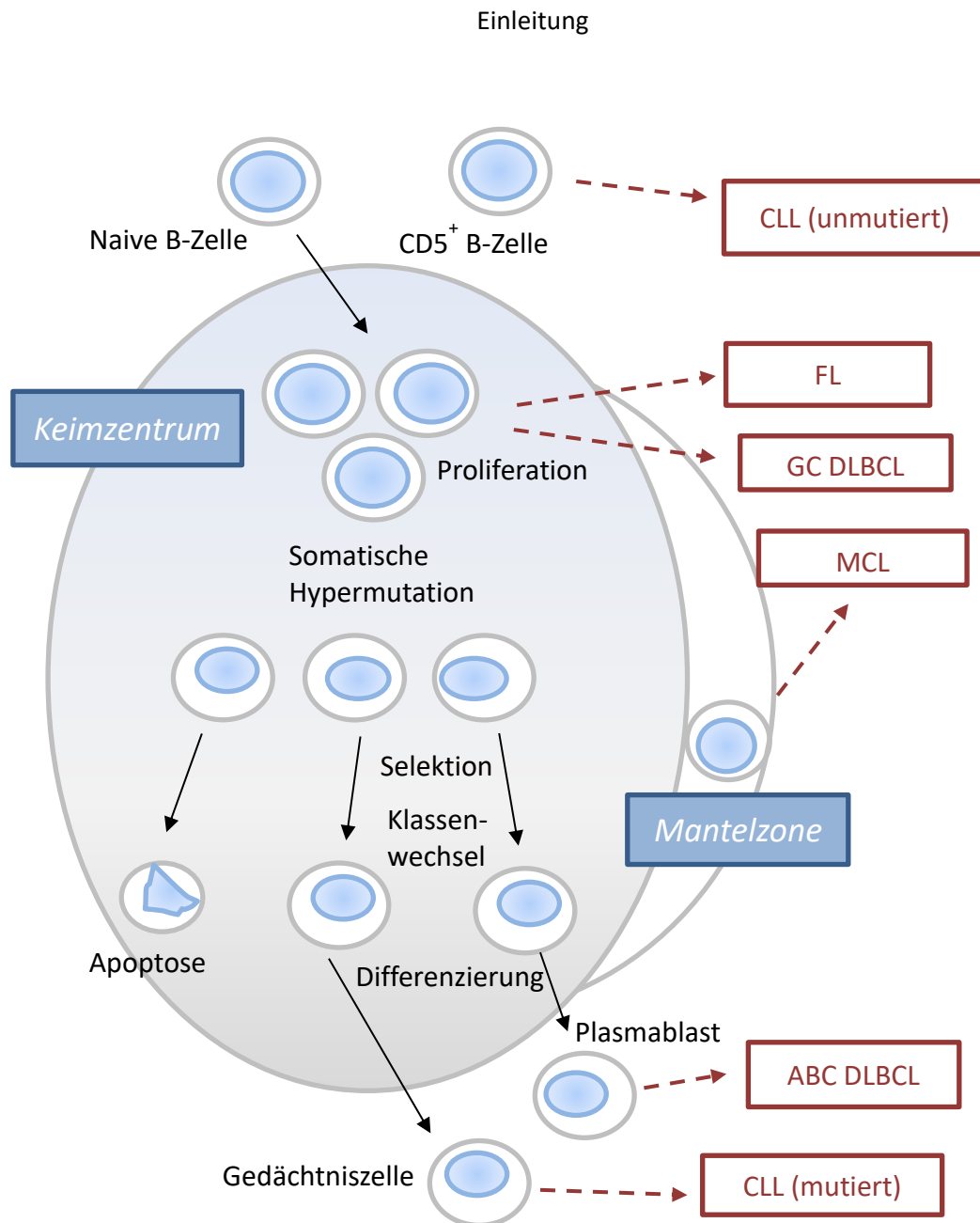


Abbildung 3: Ursprung ausgewählter B-NHL in Relation zur Keimzentrumsreaktion. Die Ursprünge der meisten B-NHL stehen in Zusammenhang mit der Keimzentrumsreaktion. Die unmutierte CLL hat ihren Ursprung vermutlich in unmutierten CD5⁺ reifen B-Zellen, während die mutierte CLL von B-Zellen abstammt, die bereits die Keimzentrumsreaktion durchlaufen haben. Das FL und das GC DLBCL entstehen aus stark proliferierenden Zellen der Keimzentrumsreaktion, während das ABC DLBCL möglicherweise aus Immunblasten jenseits der Keimzentrumsreaktion entsteht. Das MCL entstammt vermutlich B-Zellen der Mantelzone, die die Keimzentrumsreaktion nicht durchlaufen haben. Abbildung erstellt in Anlehnung an Seifert et al. [7].

2.3 Monoklonale BCR-Rearrangements als diagnostischer Marker maligner Lymphome

2.3.1 Monoklonalität in malignen Lymphomen

Krebs ist eine klonale Erkrankung [81]. Wie in Krebserkrankungen anderer Gewebe, lassen sich alle Zellen einer malignen Lymphoproliferation auf eine gemeinsame Ursprungszelle zurückführen. In malignen Lymphomen kommt diesem Umstand eine herausragende diagnostische Bedeutung zu. Im Laufe seiner Reifung generiert jeder Lymphozyt einen einzigartigen Antigenrezeptor durch das Rearrangement bestimmter DNA-Abschnitte. Diese Rearrangements werden im Falle der malignen Entartung und unkontrollierten klonalen Expansion an jede Zelle der neoplastischen Population weitergegeben. Das überwiegende Vorhandensein einer einzigen Konfiguration der Antigenrezeptor-Gene innerhalb einer Lymphozytenpopulation wird als Monoklonalität bezeichnet und ist hinweisend auf das Vorliegen einer malignen Lymphoproliferation. Im Gegensatz hierzu ist in einer reaktiven Lymphozytenpopulation eine Vielzahl verschiedener Antigen-Rezeptor-Konfigurationen nachweisbar, was als Polyklonalität bezeichnet wird.

2.3.2 Prinzipien der molekularen Klonalitätsdiagnostik

In der Mehrzahl der Fälle verdächtiger Lymphoproliferationen, kann anhand von Histo- und Zytomorphologie in Kombination mit Immunhistochemie und Durchflusszytometrie eine Diagnose gestellt werden [82]. Jedoch bestehen, trotz Anwendung dieser Methoden, in einem gewissen Anteil von etwa 5 bis 10 % der Fälle Probleme in der Diagnosestellung von Lymphoproliferationen [36]. In solchen Fällen wird auf molekularpathologische Diagnostik zum Klonalitätsnachweis zurückgegriffen. Lange Zeit kam hierfür in erster Linie die Southern-Blot-Analyse zum Einsatz, die allerdings, aufgrund hohen zeitlichen und technischen Aufwandes, hoher notwendiger DNA-Ausgangsmenge und -

Qualität und begrenzter Sensitivität, zunehmend durch PCR-basierte Methoden verdrängt wurde [83].

Um zwischen polyklonalen Lymphozytenpopulationen und monoklonalen Lymphoproliferationen mittels PCR unterscheiden zu können, sollten die PCR-Primer alle möglichen V(D)J-Rearrangements amplifizieren können und die erzeugten Amplifikate die Verbindungsregion von V- und J-Segment, im Falle des IGH-Rearrangements das gesamte D-Segment umfassend, als Region mit der höchsten Variabilität beinhalten. Wird das V(D)J-Repertoire einer polyklonalen Lymphozytenpopulation auf diese Weise amplifiziert, so entstehen, aufgrund der Vielfalt verschiedener V(D)J-Rearrangements, Produkte verschiedener Länge, die eine Gaußsche Verteilung aufweisen. Im Gegensatz hierzu liefert die Amplifikation einer monoklonalen Lymphoproliferation nur Amplifikate gleicher Länge, der Länge des V(D)J-Rearrangements der malignen Zellen (Abbildung 4).

Durch die BIOMED-2 Concerted Action BMH4 CT98-3936 wurden standardisierte PCR-Protokolle zur Klonalitätstestung in rearrangierten Ig- und T-Zell-Rezeptor- (TCR-)Genen entwickelt und 2003 veröffentlicht, die möglichst alle Rearrangements amplifizieren sollten. Zu den entwickelten Assays zählten drei für die Amplifikation kompletter IGH-Rearrangements (VH-JH), zwei für inkomplette DH-JH-Rearrangements, zwei für IGK und eines für IGL (Gen-Rearrangements des κ - bzw. λ -Isotyps der leichten Kette) [36].

Basierend auf Sequenzhomologien können V- und D-Gensegmente in Familien eingeteilt werden. Im Fall des IGH-Locus können die Segmente auch als VH-, D- und JH-Segmente bezeichnet werden. So verteilen sich die 44 funktionalen VH-Segmente des IGH-Locus auf sieben Familien [84], was man sich beim Design passender PCR-Primer zunutze machte. Die drei Primersets zur Bindung an den verschiedenen FRs bestehen aus je einem Primer pro VH-Segmentfamilie, der in einem homologen Bereich der Familie bindet. Dies sind für die Primersets FR2 und FR3 je sieben Primer, für das Primerset FR1 nur sechs, da sich die Segmentfamilien VH1 und VH7 mit demselben Primer amplifizieren lassen. Der Reverse-Primer ist für alle Primersets derselbe und bindet an die Stelle der höchsten Homologie der sechs JH-Segmente [36]. Die Detektion der erzeugten

Amplifikate kann beispielsweise mittels GeneScan-Fragment-Analyse erfolgen. Zu diesem Zweck sind die Reverse-Primer der BIOMED-2 Primersets, wie beispielsweise der JH-Konsensus-Primer, fluoreszenzmarkiert.

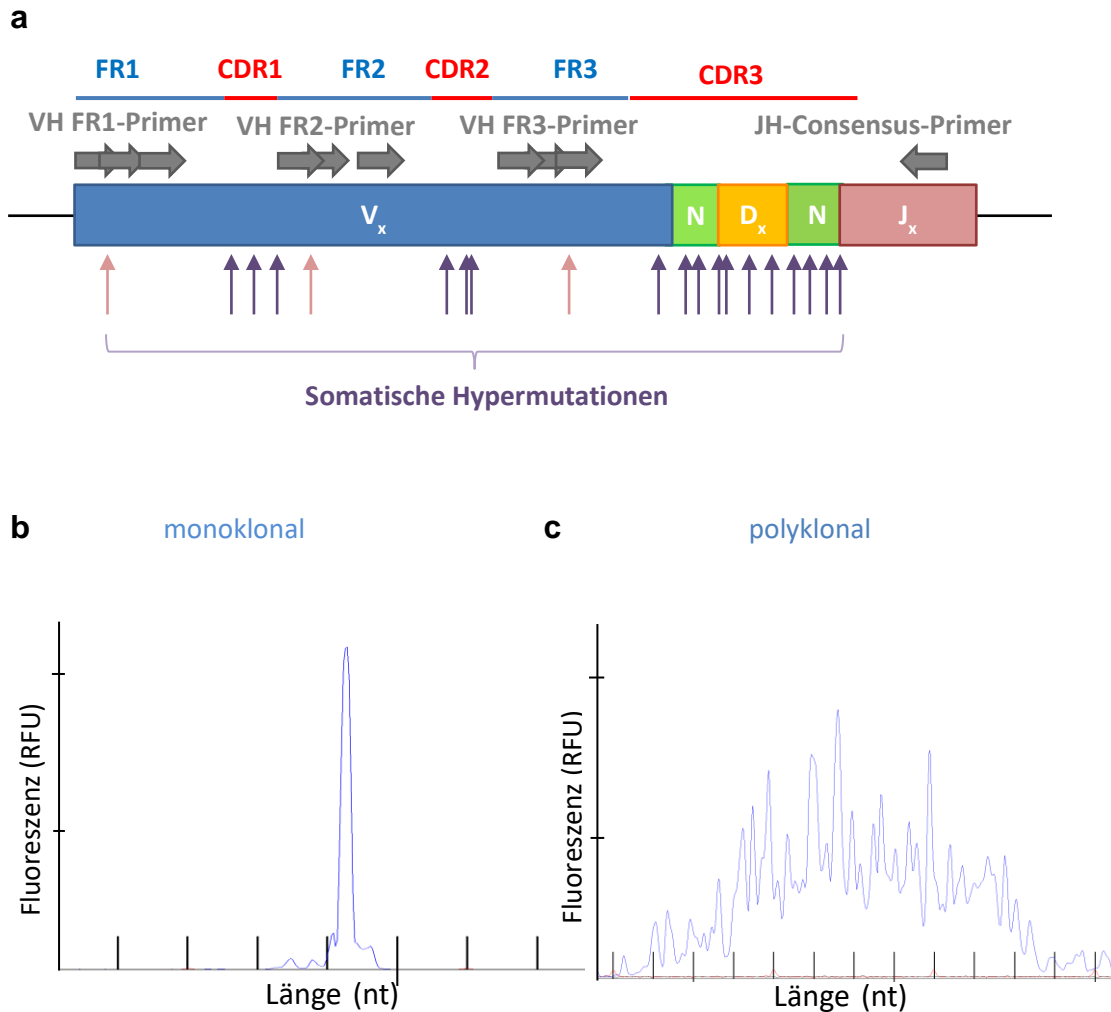


Abbildung 4: Klonalitätsdiagnostik mittels IGH-Amplifikation und Fragmentanalyse. a: IGH-Rearrangement mit BIOMED-2 Primer-Bindungsstellen und somatischen Hypermutationen. Das IGH-Rearrangement besteht aus einem V-, einem D- und einem J-Gensegment. An beiden Enden des D-Segments können sich N-Insertionen und -Deletionen befinden. Die drei verschiedenen Primersets zur Amplifikation von IGH-Rearrangements binden in je einer der drei konservierten Framework Regionen FR1, FR2 und FR3 in Kombination mit dem JH-Konsensus-Primer, der in einem konservierten Bereich des J-Gensegments bindet. Somatische Hypermutationen finden sich bevorzugt in den CDRs (lila Pfeile), sie können aber auch in den FRs auftreten (rosa Pfeile) und dann die Primer-Bindung beeinträchtigen. Abbildung erstellt in Anlehnung an die ImMunoGeneTics-Datenbank und van Dongen et al. [35, 36]. **b: Repräsentatives Elektropherogramm eines monoklonalen Rearrangements.** Entsprechend der identischen Länge der IGH-Rearrangements einer klonalen Lymphozytenpopulation resultiert ein monoklonaler Peak. **c: Repräsentatives Elektropherogramm einer polyklonalen Rearrangement-Verteilung.** Die verschiedenen langen IGH-Rearrangements einer reaktiven Lymphozytenpopulation ergeben ein normalverteiltes polyklonales Peakmuster.

In den mehr als zehn Jahren seit Einzug der BIOMED-2 Primer in die Klonalitätsdiagnostik wurden ihnen in zahlreichen Studien an verschiedenen Entitäten und Materialien eine hohe Sensitivität und Spezifität, sowie eine zuverlässige Reproduzierbarkeit der Assays bestätigt [85-94]. In Neoplasien der B-Zellreihe weist die Kombination von IGH- und IGK-Assays die höchste Sensitivität auf. Die BIOMED-2 Assays zum Klonalitätsnachweis sind kommerziell verfügbar und gelten mittlerweile als weltweiter Standard [95].

2.3.3 Probleme der PCR-basierten Klonalitätsdiagnostik

Wie oben ausgeführt, konnte die Klonalitätsdiagnostik mittels BIOMED-2 Primersets andere PCR-Verfahren und den Southern Blot durch ihre Zuverlässigkeit und gute Praktikabilität ersetzen. Jedoch ist die Methode nicht ohne ihre Schwächen.

So können SHM im Bereich einer Primer-Bindungsstelle zu einem falsch negativen Befund mit dem jeweiligen Primerset führen. Dieses Problem betrifft Lymphome, deren Ursprungszellen die Keimzentrumsreaktion bereits durchschritten hatten, oder sich zum Zeitpunkt der malignen Entartung eben in jener befanden, also insbesondere FL und DLBCL in denen immer, sowie CLL, in denen in einem gewissen Teil der Fälle SHM vorliegen [47, 65, 96-98]. Hier ist die Verwendung mehrerer Primersets zum Klonalitätsnachweis häufig notwendig.

Ein weiteres Kollektiv problematischer Fälle stellen solche aus Formalin-fixiertem Paraffin-eingebettetem (FFPE-) Material dar, die eine ungenügende DNA-Qualität aufweisen können. Je nach verwendetem BIOMED-2 Primerset sind eventuell keine ausreichend großen DNA-Fragmente enthalten. Zur Amplifikation mit dem VH-Primerset FR1 beispielsweise sollten DNA-Fragmente von mindestens 310 bis 360 bp Länge in der Probe enthalten sein [36].

Eine geringe Gesamtzahl an Lymphozyten in einer Probe kann nicht nur zu falsch negativen, sondern auch zu falsch positiven Befunden führen. Im Falle eines eingeschränkten Rezeptorrepertoires bei geringer Lymphozytenzahl können

aufgrund der hohen Sensitivität der PCR einige wenige Rearrangements reaktiver Lymphozyten amplifiziert werden und zu einer oder mehreren prominenten Banden führen, was als Pseudoklonalität bezeichnet wird [36]. Auch reaktive Lymphozytenproben mit höherer Zellzahl können eine eingeschränkte Diversität und schwer interpretierbare PCR-Ergebnisse aufweisen, wenn einige wenige Antigen-selektierte Subklone stark vertreten sind, was als Oligoklonalität bezeichnet wird [36] (Abbildung 5).

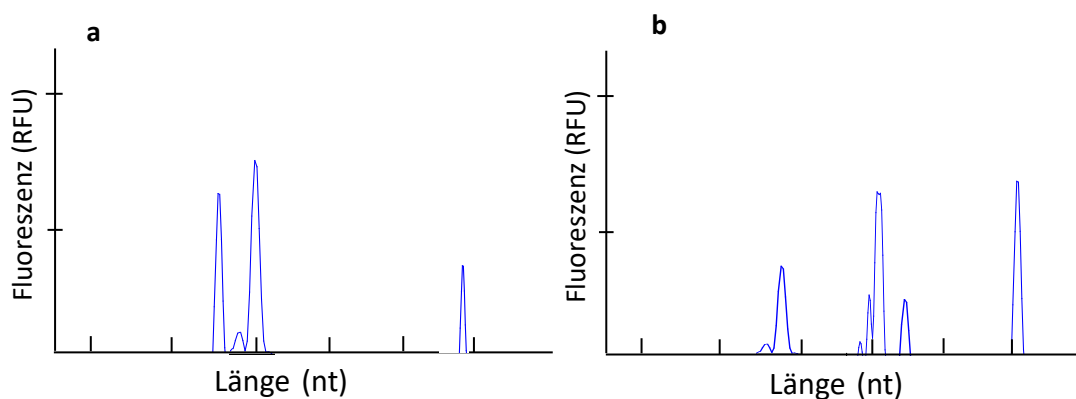


Abbildung 5: Repräsentatives Elektropherogramm einer pseudoklonalen bzw. oligoklonalen Rearrangement-Verteilung. a und b stellen Ansätze mit unterschiedlicher DNA-Ausgangsmenge (beispielsweise 50 und 200 ng) derselben untersuchten Probe dar. Es resultiert jeweils ein Elektropherogramm mit mehreren prominenten Peaks, die sich jedoch im dargestellten Doppelansatz nicht reproduzieren lassen. Ein solcher Befund kann Resultat ungenügender DNA-Qualität oder -Menge sein, beispielsweise aus Proben mit geringer Lymphozytenzahl oder aus FFPE-Gewebe. Je nach Verteilungsmuster der pseudoklonalen Peaks besteht die Gefahr der falsch positiven Interpretation als klonal.

Auch ein geringes Verhältnis von malignen Zellen zu reaktiver Hintergrundpopulation kann der Detektion von Klonalität entgegenstehen. Eine klonale Population, die weniger als 5 bis 10 % der Lymphozyten einer Probe ausmacht, kann in einer PCR zur Klonalitätstestung ein falsch negatives Ergebnis liefern [36]. Dies schließt folglich eine Anwendung zur Detektion von „minimal residual disease“ (MRD), also dem Verbleib kleinster Zahlen maligner Lymphomzellen nach erfolgter Therapie im Remissionszustand, aus.

Abschließend ist zu ergänzen, dass der Nachweis von Klonalität nicht gleichzusetzen ist mit dem Vorliegen von Malignität. So existieren benigne Lymphoproliferationen mit klonalem Ursprung, benigne monoklonale Gammopathien und klonale Expansionen im Rahmen der Immunantwort [36].

2.4 Klonalitätsnachweis mittels „Next Generation Sequencing“

2.4.1 Grundlagen und -überlegungen

Lange Zeit war die Sanger-Sequenzierung, basierend auf der Ketten-Abbruch-Synthese, die Methode der Wahl, um die Sequenz von DNA-Molekülen zu bestimmen [99]. Es folgte die Entwicklung von sog. „Next Generation Sequencing“-Methoden („Sequenziermethoden der nächsten Generation“, NGS), auch Hochdurchsatz-Sequenzierung (HTS) oder massiv parallele Sequenzierung genannt, die durch Parallelisierung, Automatisierung und Computerisierung einen enormen Sprung im Sequenzierdurchsatz ermöglichten [100-102]. Grundlage vieler unter den Begriff NGS gefasster Technologien ist dabei zunächst die Generierung einer sog. Bibliothek („library“). In der vorliegenden Arbeit versteht sich hierunter die PCR-Amplifikation der zu untersuchenden DNA-Abschnitte in einer Probe mittels passender Primersets. Im darauffolgenden Sequenzierungsschritt kann nun prinzipiell jedes einzelne derart erzeugte Amplikon parallel sequenziert werden. Dem Sequenzierungsabschnitt vorgeschaltet ist bei aktuellen Plattformen ein Schritt der lokalen klonalen Amplifikation der Amplikons auf einer Flow-Cell, die dazu dient, lokale Cluster identischer DNA-Moleküle zu generieren, von denen im Verlauf der Sequenzierreaktion ein detektierbares Signal ausgeht. Durch Verwendung spezieller Barcode-Primer bei der Erstellung der Library ist auch die Sequenzierung der Amplikons mehrerer Proben gleichzeitig und die anschließende Zuordnung der generierten Sequenzen zur Ausgangsprobe möglich. Die gewonnenen Rohdaten können in einem dritten Schritt einer bioinformatischen Auswertung unterzogen werden, die beispielsweise Aussagen zur Validität der Detektion und zur Häufigkeit einzelner Sequenzvarianten liefern kann. Durch die Anwendung von NGS-Verfahren eröffneten sich rasant neue Einsichten in zahlreiche medizinische Fragestellungen, nicht zuletzt in den Bereichen der Onkologie und Hämatopathologie [103].

Diverse Studien zur NGS-basierten Klonalitätsuntersuchung erprobten in den letzten Jahren unterschiedliche Vorgehensweisen und konnten den Nutzen der

Hochdurchsatzsequenzierung von IGH- und TCR-Rearrangements für verschiedene Fragestellungen in diesem Kontext demonstrieren und verschiedene Vorteile der Betrachtung von Antigenrezeptor-Repertoires auf Sequenzebene demonstrieren [104-112]. So ließen sich klonale Rearrangements zuverlässig mittels NGS detektieren und darüber hinaus quantifizieren [104-106], teils mit größerer Sensitivität als mittels GeneScan-Analyse [107, 112]. Dieser Umstand, in Kombination mit der Möglichkeit der Charakterisierung detektierter Rearrangements auf Sequenzebene, qualifiziert NGS-basierte Vorgehensweisen darüber hinaus für den Einsatz in der MRD-Diagnostik [108-112]. Auch die Beurteilung von Mutationen innerhalb detektierter Rearrangements, deren Vorliegen im Falle mancher Entitäten prognostisch bedeutsam ist, ist im Zuge der Klonalitätsdiagnostik mittels NGS möglich und scheint ein aussichtsreiches Anwendungsgebiet darzustellen [113].

Bisher konnte sich ein universell anwendbares Vorgehen jedoch nicht durchsetzen und es werden weiterhin verschiedene Ansätze und Vorgehensweisen erprobt. Insbesondere zur Anwendbarkeit NGS-basierter Klonalitätsstestung bei Keimzentrums- und Post-Keimzentrums-B-NHL mit mutierten IGH-Rearrangements sowie zur Anwendbarkeit an FFPE-Material ist die Studienlage bisher dünn.

2.4.2 Zielsetzung der Arbeit

In dieser Studie sollte ein NGS-Verfahren zum Klonalitätsnachweis in B-NHL etabliert werden, in welchem zunächst mittels modifizierter BIOMED-2 FR1- und FR2-Primersets das IGH-Repertoire einer Lymphomprobe amplifiziert, und in einem zweiten PCR-Schritt Barcode-Adapter für den NGS-Prozess hinzugefügt wurden. Die Sequenzierung erfolgte auf der Illumina MiSeq-Plattform (Abbildung 6). Die gewonnenen Datensätze wurden mittels IMGT/HighV-Quest analysiert, einem Online-Tool, das es erlaubt NGS-Datensätze mit großen Zahlen an Sequenzen rearrangierter Immunglobulin- und T-Zell-Rezeptor-Gene zu analysieren und zu charakterisieren [114].

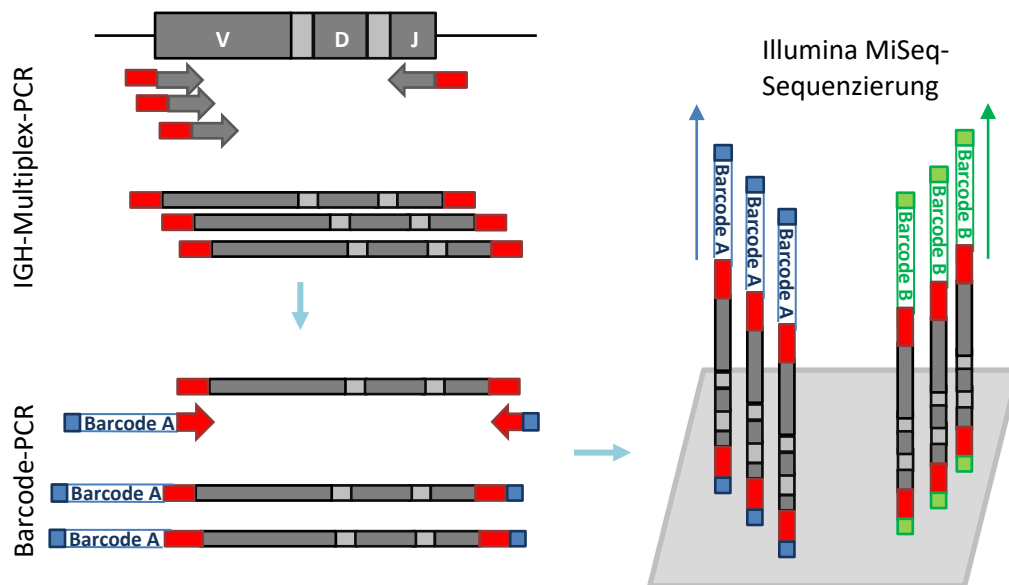


Abbildung 6: Schematische Darstellung der PCR-Schritte und Sequenzierung im Rahmen dieser Arbeit. Zunächst erfolgte die IGH-Amplifikation mittels BIOMED-2 VH-Primersets (graue Pfeile), die mit einem Fluidigm-Anhang modifiziert waren (rot), der die Bindung von Barcode-Primern (rote Pfeile) erlaubte. Über diese wurden die IGH-Amplifikate in einem zweiten PCR-Schritt mit Barcode-Sequenzen versehen (blau bzw. grün). Diese wiesen für jeden untersuchten Ansatz eine individuelle Sequenz auf, was die massiv parallele Sequenzierung auf der Illumina MiSeq-Plattform erlaubte: Anhand der Barcode-Sequenzen konnten die detektierten IGH-Sequenzen im Anschluss an die Sequenzierung wieder eindeutig den ursprünglichen Ansätzen zugeordnet werden. Die gewonnenen NGS-Rohdaten wurden einer bioinformatischen Auswertung mittels IMG/HighV-Quest unterzogen (nicht dargestellt).

Anhand von Verdünnungsproben klonaler B-NHL-Zelllinien-DNA in polyklonaler DNA (PBMC-DNA, „peripheral blood mononuclear cells“) sollten Interpretations- und Auswertungsansätze der analysierten Daten im Hinblick auf Klonalität erarbeitet, und zentrale Parameter der Klonalitätsuntersuchung mittels NGS ermittelt werden. Anschließend sollte die Anwendung auf Probenkollektive von FFPE-Patientenmaterial erfolgen. Hierzu zählten Patientenproben mit hohem Lymphomzellgehalt, die zum Vergleich zusätzlich mittels Sanger-Sequenzierung und GeneScan-Analyse untersucht wurden. Des Weiteren wurden Proben von Patienten im zeitlichen Erkrankungsverlauf untersucht. Schließlich erfolgte die Anwendung des Verfahrens an Glaskörperaspiraten intraokulärer Lymphome (IOL) mit geringem Lymphomzellgehalt.

Ziel der Studie war die Etablierung des geschilderten Verfahrens für die molekularpathologische Klonalitätsdiagnostik. Insbesondere sollten Vorteile gegenüber der GeneScan-basierten Klonalitätsuntersuchung und der

konventionellen Sequenzierung nach Sanger herausgearbeitet werden. Hierfür zentral war zunächst, ob eine Abgrenzbarkeit klonaler Populationen zu Populationen des polyklonalen Hintergrundes gegeben ist, und wie diese Abgrenzbarkeit von den quantitativen Verhältnissen in der Ausgangsprobe abhängt. Auch die Möglichkeit, mittels NGS Lymphom- und Lymphozytenpopulationen im zeitlichen Verlauf anhand ihrer IGH-Rearrangements zu verfolgen, sollte erprobt, und ermittelt werden, ob die Methode sich zur Detektion und zum Monitoring von MRD eignet. Zentral war ebenfalls die Frage, ob sich mittels NGS ein zusätzlicher Informationsgewinn bei problematischen Proben mit geringem Lymphomzellgehalt erzielen lässt. Schließlich sollte untersucht werden, ob sich sowohl im zeitlichen Verlauf als auch punktuell somatische Hypermutationen der detektierten IGH-Rearrangements erkennen und quantifizieren lassen, und ob sich eine Aussage zur Entwicklung des Mutationsstatus im zeitlichen Verlauf treffen lässt.

3. Material und Methoden

3.1 Liste der verwendeten Reagenzien und Geräte

3.1.1 Chemikalien und Reagenzien

Xylol	VWR
Ethanol	Merck, KGaA
Isopropanol	AppliChem
PBS	Thermo Fisher Scientific
SLS	Beckmann Coulter
Agencourt® CleanSEQ Kit	Beckmann Coulter
Agencourt® AMPure XP	Beckmann Coulter
LE Agarose	Biozym Scientific GmbH
1xTBE-Puffer aus	
- TRIS 0,089 M	Sigma Aldrich
- Borsäure 0,089 M	Merck, KGaA
- EDTA-Na ₂ 0,002 M	AppliChem
Gel-Red™ Nucleid Acid Stain	Biotium
GeneRuler™, 100 bp DNA-ladder	Thermo Fisher Scientific
dNTPs	Thermo Fisher Scientific

3.1.2 Enzyme

Proteinase K	Qiagen
RNAse A	Qiagen
AmpliTaq Gold DNA Polymerase	Thermo Fisher Scientific
FastStart High Fidelity PCR System	Roche

3.1.3 Kits

QIAamp®DNA FFPE Tissue Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
GenomeLab DTCS Quick Start Kit	Beckmann Coulter

3.1.4 Zelllinien

Granta-519	leukämische Transformation eines Mantelzell-Lymphoms Grad IV, 1991 aus dem peripheren Blut einer 58-jährigen Frau gewonnen [115]
OCI-LY3	DLBCL vom Nicht-Keimzentrums-Typ (ABC DLBCL), 1983 aus dem Knochenmark eines 52-jährigen Mannes zum Zeitpunkt eines Rückfalls gewonnen [116]

3.1.5 Primer (Sigma-Aldrich)

Primer zur Überprüfung der Amplifizierbarkeit der DNA (nach van Dongen et al. [36])

AF4/X3U	GGAGCAGCATTCCATCCAGC
AF4/X3L	CATCCATGGGCCGGACATAA
AF4/X11U	CCGCAGCAAGCAACGAACC
AF4/X11L	GCTTTCCTCTGGCGGCTCC
PLZF/X1U	TGCGATGTGGTCATCATGGTG
PLZF/X1L	CGTGTCATTGTCGTCTGAGGC
RAG1/X2U	TGTTGACTCGATCCACCCCA
RAG1/X2L	TGAGCTGCAAGTTTGGCTGAA
TBXAS1/X9U	GCCCGACATTCTGCAAGTCC
TBXAS1/X9L	GGTGTGCCGGGAAGGGTT

Primer für die Sanger-Sequenzierung (Fluidigm)

Fluidigm_F	ACACTGACGACATGGTTCTACA
Fluidigm_R	TACGGTAGCAGAGACTTGGTCT

Primer zur GeneScan-Analyse (nach van Dongen et al. [36])

Framework Region 2	
VH1-FR2	CTGGGTGCGACAGGCCCTGGACAA
VH2-FR2	TGGATCCGTGAGCCCCAGGGAAGG
VH3-FR2	GGT CCG CCA GGC TCC AGG GAA
VH4-FR2	TGGATCCGCCAGCCCCAGGGAAGG
VH5-FR2	GGGTGCGCCAGATGCCCCGGGAAAG G
VH6-FR2	TGGATCAGGCAGTCCCCATCGAGAG
VH7-FR2	TTGGGTGCGACAGGCCCTGGACAA
JH-Konsensus-Primer	
JH Consensus	D4-CTTACCTGAGGAGACGGTGACC

Primer zur IGH-Amplifikation mit Fluidigm-Anhang (Fluidigm)

Framework Region 1

VH1-FR1_Flui ACACTGACGACATGGTTCTACAGGCCTCAGTGAAGGTCTCCTGCAAG
 VH2-FR1_Flui ACACTGACGACATGGTTCTACAGTCTGGTCCTACGCTGGTAAACCC
 VH3-FR1_Flui ACACTGACGACATGGTTCTACACTGGGGGGTCCCTGAGACTCTCCTG
 VH4-FR1_Flui ACACTGACGACATGGTTCTACACTTCGGAGACCCTGTCCCTCACCTG
 VH5-FR1_Flui ACACTGACGACATGGTTCTACACGGGGAGTCTCTGAAGATCTCCTGT
 VH6-FR1_Flui ACACTGACGACATGGTTCTACATCGCAGACCCTCTCACTCACCTGTG

Framework Region 2

VH1-FR2_Flui ACACTGACGACATGGTTCTACACTGGGTGCGACAGGCCCTGGACAA
 VH2-FR2_Flui ACACTGACGACATGGTTCTACATGGATCCGTACCCCCAGGGAAGG
 VH3-FR2_Flui ACACTGACGACATGGTTCTACAGGTCCGCCAGGCTCCAGGGAA
 VH4-FR2_Flui ACACTGACGACATGGTTCTACATGGATCCGCCAGCCCCAGGGAAGG
 VH5-FR2_Flui ACACTGACGACATGGTTCTACAGGGTGCGCCAGATGCCCGGGAAAGG
 VH6-FR2_Flui ACACTGACGACATGGTTCTACATGGATCAGGCAGTCCCCATCGAGAG
 VH7-FR2_Flui ACACTGACGACATGGTTCTACATTGGGTGCGACAGGCCCTGGACAA

JH-Konsensus-Primer

JH_consensus_Flui TACGGTAGCAGAGACTTGGTCTTACCTGAGGAGACGGTGACC

Barcode-Primer (Fluidigm)

Fluidigm® Access Array™ - Access Array Barcode Library for Illumina Sequencers - 384

3.1.6 Geräte

Geräteklasse	Gerätebezeichnung	Hersteller
Elektrophorese-system	Electrophoresis Systems	Owl Separation Systems Inc.
	PerfectBlue Breitformat-Gelsystem	VWR Life Science Competence Center
Geldokumentations-system	Quantum Geldokumentationssystem	VWR Life Science Competence Center
Konzentrator	Eppendorf Concentrator 5301	Eppendorf
Magnetplatte	Agencourt SPRIPlate SP 384 MagnetPlate	Beckman Coulter
Magnetrührer	IKAMAG®RCT	IKA-Werke
Mikrowellengerät	Microwave 800	Severin Elektrogeräte GmbH
Pipetten	Eppendorf	Eppendorf
	Pipetman Gilson	Gilson
Thermocycler	Thermocycler ABI 9700	Applied Biosystems
Thermoschüttler	Thriller® Inkubationsschüttler	VWR Life Science Competence Center
Spectrometer	NanoDrop 2000	Thermo Fisher Scientific

UV-Bestrahlungsgerät	HERAsafe®	Thermo Fisher Scientific
Vortexer	Mixer UZUSIO VTX-3000L	LMS Co., Ltd.
	Vortex-2-Genie®	Scientific Industries, Inc.
Zentrifugen	Biozym Sproud	Biozym Scientific GmbH
	Biofuge fresco	Heraeus
	COMBI-SPIN FVL-2400	VWR Life Science Competence Center
	Minizentrifuge	neoLab
	Mini Centrifuge	Labnet International, Inc.
	Microfuge®16 Centrifuge	Beckman Coulter

3.1.7 Verbrauchsmaterialien

Material	Bezeichnung	Hersteller
Pipettenspitzen	Safe Seal Tips	Biozym Scientific GmbH
	Sapphire Pipette Tips	Greiner Bio-One GmbH
Reaktionsgefäße	1,5 ml, 2 ml Reaktionsgefäße	Eppendorf
	0,2 ml Multiply®-µStrip pro 8-strip	Sarstedt AG & Co

3.1.8 Programme und Tools

GenomeLab GeXP Genetic Analysis System	Beckman Coulter Brea
GenomeLab GeXP Software 10.2	Beckman Coulter Brea
IMGT/V-Quest	IMGT
IMGT/HighV-Quest	IMGT
BLAST	NCBI

3.1.9 Hersteller

AppliChem	Darmstadt
Beckmann Coulter	Krefeld
Biotium	Fremont, CA, USA
Biozym Scientific GmbH	Hessisch Oldendorf
Eppendorf	Hamburg
Fluidigm	South San Francisco, CA, USA
Gilson	Middleton, WI, USA
Greiner Bio-One GmbH	Kremsmünster, Österreich

Heraeus	Hanau
IKA-Werke	Staufen im Breisgau
Labnet International, Inc.	Woodbridge, NJ, USA
Merck, KGaA	Darmstadt
neoLab	Heidelberg
Owl Separation Systems Inc.	Portsmouth, NH, USA
Qiagen	Hilden
Roche	Penzberg
Sarstedt AG & Co	Nümbrecht
Severin Elektrogeräte GmbH	Sundern
Sigma Aldrich	Steinheim
Thermo Fisher Scientific	Waltham, MA, USA
VWR	Darmstadt
VWR Life Science Competence Center	Erlangen

3.2 Extraktion und Aufreinigung von DNA

3.2.1 Extraktion und Aufreinigung von DNA aus FFPE-Gewebe

Um DNA aus FFPE-Gewebe molekularpathologischen Untersuchungen zugänglich zu machen, musste zunächst eine Entparaffinierung durchgeführt werden. Als Ausgangsmaterial dienten ein bis zehn Paraffinschnitte von 5 µm Dicke in einem 1,5 ml Reaktionsgefäß. Zunächst wurde 1 ml 100%iges Xylol zugegeben, 5 s gevortext und im Anschluss für 5 min bei 13.000 rpm zentrifugiert. Am Boden des Reaktionsgefäßes bildete sich ein Pellet, der Überstand wurde abgenommen und verworfen. Dieses Vorgehen wurde ein weiteres Mal mit 1 ml Xylol und anschließend zwei weitere Male mit 100%igem Ethanol wiederholt. Dann wurde das Pellet bei 50 °C im Thermoschüttler getrocknet.

Die entparaffinierten Zellen der Probe mussten nun lysiert werden. Im Folgenden wurde das QIAamp®DNA FFPE Tissue Kit verwendet. Zunächst wurde das Pellet in 180 µl Puffer ATL resuspendiert, es wurden 20 µl Proteinase K hinzugefügt und bei 56 °C und 500 rpm im Thermoschüttler 18 h inkubiert. Es folgte ein weiterer Inkubationsschritt bei 90 °C für 1 h. Nach kurzem Zentrifugieren wurden 2 µl RNase A hinzugefügt und das Gemisch 2 min bei Raumtemperatur inkubiert. Nun wurden 200 µl Puffer AL und 200 µl 100%iges

Ethanol zugegeben, das Reaktionsgefäß jeweils gevortext und im Anschluss kurz zentrifugiert. Zur Aufreinigung des Lysats wurde dieses nun in eine QIAamp MinElute Säule überführt, die sich wiederum in einem 2 ml Reaktionsgefäß befand, und für 1 min bei 8.000 rpm zentrifugiert. Die QIAamp MinElute Säule wurde in ein neues 2 ml Reaktionsgefäß überführt und das alte Reaktionsgefäß mitsamt Durchfluss verworfen. 500 µl Puffer AW1 wurden auf die QIAamp MinElute Säule gegeben und bei 8.000 rpm zentrifugiert. Erneut wurde die QIAamp MinElute Säule in ein neues 2 ml Reaktionsgefäß überführt und das alte Gefäß mit Durchfluss verworfen. 500 µl Puffer AW2 wurden auf die QIAamp MinElute Säule gegeben und analog zum Vorgehen nach Zugabe von Puffer AW1 verfahren. Um die Membran der QIAamp MinElute Säule zu trocknen und von Ethanolresten zu befreien schloss sich ein weiterer Zentrifugationsschritt für 3 min bei 13.000 rpm an.

Zuletzt wurde die QIAamp MinElute Säule in ein 1,5 ml Reaktionsgefäß überführt, 30 µl ddH₂O auf das Zentrum der Membran gegeben und nach 5 min Inkubation für 1 min bei 13.000 rpm zentrifugiert. Um die Menge an gelöster DNA im so gewonnenen Eluat zu erhöhen, wurde dieses nach erfolgter Zentrifugation erneut auf die QIAamp MinElute Säule gegeben, nochmals 5 min inkubiert und der einminütige Zentrifugationsschritt bei 13.000 rpm wiederholt. Nun wurde spektrophotometrisch die DNA-Konzentration bestimmt.

3.2.2 Extraktion und Aufreinigung von DNA aus Zellen einer B-NHL-Zelllinie

Um die Verdünnungen von DNA der B-NHL-Zelllinie Granta-519 in PBMC-DNA herzustellen, musste die DNA zunächst aus den Lymphomzellen isoliert und aufgereinigt werden. Hierzu wurde im Folgenden das QIAamp®DNA FFPE Tissue Kit verwendet.

20 µl Proteinase K wurden in ein 1,5 ml Reaktionsgefäß gegeben und 200 µl PBS hinzugefügt, in dem sich die Lymphomzellen befanden. Nun wurden 200 µl Puffer AL dazugegeben und für 15 s gevortext. Es folgte ein Inkubationsschritt von 1 h

bei 56 °C und 500 rpm im Thermoschüttler. Das Reaktionsgefäß wurde kurz zentrifugiert, um Tropfen von der Deckelinnenseite zu entfernen. 200 µl 100%iges Ethanol wurden dem Gefäß hinzugefügt, es wurde für 15 s gevortext und erneut kurz zentrifugiert.

Nun wurde die Mixtur auf eine QIAamp Mini Spin Säule übertragen, die sich ihrerseits in einem 2 ml Reaktionsgefäß befand und bei geschlossenem Deckel für 1 min bei 8.000 rpm zentrifugiert. Die QIAamp Mini Spin Säule wurde in ein neues 2 ml Reaktionsgefäß platziert und das alte mitsamt Durchfluss verworfen. Nun wurden 500 µl Puffer AW1 auf die QIAamp Mini Spin Säule gegeben, diese erneut für 1 min bei 8.000 rpm zentrifugiert, in ein neues Reaktionsgefäß übertragen und das alte mit Durchfluss verworfen. 500 µl Puffer AW2 wurden auf die QIAamp Mini Spin Säule gegeben, für 3 min bei 13.000 rpm zentrifugiert, das alte Reaktionsgefäß mit Durchfluss verworfen. Ein zusätzlicher Zentrifugationsschritt in einem neuen 2 ml Reaktionsgefäß für 1 min bei 13.000 rpm schloss sich an, auch dieses Reaktionsgefäß wurde verworfen und die QIAamp Mini Spin Säule in ein letztes 2 ml Reaktionsgefäß platziert. Hier wurden 100 µl ddH₂O auf die Säule gegeben, für 1 min inkubiert und anschließend für 1 min bei 8.000 rpm zentrifugiert. Dieser Schritt wurde mit weiteren 100 µl ddH₂O wiederholt. Im so entstandenen Eluat wurde spektrophotometrisch die Konzentration der Lymphom-DNA bestimmt und im Anschluss mit der entsprechenden Menge ddH₂O auf 50 ng/µl verdünnt.

3.2.3 PCR zur Bestimmung der Amplifizierbarkeit der DNA

Formalin-Fixierung über längere Zeit führt zu gesteigerter Fragmentierung von Nukleinsäuren [117], so dass DNA, die aus FFPE-Gewebe isoliert wurde, häufig aus kurzen Fragmenten von nur 100 bis 300 bp besteht [118]. Dies stellt beispielsweise ein Problem für molekularpathologische Diagnoseverfahren dar, wenn ein mittels PCR zu vervielfältigender DNA-Abschnitt länger ist als die in der Probe enthaltenen DNA-Fragmente [119]. Daher ist es notwendig die aus FFPE-Gewebe gewonnene DNA zunächst auf ihre Amplifizierbarkeit, d.h. die Länge der

vorliegenden Fragmente, zu untersuchen, bevor eine Klonalitätsanalyse oder andere molekularpathologischen Diagnoseverfahren zur Anwendung kommen können.

Zu diesem Zweck wurde durch das BIOMED-2 Konsortium ein PCR-Protokoll entwickelt, in dem durch fünf Primerpaare in einem Multiplexansatz DNA-Produkte von exakt 100, 200, 300, 400 und 600 bp amplifiziert werden sollen [36]. Die Amplifizierbarkeit der verschiedenen langen DNA-Abschnitte spricht dabei für das Vorliegen entsprechend langer DNA-Fragmente in der untersuchten Probe. Im Rahmen dieser Arbeit sind die durch das BIOMED-2 Konsortium entwickelten Primerpaare in einem modifizierten PCR-Protokoll zur Anwendung gekommen (Tabelle 1).

Tabelle 1: PCR zur Bestimmung der Amplifizierbarkeit der DNA.

Reagenzien	Volumen pro Ansatz [µl]	Endkonzentration
ddH ₂ O	16,7	
10 x Puffer II	2,5	1 x
25 mM MgCl ₂	2	2 mM
dNTPs (10 mM)	0,5	0,2 mM
Primer-Mix	2	AF4/X3 2,5 µM, alle anderen 1,25 µM
AmpliTaq Gold DNA Polymerase	0,3	1,7 U/µl
DNA (50 ng/µl)	1	
	<hr/> 25	
PCR-Programm		
7 min	95 °C	} 35 Zyklen
45 s	95 °C	
45 s	60 °C	
1 min 30 s	72 °C	
10 min	72 °C	
Hold	12 °C	

Die Auftrennung der amplifizierten DNA-Abschnitte nach ihrer Länge erfolgte mittels Agarose-Gelelektrophorese, die Detektion unter UV-Licht (s. 3.4.2).

3.2.4 Herstellung von Verdünnungsproben von B-NHL-DNA in PBMC-DNA

In der ersten Phase dieses Projektes wurden Verdünnungsreihen von DNA zweier B-NHL-Zelllinien in PBMC-DNA einer gesunden Spenderin gefertigt, um Proben mit bekanntem Lymphom-DNA-Gehalt zu erhalten. Die PBMC-DNA, die bereits im Rahmen eines früheren Projektes der Arbeitsgruppe isoliert worden war, sollte eine polyklonale Hintergrundpopulation reaktiver Lymphozyten simulieren. Anhand der so erzeugten Proben sollten Parameter der Klonalität in NGS-Datensätzen von B-NHL definiert und die Nachweisgrenze der Methode ermittelt werden.

Zunächst wurden durch entsprechende Verdünnung mit ddH₂O Proben mit einer DNA-Konzentration von 50 ng/μl hergestellt und anschließend die gelöste B-NHL-DNA der Zelllinien in entsprechenden Volumina der PBMC-DNA-Lösung verdünnt. So wurde für jede Zelllinie Verdünnungen von 1:5, 1:10, 1:20, 1:50, 1:100 und 1:1.000 hergestellt.

3.3 Klonalitätsuntersuchung mittels Fragmentanalyse

3.3.1 Multiplex-PCR mit den Primern des BIOMED-2 Konsortiums

Für die Amplifikation und Detektion von klonalen IGH-Rearrangements in Patientenproben mit hohem Tumorzellgehalt wurde auf die Primersets zurückgegriffen, die durch das BIOMED-2 Konsortium zum Klonalitätsnachweis entwickelt wurden [36]. Es wurden pro Probe zwei PCR-Ansätze gefertigt, jeweils mit 50 und 200 ng Proben-DNA (Tabelle 2). Als Primer dienten das VH-Primerset für die FR2-Region sowie ein JH-Konsensus-Primer, der für die spätere Fragmentanalyse fluoreszenzmarkiert war. Die kapillarelektrophoretische Fragmentanalyse geschah mit dem GenomeLab GeXP Analyser.

Tabelle 2: PCR zur Amplifikation klonaler IGH-Rearrangements.

Reagenzien	Volumen pro Ansatz [μ l]	Endkonzentration
ddH ₂ O	19,2	
Puffer (mit 15mM MgCl ₂)	2,5	1,5 mM
dNTPs (10 mM)	0,5	0,2 mM
Primer-Mix- FR2 (10 μ M/Primer)	0,25	0,1 μ M/Primer
JH-Konsensus-Primer (10 μ M)	0,25	0,1 μ M
Ampli Taq Gold Polymerase (5 U/ μ l)	0,3	0,06 U/ μ l
DNA (50 ng/ μ l)	1 bzw. 4	
	25	
PCR-Programm		
7 min	95 °C	
45 s	95 °C	} 40 Zyklen
45 s	60 °C	
1 min 30 s	72 °C	
10 min	72 °C	
Hold	12 °C	

3.3.2 Klonalitätsanalyse am GenomeLab GeXP Analyzer

Zur Beurteilung der amplifizierten Produkte wurden die durch die Multiplex-PCR mit den Primern des BIOMED-2 Konsortiums vervielfältigten IGH-Rearrangements kapillarelektrophoretisch untersucht. Dies geschah am GenomeLab GeXP Analyzer.

Das Prinzip der Kapillarelektrophorese (GeneScan-Analyse) basiert auf der Wanderung der negativ geladenen DNA-Moleküle entlang eines elektrischen Feldes durch mit einem Polymer gefüllte Kapillaren, bei der längere DNA-Moleküle langsamer durch die Matrix wandern als kürzere [18]. Daraus resultiert eine Aufreihung der verschiedenen DNA-Fragmente entsprechend ihrer Größe. Da die in der vorangegangenen PCR verwendeten Primer eine Fluoreszenzmarkierung aufweisen, können die wandernden DNA-Fragmente durch einen Laser zur Fluoreszenz angeregt werden. Ein Detektor registriert Wellenlänge und Stärke des abgegebenen Signals.

Die Stärke des detektierten Signals, gemessen in relativen Fluoreszenzeinheiten (RFU), wird nun von der zugehörigen Software gegen die Elektrophoresedauer aufgetragen, sodass sich im resultierenden Elektropherogramm sog. Peaks

bilden, deren Höhe es erlaubt, die Menge der passierenden DNA-Fragmente abzuschätzen. In jeder Kapillarelektrophorese werden außerdem DNA-Fragmente bekannter Länge mitgeführt (GenomeLab DNA Size Standard), anhand derer sich die Länge der in der Probe enthaltenen und den Detektor passierenden DNA-Fragmente, gemessen in Nukleotiden (nt), ermitteln lässt. So entstehen im Falle der Klonalitätsanalyse die spezifischen Kurvenmuster, die beispielsweise für eine poly- oder monoklonale B-Zell-Population charakteristisch sind.

3.3.3 Kriterien zur Auswertung

Zur korrekten Interpretation der gelieferten Kurven, wurde auf die EuroClonality/BIOMED-2 Richtlinien zur Interpretation und Befundung von Ig/TCR-Klonalitätstestungen bei vermuteten lymphoproliferativen Erkrankungen zurückgegriffen [95]. Siehe hierzu Tabelle 3.

Tabelle 3: Standardisierte EuroClonality-Terminologie zur technischen Beschreibung der PCR-Ergebnisse. Verändert nach Langerak AW, et al. [95].

Profil / PCR-Ansatz (im Doppelansatz)	Technische Beschreibung / PCR-Ansatz
- kein Peak / keine Bande (aber: unzureichende DNA-Qualität)	- kein (spezifisches) PCR-Produkt, unzureichende DNA-Qualität
- kein Peak / keine Bande (mit/ohne Hintergrund)	- kein spezifisches PCR-Produkt
- 1 oder 2 reproduzierbare klonale Peaks / Banden	- klonal
- 1 oder 2 nicht-reproduzierbare (deutliche) Peaks / Banden	- pseudoklonal
- multiple (n ≥ 3) nicht-reproduzierbare (deutliche) Peaks / Banden	- pseudoklonal
- mehrere (n ≥ 3) reproduzierbare (deutliche) Peaks / Banden	- mehrere PCR-Produkte (n= ...)
- Gauß-Verteilung / Schmier (ohne/mit diskrete reproduzierbare Peaks/Banden)	- polyklonal (nicht klonal)
- nach den oben genannten Kriterien nicht klassifizierbares PCR-Profil	- nicht auswertbar

3.4 Sequenzierung klonaler IGH-Rearrangements und Charakterisierung mittels IMGT/V-Quest

3.4.1 Ermittlung der VH-Segmentfamilie und Amplifikation klonaler IGH-Rearrangements

Der Versuch, das IGH-Rearrangement des in einer Probe enthaltenen malignen Klons mittels Kettenabbruchsynthese nach Sanger zu sequenzieren, gestaltet sich oft problematisch, da ein mitamplifizierter polyklonaler Hintergrund reaktiver Lymphozyten die Elektrophoreseanalyse überlagert. Daher wurde in dieser Arbeit versucht, möglichst selektiv das IGH-Rearrangement des malignen Klons zu amplifizieren und dann zu sequenzieren.

Dazu wurde in einem ersten Schritt die unter 3.3.1 beschriebene PCR zur Klonalitätsanalyse mit VH-Primern für die FR2-Region durchgeführt, allerdings nicht in einem Multiplex-Setting, sondern für jede Probe in sieben separaten Ansätzen, die jeweils einen der sieben VH-Primer für die verschiedenen VH-Segmentfamilien sowie den fluoreszenzmarkierten JH-Konsensus-Primer enthielten. Die Ansätze wurden nun kapillarelektrophoretisch untersucht, um festzustellen, durch Zugabe welchen VH-Primers sich der klonale Peak der bereits vorhandenen Standard-Klonalitätsanalyse reproduzieren ließ. Der entsprechende VH-Primer wurde als derjenige betrachtet, der am ehesten das IGH-Rearrangement des malignen Klons amplifizierte.

Für jede Probe wurde nun eine weitere PCR mit dem passenden VH-Primer mit Fluidigm-Anhang und dem Fluidigm-JH-Konsensus-Primer durchgeführt (PCR-Protokoll siehe 3.5.1, S.41). Die Amplifikation wurde mittels Agarose-Gelelektrophorese überprüft.

3.4.2 Überprüfung der Amplifikation mittels Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese ermöglicht die Auftrennung von DNA-Fragmenten nach ihrer Größe. Die negativ geladene DNA wandert dabei entlang

eines elektrischen Feldes durch die Poren des Agarosegels. Kurze DNA-Fragmente werden weniger stark zurückgehalten als lange, sodass Fragmente verschiedener Länge nach Abschalten des elektrischen Feldes unterschiedlich weite Strecken im Gel zurückgelegt haben und dort Banden bilden. Die Basenpaarlänge der untersuchten DNA-Fragmente lässt sich abschätzen, indem man eine Referenzprobe mit DNA-Fragmenten bekannter Basenpaarlänge im Gel mitlaufen lässt. Die Sichtbarmachung der Banden gelingt durch Zugabe eines DNA-bindenden Farbstoffes, der in die DNA interkaliert und unter UV-Belichtung fluoresziert. Im Rahmen der vorliegenden Arbeit kam die Agarose-Gelelektrophorese zum Einsatz, um die Amplifikation und Länge von DNA-Abschnitten in der PCR zur Bestimmung der DNA-Qualität zu ermitteln, um die Amplifikation von IGH-Rearrangements mittels PCR zu überprüfen und um die erfolgreiche Aufreinigung von PCR-Produkten nachzuweisen.

Zur Herstellung eines 2%igen Agarosegels wurden 2 g LE Agarose in 100 ml TBE Puffer gegeben und das Gemisch im Mikrowellengerät aufgekocht. Zur späteren Sichtbarmachung der Banden wurden 7 µl GelRed hinzugegeben. Das Gel wurde in eine spezielle Vorrichtung gegossen und aushärten gelassen.

Als Laufpuffer in der Elektrophoresekammer wurde ebenfalls TBE Puffer verwendet. 10 µl Probe wurden mit 2 µl 6x DNA Loading Dye, kleinere Proben von 5 µl entsprechend mit 1 µl 6x DNA Loading Dye versetzt und in eine Tasche des Gels pipettiert. In jeder Agarose-Gelelektrophorese wurden 5 µl GeneRuler 100 bp Ladder als Größenmarker mitgeführt. Eine Spannung von 140 V wurde für 35 min angelegt, während derer die elektrophoretische Auftrennung der DNA-Fragmente erfolgte. Unter UV-Licht konnten die Banden sichtbar gemacht und mit einem Geldokumentationssystem dokumentiert werden.

3.4.3 DNA-Isolation aus Agarosegel zur Isolation der klonalen Bande

Trotz möglichst selektiver Amplifikation von IGH-Rearrangements mit dem segmentfamilien-spezifischen VH-Primer, resultierte in einigen Ansätzen ein polyklonaler Hintergrund, der die Sequenzierung nach Sanger behinderte. In

diesen Fällen wurde der Versuch unternommen, die PCR-Produkte mittels Agarose-Gelelektrophorese aufzutrennen und die klonale Bande zu isolieren. Dazu wurden die Banden im Gel unter einem UV-Schirm sichtbar gemacht und die dem klonalen Rearrangement zugeordnete Bande mit einem sauberen Skalpell herausgeschnitten. Anschließend folgte die Extraktion und Aufreinigung der DNA aus dem Agarosegel, wozu das QIAquick Gel Extraction Kit verwendet wurde.

Dazu wurde das Gel zunächst gewogen, um festzulegen, wie viel Puffer GC hinzugegeben werden musste. Pro 100 mg Gel wurden 100 µl Puffer hinzugegeben. Nun wurde bei 50 °C für 10 min inkubiert, wobei alle 3 min die Probe kurz gevortext wurde. Nachdem sich das Gelstück komplett aufgelöst hatte, wurde die Farbe der Lösung kontrolliert. Pro 100 mg Gel wurden 100 µl 100%iges Isopropanol hinzugegeben. Das Gemisch wurde nun in eine QIAquick Spin Column gegeben, die sich in einem 2 ml Reaktionsgefäß befand, und es wurde für 1 min bei 13.000 rpm zentrifugiert. Das Filtrat wurde verworfen und die QIAquick Spin Column in ein neues 2 ml Reaktionsgefäß überführt. 0,70 ml Puffer BE wurden auf die Membran der QIAquick Spin Column gegeben und 5 min inkubiert, bevor erneut für 1 min bei 13.000 rpm zentrifugiert wurde. Das Filtrat wurde verworfen und die Säule im Reaktionsgefäß für eine weitere Minute bei 13.000 rpm zentrifugiert. Schließlich wurde die QIAquick Spin Column in ein sauberes 1,5 ml Reaktionsgefäß platziert, 30 µl ddH₂O wurden auf die Membran gegeben und für 1 min inkubiert bevor für 1 min bei 13.000 rpm zentrifugiert wurde. Im so gewonnenen Eluat wurde spektrophotometrisch die DNA-Konzentration bestimmt.

3.4.4 Sequenzierung von IGH-Rearrangements nach Sanger

Zur Ermittlung der Sequenz klonaler IGH-Rearrangements wurden diese in einigen Fällen mittels Kettenabbruchsynthese sequenziert. Diese wurde 1977 von Sanger et al. beschrieben und seitdem für die breite Anwendung in der Molekularbiologie und -pathologie optimiert [99].

Als Templates für die Sequenzierreaktion dienten PCR-Produkte, die durch Amplifikation mit dem für das jeweilige IGH-Rearrangement passenden VH-Primer mit Fluidigm-Anhang erzeugt wurden. Die Bandenstärke in der darauffolgenden Agarose-Gelelektrophorese bestimmte das in der Sequenzier-PCR eingesetzte Probenvolumen. Bei Vorliegen eines starken polyklonalen Hintergrundes folgte der gelelektrophoretischen Auftrennung der DNA-Fragmente eine Extraktion der zu sequenzierenden Bande aus dem Agarosegel. Vor Einsetzen in die Sequenzierreaktion mussten die PCR-Produkte zudem mittels Agencourt® AMPure® XP aufgereinigt werden (siehe unten).

Durch die Amplifikation der IGH-Rearrangements mit Primern mit einem universellem Sequenzanhang (Fluidigm), war jedes Amplikon an beiden Enden mit einer zusätzlichen Sequenz versehen, die als Zielsequenz für die Sequenzierprimer diente. In der Sequenzierreaktion (siehe Tabelle 4) wurden nun ein Fluidigm-Primer eingesetzt, der an diese Sequenz am 3'-Ende eines Templates band und von dem aus die Kettenabbruchsynthese startete.

Tabelle 4: PCR zur Kettenabbruchsynthese.

Reagenzien	Volumen pro Ansatz [µl]
Proben-DNA	je nach Bandenstärke 3 – 7
DTCS-Mix	2
Fluidigm-Primer F/R (10 µM)	1
ddH ₂ O	ergänzend 0 – 4
	10
PCR-Programm	
20 s	96 °C
20 s	50 °C
4 min	60 °C
Hold	12 °C

Dem Reaktionsansatz waren neben Desoxynukleotidtriphosphaten (dNTPs) auch vier verschiedene Didesoxynukleosidtriphosphate (ddNTPs) zugesetzt, bei deren Einbau kein weiteres Nukleotid der wachsenden Kette hinzugefügt werden konnte und es zum Kettenabbruch kam. Diese ddNTPs wiesen zusätzlich verschiedenfarbige Fluoreszenzmarkierungen auf. Pro Probe wurden zwei PCR-Ansätze mit verschiedenen Primern (Fluidigm F und R) gefertigt, die jeweils

komplementär zum 3'-Enden eines der beiden DNA-Einzelstränge waren. So erfolgte die Kettenabbruchsynthese sowohl in „Forward“- als auch in „Reverse“-Richtung.

Nach der PCR mussten überschüssige dNTPs und ddNTPs durch Aufreinigung mit Agencourt® CleanSeq® entfernt werden (siehe unten). Nun konnte am GenomeLab GeXP Analyzer die kapillarelektrophoretische Auftrennung der verschieden langen PCR-Produkte erfolgen. Das Prinzip der Kapillarelektrophorese wurde bereits unter 3.3.2 umrissen. Dank der verschiedenfarbigen Fluoreszenzmarkierungen der ddNTPs, konnte zusätzlich ermittelt werden, welches ddNTP zum Kettenabbruch geführt hatte. Aus dem daraus resultierenden Elektropherogramm, in dem die Basenpaarlänge jedes Fragments gegen die Farbe und Stärke des korrespondierenden Fluoreszenzsignals aufgetragen war, konnte mit der zugehörigen Software die Sequenz des untersuchten Amplikons festgestellt werden. Mit dem Alignment-Tool BLAST wurden „Forward“- und „Reverse“-Sequenz eines IGH-Rearrangements miteinander verglichen. Bei nicht übereinstimmenden Basen wurde das Elektropherogramm an der entsprechenden Position kontrolliert und die Sequenz gegebenenfalls manuell korrigiert.

Aufreinigung von PCR-Produkten mit Agencourt® AMPure® XP

Das Agencourt® AMPure® XP Reinigungssystem ermöglicht eine Aufreinigung von PCR-Produkten nach dem Prinzip der SPRI (solide-phase reverse immobilization). Dieses basiert auf der Bindung von DNA-Fragmenten größer 100 bp an die Carboxylbeschichtung paramagnetischer Partikel, die mittels einer Magnetplatte im Reaktionsgefäß fixiert werden, während in mehreren Waschschritten Nebenprodukte der PCR entfernt werden [120, 121]. Hierzu zählen überschüssige Primer, Nukleotide, Salze und Enzyme, die die weitere Analyse der amplifizierten DNA-Abschnitte behindern.

Als erster Schritt wurde dem PCR-Ansatz in einem 0,2 ml Reaktionsgefäß das 1,6-Fache an Volumen des AMPure® Reagenz hinzugefügt, gevortext und 5 min inkubiert. Nun wurde das Reaktionsgefäß in die Magnetplatte gestellt, sodass die

paramagnetischen Partikel mit den gebundenen PCR-Produkten an die Wand des Reaktionsgefäßes wanderten und dort fixiert wurden. Der Überstand konnte nun abpipettiert und verworfen werden. Es folgten zwei Reinigungsschritte mit jeweils 200 µl 70%igem Ethanol, indem dieses hinzugegeben und nach 3 min Inkubation abpipettiert wurde. Auch kleine Reste Ethanol wurden entfernt und die Beads anschließend trocknen gelassen. Nun folgte die Resuspension der DNA-Fragmente in ddH₂O, wobei bei großer Bandenstärke des erzielten PCR-Produktes in der vorangegangenen Gelelektrophorese 40 µl ddH₂O, bei schwächerer Bandenstärke 30 µl ddH₂O verwendet wurden.

Aufreinigung von PCR-Produkten mit Agencourt® CleanSeq®

Die Aufreinigung der PCR-Produkte aus der Kettenabbruchsynthese erfolgte nach dem bereits oben beschriebenen SPRI-Prinzip und diente der Entfernung von überschüssigen Desoxy- und Didesoxynukleotiden.

Im 0,2 ml Reaktionsgefäß wurden zu 10 µl PCR-Produkt 10 µl CleanSeq®-Reagenz sowie 42 µl 85%iges Ethanol gegeben, gevortext und 5 min inkubiert. Durch Einsetzen des Reaktionsgefäßes in die Magnetplatte und Inkubation für 3 min wanderten die paramagnetischen Beads an die Wand des Reaktionsgefäßes und waren dort für den folgenden Reinigungsschritt fixiert. Der Überstand wurde entfernt und 100 µl 85%iges Ethanol hinzugegeben. Dieses wurde entfernt, es wurden zügig 40 µl SLS zugegeben und die Produkte der Kettenabbruchsynthese darin eluiert. Durch entsprechende Positionierung des Reaktionsgefäßes auf der Magnetplatte bildeten die paramagnetischen Partikel am Boden ein Pellet.

3.4.5 Analyse des klonalen IGH-Rearrangements mittels IMGT/V-Quest

Das Online-Tool IMGT/V-Quest ist ein Werkzeug für die Analyse von rearrangierten Immunglobulin- und T-Zell-Rezeptor-Sequenzen. Es identifiziert die V-, D- und J-Gene und -Allele mittels Alignment mit den Keimbahn Ig- und TCR-Genen im IMGT-Referenzverzeichnis [122]. Bei gleichem

Übereinstimmungsgrad eines V-, D-, oder J-Gensegmentes mit mehreren Genen bzw. Allelen des Referenzverzeichnisses, wird jedes dieser möglichen Gene und Allele in der Ausgabe „Gene and Allele“ angegeben. Der prozentuale Wert der Übereinstimmung einer IGH-Sequenz mit der Sequenz des Referenzverzeichnisses wird für V- und J-Gen und –Allel angegeben. Das Tool liefert auch eine Beschreibung der CDR3, der Mutationen in der V-Region und detektiert Insertionen und Deletionen (Abbildung 7).

Im Rahmen dieser Arbeit wurde IMGT/V-Quest genutzt, um mittels Sanger-Sequenzierung gewonnene Sequenzen von klonalen IGH-Rearrangements zu analysieren und diese später im NGS-Datensatz identifizieren zu können.

```
>IGH_DLBC1
gggctccagggaaggggctggagtgggtggcccaacataaatcaagatggaagtgagagaa
aatatgtggactctgtgaagggccgattcaccatctccagagacaacgccaacaattcac
tgtatctacaaatgaacagcctgagagccgaggacacggccgtgtattatgtgogagag
accogagaaaggactggggatcagagcacaactacttcgacccctggggccaggagccc
tggtcaccgtctctcctcaggttaagagaccaagtctctgctaccgtaaa
```

Result summary:	Productive IGH rearranged sequence (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV3-7*03 F	score = 779	identity = 94.80% (164/173 nt)
J-GENE and allele	Homsap IGHJ5*02 F	score = 228	identity = 94.12% (48/51 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHJ5*02 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[X.11.38.11]	[X.8.17]	CARDPRKDWGSEHNYFDPW

Abbildung 7: Beispiel der Analyse der IGH-Sequenz eines DLBCL mittels IMGT/V-Quest, Ergebniszusammenfassung. Die analysierte Sequenz weist kein Stop-Codon auf („productive IGH rearranged sequence“). Es werden das V-, D- und J-Gen und –Allel der Keimbahn mit der größten Ähnlichkeit zum analysierten Rearrangement aufgeführt. Als Maß dieser Übereinstimmung dienen der prozentuale „Identity“-Wert und der jeweilige Alignment-Score. Letzterer errechnet sich durch Addition von fünf pro übereinstimmendem und Subtraktion von vier pro von der Referenzsequenz abweichendem Nukleotid. Außerdem aufgeführt sind die Längen der Framework- und CD-Regionen (FR und CDR), sowie die Aminosäuresequenz der Verbindungsregion (AA JUNCTION).

3.5 Next Generation Sequencing des amplifizierten IGH-Repertoires

3.5.1 Multiplex-PCR mit VH-Primern mit Fluidigm-Anhang

Um IGH-Rearrangements zu amplifizieren und das amplifizierte Repertoire anschließend sequenzieren zu können, wurden spezielle Fluidigm-PCR-Primer verwendet. Diese basierten auf den durch das BIOMED-2 Konsortium

entwickelten Primern zur Klonalitätsanalyse (siehe 3.3.1) und waren mit einem zusätzlichen Anhang versehen, der es ermöglichte, dass in einer zweiten PCR die Primer für die Kettenabbruchsynthese (siehe 3.4.4) bzw. die Barcode-Primer für den Next Generation Sequencing Prozess an die Amplikons binden konnten. Jede Probe wurde mit zwei verschiedenen Primersets amplifiziert; mit dem einen erfolgte dies mit Forward-Primern, die an die FR1-Region hybridisierten, während mit dem anderen die Forward-Primer an die FR2-Region anlagerten. In beiden Reaktionen kam außerdem ein JH-Konsensus-Primer zum Einsatz.

Für die Amplifikation der B-NHL-Zelllinienverdünnungs- und Patientenproben konnte auf das gleiche PCR-Protokoll wie unter 3.3.1 zurückgegriffen werden (Tabelle 5).

Tabelle 5: PCR zur IGH-Amplifikation in Zelllinienverdünnungs- und Patientenproben.

Reagenzien	Volumen pro Ansatz [µl]	Endkonzentration
ddH ₂ O	19,2	
Puffer (mit 15mM MgCl ₂)	2,5	1,5 mM
dNTPs (10 mM)	0,5	0,2 mM
Primer-Mix-Flui-FR1/FR2 (10 µM/Primer)	0,25	0,1 µM/Primer
JH-Konsensus-Flui-Primer (10 µM)	0,25	0,1 µM
Ampli Taq Gold Polymerase (5 U/µl)	0,3	0,06 U/µl
DNA (50 ng/µl)	1	
	<hr/>	
	25	
PCR-Programm		
7 min	95 °C	
45 s	95 °C	} 40 Zyklen
45 s	60 °C	
1 min 30 s	72 °C	
10 min	72 °C	
Hold	12 °C	

Die Amplifikation wurde mittels Agarose-Gelelektrophorese überprüft, und die PCR-Produkte mit Agencourt® AMPure® XP aufgereinigt.

3.5.2 Barcode-PCR und Proben-Pooling

In einer zweiten PCR (Tabelle 6) wurde für jeden Ansatz ein anders Primerpaar (Barcode-Primer, Fluidigm® Access Array™ - Access Array Barcode Library for Illumina Sequencers - 384) verwendet, um jedes Amplikon mit einer für den Ansatz einzigartigen Adaptersequenz zu versehen. Diese diente als „Barcode“, anhand dessen die später im massiv parallelen Sequenzierungsprozess ermittelten Sequenzen wieder eindeutig dem jeweiligen Ansatz zugeordnet werden konnten.

Tabelle 6: Barcode-PCR. Für jeden Ansatz wurde jeweils ein anderer Barcode-Primer verwendet. Die Verdünnung des Amplifikats war abhängig von der Bandenstärke des jeweiligen Amplifikats in der vorangegangenen Gelelektrophorese und betrug maximal 1:10.

Reagenzien	Volumen pro Ansatz [µl]	Endkonzentration
ddH ₂ O	14	
Fast Start High Fidelity Reaction Buffer (mit 18 mM MgCl ₂)	2	1x/1,8 mM MgCl ₂
DMSO	0,4	
dNTPs (10 mM)	0,4	0,2 mM
Fast Start High Fidelity Enzyme Blend	0,2	
Barcode-Primer (2 µM)	2	0,2 µM
Amplifikat (Verdünnung)	1	
	20	
PCR-Programm		
10 min	95 °C	
15 s	95 °C	} 10 Zyklen
30 s	60 °C	
1 min	72 °C	
3 min	72 °C	
Hold	12 °C	

Die Amplifikation wurde mittels Agarose-Gelelektrophorese überprüft und der Ansatz mittels Agencourt® AMPure® XP aufgereinigt. Die Bestimmung der DNA-Konzentration erfolgte spektrophotometrisch.

Nun wurden die Ansätze für den NGS-Prozess in einem Pool gesammelt. Im Laufe dieser Arbeit wurden drei Pools gefertigt, einer mit den Ansätzen der Zelllinienverdünnungsproben, einer mit den Ansätzen der Patientenproben mit hohem Lymphomzellgehalt und ein dritter Pool mit den Ansätzen der Verlaufs- und der Glaskörperproben. In Abhängigkeit des Endvolumens des Pools und der

festgelegten DNA-Konzentration wurden Volumina jedes Ansatzes kalkuliert, die in einem 1,5 ml Low-Binding-Tube in ATE Elusion Buffer gegeben wurden. Jeder Pool hatte ein Endvolumen von 300 µl und eine DNA-Konzentration von 8 nmol/l pro Ansatz. Die Probenpools wurden zur Sequenzierung auf der Illumina MiSeq-Plattform an das Institut für Medizinische Genetik und angewandte Genomik der Universität Tübingen übergeben.

3.5.3 Sequenzierung auf der Illumina MiSeq-Plattform

Der Sequenzierung am Illumina MiSeq liegt das Prinzip des „sequencing by synthesis“ (Sequenzierung durch Synthese) zugrunde: Base für Base wird für jedes Template der komplementäre Strang synthetisiert und die jeweils eingebauten Basen über ein Fluoreszenzsignal bestimmt. Da ein einzelnes eingebautes Molekül allerdings kein Signal von detektierbarer Stärke erzeugt, geht dem Sequenzierungsprozess die Generierung sogenannter Cluster voraus, lokale Ansammlungen klonal amplifizierter Templates auf der Flow-Cell [123].

Als erster Schritt der Cluster-Generierung binden die zu sequenzierenden Einzelstrang-DNA-Templates mit einem Ende an die Oberfläche der Flow-Cell. Diese Immobilisation wird vermittelt durch eine dichte Beschichtung der Oberfläche mit Forward- und Reverse-Primern, die komplementär zu den Adaptersequenzen sind, mit denen die Templates in der vorausgegangenen PCR versehen wurden [102]. Das jeweils freie Ende kann nun ebenfalls an den komplementären Primer an der Oberfläche hybridisieren und so eine „Brücke“ bilden, die folgende PCR-Amplifikation wird daher auch als „Brückenamplifikation“ bezeichnet [124]. So entsteht pro Template eine Kolonie von ca. einer Million klonalen Amplifikaten, ein sogenanntes Cluster [100].

Nun beginnt der eigentliche Sequenzierungsprozess, indem die DNA-Polymerase und die vier nötigen Nukleotide gleichzeitig auf die Oberfläche der Flow-Cell gegeben werden. Die Nukleotide tragen eine basenspezifische Fluoreszenzmarkierung und sind außerdem an der 3'OH-Gruppe chemisch blockiert [100], sodass jeder Inkorporationsschritt einzeln stattfindet. Jedem Einbausschritt folgt ein Bildgebungsschritt, in dem für jedes Cluster die eingebaute

Base detektiert wird. Nun wird die chemische Blockade an der 3'OH-Gruppe der eingebauten Base, sowie deren Fluoreszenzmarkierung entfernt und die nächste fluoreszenzmarkierte, 3'OH-blockierte Base kann in den wachsenden komplementären Strang eingebaut werden. Dieser Vorgang wird für eine benutzerdefinierte Anzahl von Zyklen wiederholt. Ein Base-Calling-Algorithmus ordnet Sequenzen zu, versieht jeden Read mit Qualitätswerten und die Illumina-Daten jedes Laufs werden bewertet, um Sequenzen von schlechter Qualität zu verwerfen [101].

Nach Vollendung des ersten Reads, kann auch die Reverse-Sequenz ermittelt werden. Hierzu werden zunächst die eben sequenzierten DNA-Stränge enzymatisch abgetrennt und die komplementären Stränge durchlaufen die Brücken-Amplifikation, um Cluster zu formen. Der Sequenzierungsprozess kann nun von neuem ablaufen [101].

Jeder Read lieferte etwa 180 bp eines jeden IGH-Rearrangements, sowohl in Forward- als auch in Reverse-Richtung. Dieses Vorgehen, auch Paired-End Sequencing genannt, ermöglichte es, IGH-Rearrangements in einem Durchlauf in kompletter Länge zu sequenzieren, indem durch ein In-House-Tool der Medizinischen Genetik die zueinander gehörenden Forward- und Reverse-Sequenzen zu einer Gesamtsequenz zusammengefügt wurden. Es resultierten Ausgabedateien im FASTA-Format, welches Nukleotidsequenzen buchstabenkodiert repräsentiert.

3.5.4 Analyse des amplifizierten IGH-Repertoires mittels IMGT/HighV-Quest

Die mittels NGS gewonnen Datensätze der einzelnen Proben wurden mit IMGT/HighV-Quest analysiert, einem Online-Tool, das es erlaubt analog zur Analyse einzelner rearrangierter Ig- und TCR -Gene durch IMGT/V-Quest, NGS-Datensätze mit großen Zahlen an Sequenzen zu analysieren und charakterisieren [114]. Die Funktionsweise und Handhabung des Tools sollen im

Folgenden mit Fokus auf die in dieser Arbeit genutzten Funktionen beschrieben werden.

Im IMGT/HighV-Quest-Suchfenster wird für jede Analyse zunächst ein Titel festgelegt sowie Rezeptortyp und die Herkunftsspezies (in dieser Arbeit IGH, Mensch) angegeben. Nun wird der zu analysierende NGS-Datensatz im FASTA-Format ausgewählt, der bis zu 150.000 Sequenzen umfassen darf. Durch Wählen von „Start“ beginnt die Analyse mit Standardeinstellungen, unter der Rubrik „Display Results“ sowie „Advanced Parameters“ lassen sich die Parameter der Analyse sowie die Zusammenstellung an Ergebnisdateien individuell weiter anpassen, was für die Fragestellungen dieser Arbeit allerdings nicht nötig war. Beim Hochladen der FASTA-Datei wird eine Syntaxanalyse durchgeführt, um den Nutzer auf syntaktisch inkorrekte Sequenzen aufmerksam zu machen [114]. Werden alle Sequenzen akzeptiert, gelangt der hochgeladene Datensatz in eine Warteschleife, bis Rechnerkapazitäten zu Analyse zur Verfügung stehen. Je nach getroffener Auswahl erfolgt eine Benachrichtigung per Mail, sobald die Analyse abgeschlossen ist.

Auf Grundlage der genauen Charakterisierung aller in einem Datensatz enthaltenen IGH-Sequenzen nach dem gleichen Algorithmus wie er in IMGT/V-Quest Anwendung findet, liefert IMGT/HighV-Quest in einem nächsten Schritt eine statistische Auswertung [125]. Diese bietet auch die Möglichkeit, mehrere Datensätze zusammenzulegen und übergreifend statistisch zu analysieren, eine Funktion, für die in dieser Arbeit keine Verwendung bestand. Im Folgenden wird nur auf jene Funktionen und Ausgabedateien eingegangen, die für die Fragestellung relevante Ergebnisse lieferten.

Zentral für die Interpretation der untersuchten Datensätze hinsichtlich Klonalität war die Charakterisierung sogenannter „IMGT-Clonotypes (AA)“ innerhalb eines Datensatzes und die Anzahl von IGH-Sequenzen die einem solchen Clonotype (AA) jeweils zugeordnet wurden. Durch IMGT ist ein Clonotype (AA) definiert als eine einzigartige Kombination eines V(D)J-Rearrangements (bestimmt auf Nukleotidebene [114]) mit einer CDR3-IMGT-Verbindungssequenz („sequence of junction“, bestimmt auf Aminosäureebene („amino acid“, AA)) [125, 126]

innerhalb eines Datensatzes. Während also die Unterscheidung der Clonotypes (AA) im Bereich der CDR3 zunächst nur auf AA-Ebene erfolgt, wird in den restlichen Abschnitten der Rearrangements, wie dem V- und dem J-Segment, auf Nukleotidebene unterschieden. Jedem Clonotype (AA) werden alle passenden Sequenzen innerhalb des Datensatzes zugeordnet, die entweder der Kategorie „1 copy“ oder „more than 1“ angehören. Zu „1 copy“ zählt jede erste individuelle Sequenz, die im Detail analysiert wird. Stößt das Programm im Analyseprozess auf eine Sequenz, die exakt mit einer bereits charakterisierten Sequenz übereinstimmt, wird diese als „more than 1“ dem gleichen Clonotype (AA) wie das zuvor unter „1 copy“ charakterisierte Ebenbild zugeordnet. Die Summe aller Sequenzen dieser beiden Kategorien ergibt die Gesamtzahl aller Sequenzen („total“), die innerhalb eines Datensatzes einem bestimmten Clonotype (AA) zugeordnet wurden.

Auch in der Definition der Clonotypes (AA) werden bei gleichem Übereinstimmungsgrad eines V-, D-, oder J-Gensegmentes mit mehreren Genen bzw. Allelen des Referenzverzeichnisses alle dieser möglichen Gene und Allele angegeben. Der Übereinstimmungsgrad der V-Region wird als „V-Region Identity“ (V%) pro Clonotype (AA) in Prozent angegeben.

In der Ausgabedatei „IMGT clonotypes (AA) per Nb“ (Abbildung 8) sind die Clonotypes (AA) des jeweiligen Datensatzes geordnet nach der Gesamtzahl ihnen zugeordneter Sequenzen (nach Sequenzstärke). Sie war daher am besten dazu geeignet, eine Aussage über die Repräsentation und den Anteil eines möglichen klonalen IGH-Rearrangements innerhalb eines mittels PCR amplifizierten IGH-Repertoires einer Lymphomprobe zu treffen. Es war zu vermuten, dass ein auf Rang 1 stehender Clonotype (AA), dem ein hoher Anteil aller IGH-Sequenzen eines Datensatzes zugeordnet worden waren, das IGH-Rearrangement einer klonal expandierten Lymphozytenpopulation charakterisiert.

Bsp. CLL IGH IMGT clonotypes (AA) per Nb

Nb. of sequences assigned to a IMGT clonotype (AA) : 27102

Analysis date: Tue Sep 15 22:05:42 2015

ID		Nb			IMGT clonotype (AA) definition						IMG
#	Exp. ID	Total nb of '1 copy'	Total nb of 'More than 1'	Total	V gene and allele	D gene and allele	J gene and allele	CDR3-IMGT length (AA)	CDR3-IMGT sequence (AA)	Anchors 104,118	V %
1	338-030934	6543	8566	15109	Homsap IGHV3-72*01 F	Homsap IGHD2-2*01 F, or Homsap IGHD2-2*02 F, or Homsap IGHD3-9*01 F, or Homsap IGHD4-23*01 ORF	Homsap IGHJ6*02	20 AA	VRAYCVTTSCSKSRYPHGMVDV	C,W	97.2
2	1383-030934	582	826	1408	Homsap IGHV1-2*02 F	Homsap IGHD1-1*01 F, or Homsap IGHD1-14*01 ORF, or Homsap IGHD3-3*01 F, or Homsap IGHD4-17*01 F, or Homsap IGHD5-18*01 F, or Homsap IGHD6-6*01 F	Homsap IGHJ3*02	9 AA	ASTYDAFDI	C,W	100
3	621-030934	550	587	1137	Homsap IGHV1-2*02 F	Homsap IGHD2-15*01 F, Homsap IGHD1-26*01 F,	Homsap IGHJ5*02	18 AA	ATTPLQCSSGGICYFPMDP	C,W	98.9

Abbildung 8: Ausschnitt aus der Ausgabedatei „IMGT clonotypes (AA) per Nb“ der statistischen Analyse. Jede Tabellenzeile beinhaltet einen Clonotype (AA). Diese sind nach absteigender Anzahl ihnen zugeordneter Sequenzen („Total“) geordnet und mit einer ID versehen (in diesem Ausschnitt nur die drei sequenzstärksten abgebildet). In der Spalte „IMGT clonotype (AA) definition“ ist das V(D)J-Rearrangement charakterisiert sowie die CDR3 des jeweiligen Clonotypes (AA). Die D-Gensegmente der drei dargestellten Clonotypes (AA) wiesen gleiche Übereinstimmungsgrade zu mehreren Genen bzw. Allelen des Referenzverzeichnisses auf, die alle in der Definition des jeweiligen Clonotypes (AA) aufgelistet sind. „V%“ gibt analog zu IMGT/V-Quest die V-Region-Identity der für den jeweiligen Clonotype (AA) repräsentativen IGH-Sequenz an. Neben weiteren Parametern wird auch die Nukleotidlänge der repräsentativen Sequenz angegeben (hier außerhalb der Darstellung).

Da die Definition eines Clonotypes (AA) im Bereich der CDR3 lediglich auf AA-Ebene erfolgt, können innerhalb der Sequenzen, die demselben Clonotypes (AA) zugeordnet werden, in der CDR3 Unterschiede auf Nukleotidebene bestehen, solange diese nicht die Aminosäuresequenz verändern. Diese innerhalb der Sequenzen eines Clonotypes (AA) unterscheidbaren Varianten, werden als Clonotypes (nt) bezeichnet. Einen Überblick über alle einem Clonotype (AA) untergeordneten Clonotypes (nt) bietet die Ausgabedatei „IMGT clonotypes (AA) per Nb with detailed clonotypes (nt)“ (Nukleotid, nt). Ihr ist neben der Anzahl der den jeweiligen Clonotypes (nt) zugeordneter Sequenzen auch die genaue Nukleotidsequenz der Clonotypes (nt) zu entnehmen, sowie die Anzahl an

Nukleotiden der CDR3, die von der Referenzsequenz des Clonotypes (AA) abweichen. Der Ausgabedatei „IMGT clonotypes (AA) per Nb with detailed clonotypes (nt)“ wurden Informationen über Varianten innerhalb von Rearrangement-Populationen entnommen, die durch SHM, oder als Artefakte der Fixierung, des Amplifikations- oder Sequenzvorganges entstanden sein konnten.

In der Ausgabedatei „Result summary for batches“ sind für jeden innerhalb einer statistischen Analyse untersuchten Datensatz (in der vorliegenden Arbeit jeweils nur einer) verschiedene zusammenfassende Parameter aufgeführt. Hierzu zählen unter anderem die Gesamtzahl der analysierten Sequenzen, die Gesamtzahl der Sequenzen, die einem Clonotype (AA) zugeordnet wurden (d.h. die als IGH-Sequenz erkannt wurden) sowie die Anzahl verschiedener erkannter Clonotypes (AA) innerhalb eines Datensatzes. Diese Maße lieferten Hinweise auf die Effektivität der Multiplex-IGH-PCR, im Hinblick auf die Menge und Qualität der erzeugten IGH-Sequenzen, und die Vielfältigkeit des amplifizierten IGH-Repertoires.

4. Ergebnisse

4.1 Beschreibung des Probenkollektivs

4.1.1 B-NHL-Zelllinien

Um Parameter der Klonalität in NGS-Datensätzen von B-NHL zu definieren und die Nachweisgrenze der Methode zu ermitteln, wurden in der ersten Phase dieses Projektes Verdünnungsreihen von DNA zweier B-NHL-Zelllinien in PBMC-DNA gefertigt und mittels NGS untersucht. Hierzu wurden die Mantelzell-Lymphomzelllinie Granta-519 und die ABC DLBCL Zelllinie OCI-LY3 genutzt.

Wie unter 3.2.4 beschrieben wurde aus der DNA der B-NHL-Zelllinien Verdünnungsproben in PBMC-DNA hergestellt (Tabelle 7).

Tabelle 7: Übersicht über Verdünnungsproben, die aus B-NHL-Zelllinien-DNA durch Verdünnung in PBMC-DNA hergestellt wurden.

Zelllinie Granta-519			Zelllinie OCI-LY3		
Proben-Nr.	Entität	Verdünnung	Proben-Nr.	Entität	Verdünnung
S01	Mantelzell-Lymphom	1:5	S07	ABC DLBCL	1:5
S02		1:10	S08		1:10
S03		1:20	S09		1:20
S04		1:50	S10		1:50
S05		1:100	S11		1:100
S06		1:1.000	S12		1:1.000

4.1.2 Zusammenstellung der verwendeten Patientenproben

Zur Rekrutierung der Patientenproben wurde das Archiv des Instituts für Pathologie und Neuropathologie des Universitätsklinikums Tübingen durchgegangen und passende Fälle aus den Jahren 2003 bis 2014 ausgewählt. Ein Ethikvotum der Medizinischen Fakultät Tübingen hierzu liegt vor (Aktenzeichen 069/2015BO2).

Für die verschiedenen Fragestellungen zum Einsatz des NGS in der Klonalitätsuntersuchung wurden drei Unterkollektive gebildet, die jeweils andere Charakteristika aufwiesen.

Zunächst sollten Patientenproben mit hohem Tumorzellgehalt untersucht werden. Hierzu wurden 15 Fälle ausgewählt, die histologisch und immunhistochemisch als B-NHL eingeordnet worden waren und die mikroskopisch einen hohen Anteil der Lymphomzellen aufwiesen. Es wurden fünf Fälle einer B-CLL, FL und fünf ABC DLBCL ausgewählt. Innerhalb des Unterkollektivs sollten Proben aus Lymphknoten (LK), Knochenmark (KM) und aus einer extranodalen Lokalisation repräsentiert sein.

In einem zweiten Unterkollektiv wurden Verlaufspröben von Patienten, die im Zeitraum ihrer Erkrankung mehrmals bioptisch untersucht worden waren, zusammengestellt. Es wurden sechs B-NHL-Patienten ausgewählt, von denen jeweils drei Proben von unterschiedlichen Zeitpunkten mittels NGS untersucht wurden. Hier wurde ein Patient mit einer B-CLL, drei Patienten mit einem FL und zwei Patienten mit einem DLBCL (eines nicht näher charakterisiert, ein ABC DLBCL) ausgewählt. Auch in diesem Unterkollektiv waren LK-, KM- und Proben aus extranodalen Lokalisationen vertreten.

Das dritte Unterkollektiv setzte sich aus drei Glaskörperaspiraten (GK) zusammen, die bei Vorliegen eines intraokulären Lymphoms gewonnen worden waren. Diese GK-Proben wurden anhand ihres Bandenmusters in der GeneScan-Analyse ausgewählt, die aus der molekularpathologischen Diagnostik vorlagen, sodass eine Probe mit monoklonalem Muster, eine mit pseudoklonalem und eine mit oligoklonalem Bandenmuster in die Studie eingeschlossen wurden.

Insgesamt wurden also 36 Lymphomproben von 24 Patienten im Rahmen dieser Studie untersucht (Tabelle 8).

Ergebnisse

Tabelle 8: Übersicht über die in dieser Studie untersuchten Patientenproben. PE: Probeexzision.

Nr.	Diagnose	Patient			Material
Proben mit hohem Lymphomzellgehalt					
S13	B-CLL	Patient 1,	männlich,	73 Jahre	KM
S14	B-CLL	Patient 2,	weiblich,	78 Jahre	KM
S15	B-CLL	Patient 3,	weiblich,	74 Jahre	LK
S16	B-CLL	Patient 4,	männlich,	72 Jahre	Glandula parotidea
S17	B-CLL	Patient 5,	männlich,	75 Jahre	LK
S18	FL	Patient 6,	weiblich,	70 Jahre	LK
S19	FL	Patient 7,	weiblich,	74 Jahre	LK
S20	FL	Patient 8,	weiblich,	73 Jahre	LK
S21	FL	Patient 9,	weiblich,	54 Jahre	LK
S22	FL	Patient 10,	weiblich,	59 Jahre	KM
S23	ABC DLBCL	Patient 11,	weiblich,	72 Jahre	LK
S24	ABC DLBCL	Patient 12,	männlich,	56 Jahre	LK
S25	ABC DLBCL	Patient 13,	weiblich,	67 Jahre	LK
S26	ABC DLBCL	Patient 14,	männlich,	71 Jahre	LK
S27	ABC DLBCL	Patient 15,	männlich,	44 Jahre	LK
Verlaufspröben					
S28	NHL	Patient 16,	männlich,	43 Jahre, Zeitpunkt 1	Hautbiopsie
S29	DLBCL		32 Monate später,	46 Jahre, Zeitpunkt 2	Hodenbiopsie
S30	DLBCL		7 Monate später,	47 Jahre, Zeitpunkt 3	Hautbiopsie
S31	GC DLBCL	Patient 17,	weiblich,	61 Jahre, Zeitpunkt 1	LK
S32	GC DLBCL		34 Monate später,	64 Jahre, Zeitpunkt 2	Milz
S33	GC DLBCL		44 Monate später	67 Jahre, Zeitpunkt 3	LK
S34	CLL	Patient 18,	weiblich,	59 Jahre, Zeitpunkt 1	KM
S35	CLL		10 Monate später,	59 Jahre, Zeitpunkt 2	KM
S36	CLL		selber Monat,	59 Jahre, Zeitpunkt 3	KM
S37	FL	Patient 19,	weiblich,	55 Jahre, Zeitpunkt 1	KM
S38	FL		22 Monate später,	57 Jahre, Zeitpunkt 2	KM
S39	FL		21 Monate später,	58 Jahre, Zeitpunkt 3	KM
S40	FL	Patient 20,	weiblich,	37 Jahre, Zeitpunkt 1	LK
S41	FL		54 Monate später,	41 Jahre, Zeitpunkt 2	LK
S42	FL		33 Monate später,	44 Jahre, Zeitpunkt 3	LK
S43	FL	Patient 21,	männlich,	63 Jahre, Zeitpunkt 1	PE retroaurikulär
S44	FL		16 Monate später,	64 Jahre, Zeitpunkt 2	Hautbiopsie
S45	FL		85 Monate später,	71 Jahre, Zeitpunkt 3	KM
Glaskörperaspirate					
S46	B-NHL	Patient 22,	weiblich,	65 Jahre	GK
S47	B-NHL	Patient 23,	männlich,	73 Jahre	GK
S48	B-NHL	Patient 24,	weiblich,	71 Jahre	GK

4.2 Amplifizierbarkeit der DNA

4.2.1 Amplifizierbarkeit der DNA aus Formalin-fixiertem Paraffin-eingebettetem Gewebe

DNA, die aus Formalin-fixiertem Paraffin-eingebettetem Gewebe isoliert wurde, wurde einer PCR zur Überprüfung der DNA-Amplifizierbarkeit unterzogen (siehe 3.2.3). Die Basenpaarlänge des größten erzeugten Amplifikats gab dabei Aufschluss über den Fragmentierungsgrad der in der Probe enthaltenen DNA. Dabei gilt zu beachten, dass für eine erfolgreiche Amplifikation von IGH-Rearrangements eine Fragmentlänge von mindestens 300 bp vorliegen sollte. Insgesamt 33 Proben waren aus FFPE-Gewebe extrahiert worden und wurden einer PCR zur Überprüfung der DNA-Amplifizierbarkeit unterzogen (Abbildung 9).

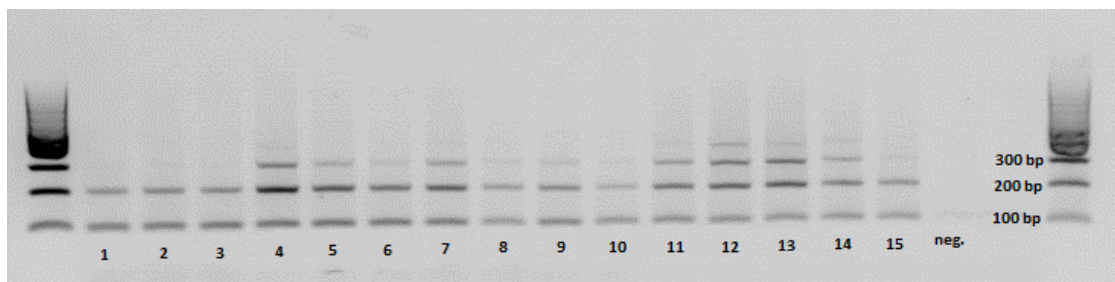


Abbildung 9: Exemplarische Darstellung der gelelektrophoretischen Auftrennung der Produkte der PCR zur Überprüfung der DNA-Qualität von Patientenproben mit hohem Lymphomzellgehalt (Ansätze 1 bis 15). In der linken und rechten äußeren Geltasche wurde ein Größenmarker mitgeführt. Der Vergleich der Banden der Patientenproben mit denen des Größenmarkers erlaubt die Größenermittlung der amplifizierbaren Produkte in den Patientenproben und damit einen Rückschluss auf die DNA-Qualität. neg.: Negativkontrolle.

In acht der 33 Proben ließen sich Produkte von über 300 bp erzeugen, in 13 Proben bis 300 bp und in zwölf Proben bis 200 bp. Damit wiesen rund 64 % der extrahierten Proben Fragmentlängen von mindestens 300 bp auf. In zwölf Proben ließen sich Fragmente bis 200 bp amplifizieren, sodass nicht klar ist, ob in diesen Proben Fragmente der geforderten 300 bp Länge vorhanden waren (Abbildung 10).

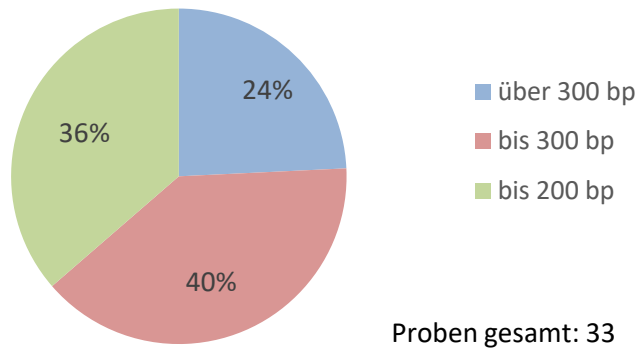


Abbildung 10: Prozentualer Anteil der aus FFPE-Gewebe isolierten DNA-Proben mit jeweiliger maximaler Fragmentlänge, die sich mittels Qualitäts-PCR amplifizieren ließ.

4.2.2 Amplifizierbarkeit der DNA aus Glaskörperaspiraten

Die maximale amplifizierbare Fragmentlänge in der PCR zur Bestimmung der DNA-Amplifizierbarkeit der drei im Rahmen dieser Arbeit verwendeten Glaskörperaspirate wurde dem entsprechenden molekularpathologischen Befund entnommen. Demnach ließen sich in einem Fall Amplifikate bis 600 bp Länge, in einem Fall bis 400 bp und in einem weiteren bis 100 bp Länge erzeugen.

4.2.3 Amplifizierbarkeit von B-NHL-DNA in Multiplex-PCR-Ansätzen mit VH-Primern mit Fluidigm-Anhang

Im Rahmen dieser Arbeit wurden für die NGS-Analysen IGH-Rearrangements in Multiplex-PCR-Ansätzen mit verschiedenen VH-Primersets vervielfältigt. Ob und in welchem Umfang eine Amplifikation erfolgt war, wurde mittels Agarose-Gelelektrophorese überprüft.

Amplifizierbarkeit von B-NHL-Zelllinienverdünnungsproben

Die DNA-Verdünnungsproben der Lymphom-Zelllinien Granta-519 sowie OCI-LY3 wurden mit dem in Tabelle 5 beschriebenen PCR-Protokoll amplifiziert. Das verwendete Enzym war die AmpliTaq Gold Polymerase von Thermo Fisher Scientific. Die Verdünnungsproben der Zelllinie Granta-519 ließen sich sowohl mit dem Primerset FR1 wie dem Primerset FR2 amplifizieren, wobei mit dem Primerset FR1 in allen Verdünnungen starke Banden erzielt wurden, mit dem Primerset FR2 hingegen mäßige Banden für die Verdünnung 1:5 und die Verdünnung 1:10 und schwache Banden für die übrigen Verdünnungen. Auch die Verdünnungsproben der Zelllinie OCI-LY3 ließen sich mit beiden Primersets amplifizieren, wobei mit dem Primermix-FR1 nur schwache Banden, mit dem Primermix-FR2 durchgehend starke Banden erzeugt wurden.

Amplifizierbarkeit von Patientenproben

Dasselbe PCR-Protokoll wurde auf alle 36 Patientenproben angewandt (Abbildung 11), pro Probe in einem Ansatz mit VH-Primern für die FR1- und einem mit VH-Primern für die FR2-Region (Abbildung 11).

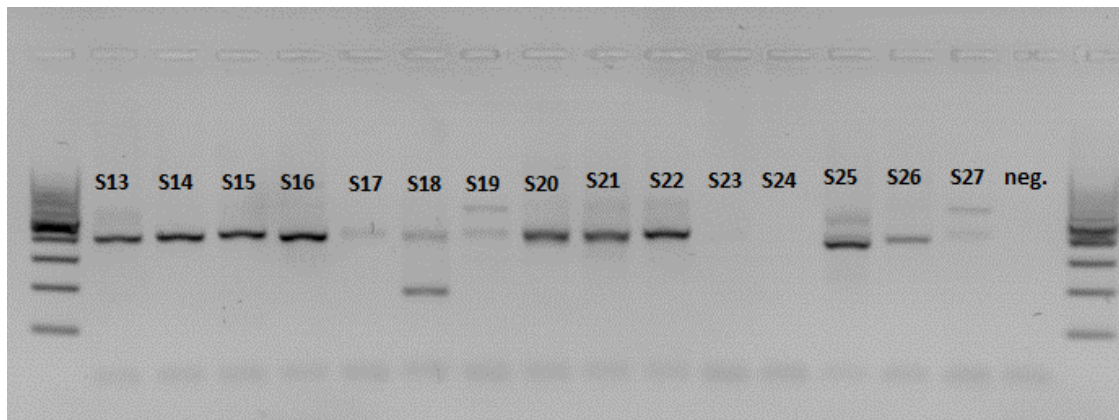


Abbildung 11: Gelelektrophoretische Darstellung der IGH-Amplifikation der Patientenproben mit hohem Lymphomzellgehalt (S13 bis S27) unter Verwendung der DNA-Polymerase AmpliTaq Gold, Primerset FR1. In den meisten Proben ließ sich eine Amplifikation erzielen. neg.: Negativkontrolle.

Mit Primern für die FR1 konnte in fünf Fällen keine Bande detektiert werden. Mit Primern für die FR2 war in drei Fällen keine Bande nachweisbar. Auch Ansätze ohne detektierbare Bande nach Multiplex-PCR mit VH-Primern wurden im Rahmen der Barcode-PCR amplifiziert und sequenziert, falls sie in der

darauffolgenden Gelelektrophorese eine Bande lieferten. In zwei Fällen war weder mit den Primern für die FR1 noch mit denen für die FR2 eine Bande zu erzeugen. Im ersten Fall handelte es sich um die aus einem Lymphknoten gewonnene Verlaufsprobe (S41) eines folliculären Lymphoms von Patientin 20, aus der bereits in der Qualitäts-PCR nur Fragmente einer maximalen Länge von 200 bp amplifiziert werden konnten. Im zweiten Fall handelte es sich um ein Glaskörperaspirat (S47), in dem histologisch und immunhistochemisch keine Lymphomzellen nachgewiesen worden waren und das ebenfalls eine stark eingeschränkte DNA-Qualität mit maximal amplifizierbarer Fragmentlänge von 100 bp aufgewiesen hatte.

4.3 Parameter der Klonalitätsanalyse in NGS-Datensätzen von Verdünnungsproben von B-NHL-Zelllinien-DNA

Als erster Schritt sollten Parameter der Klonalitätsanalyse mittels NGS und Charakteristika klonaler Proben in NGS-Datensätzen ermittelt werden. Hierzu wurde das IGH-Rearrangement zweier Lymphomzelllinien bestimmt, Verdünnungsproben verschiedener Verhältnisse von Lymphom-DNA in PBMC-DNA hergestellt, mittels NGS sequenziert und durch IMG/HighV-Quest analysiert. Herangehensweisen zur Interpretation der gelieferten Daten zur Klonalitätsanalyse wurden erarbeitet.

4.3.1 Ermittelte IGH-Rearrangements der Lymphom-Zelllinien

Zelllinie Granta-519

Wie unter 3.4 beschrieben wurde zunächst ermittelt, durch welchen VH-Primer das klonale Rearrangement der Zelllinie amplifiziert wurde. Hierzu wurden die Standard-Primer zur GeneScan-Analyse für die FR2 verwendet. Die Amplifikation gelang bei der Zelllinie Granta-519 mit dem Primer für die VH-Segmentfamilie 4 (Daten nicht gezeigt). Die darauffolgende Amplifikation mit dem

entsprechenden Primer für die FR1 mit Fluidigm-Anhang, die Sequenzierung des Amplifikats nach Sanger und die Analyse mittels IMGT/V-Quest lieferte die in Tabelle 9 auszugsweise dargestellte Charakterisierung des IGH-Rearrangements.

Tabelle 9: Charakterisierung des IGH-Rearrangements der Zelllinie Granta-519 mit dem FR1-Primer set durch IMGT/V-Quest. Für V-Gen und -Allel bestand gleiche Übereinstimmungen mit zwei Einträgen der Referenzdatenbank.

V-Gen und Allel	D-Gen und Allel	J-Gen und Allel	CDR3-IMGT-Länge (AA)	AA-Sequenz der CDR3	V-Region Identity (%)	Sequenz-Länge (nt)
IGHV4-59*01 F, or IGHV4-59*07 F	IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRSEYFDY	100	339

Hervorzuheben ist die V-Region Identity von 100 %, die anzeigt, dass das V-Gen in diesem IGH-Rearrangement sich komplett mit dem der IMGT-Datenbank deckte und in diesem Abschnitt des IGH-Rearrangements keine SHM vorlagen, die die Bindung eines VH-Primers beeinträchtigen könnten. Das Alignment der Sanger-Sequenz mit der IMGT-Referenzsequenz war im Bereich des V-Gens in 235 von 235 Nukleotiden übereinstimmend. Die Zuteilung zweier möglicher Allele durch IMGT ist der Tatsache geschuldet, dass die Sanger-Sequenz nicht die gesamte FR1 umfasst. Die Allele *01 und *07 des V-Gens IGHV4-59 unterscheiden sich durch ein Nukleotid wenige Positionen jenseits der Primer-Bindungsstelle und damit außerhalb des amplifizierten Bereiches. Hervorzuheben ist auch, dass zu drei weiteren Allelen große Ähnlichkeit bestand; so stimmte das Alignment zu den Allelen *02 und *03 in 234 von 235 Nukleotiden und zu dem Allel *08 in 233 von 235 Nukleotiden überein.

Zelllinie OCI-LY3

Die Amplifikation des klonalen IGH-Rearrangements der Zelllinie OCI-LY3 ließ sich mit dem Standard-FR2-Primer für die GeneScan-Analyse für die VH-Segmentfamilie 4 erzielen. Es folgte die Amplifikation mit dem entsprechenden Primer set für die FR1 mit Fluidigm-Anhang. Die Charakterisierung der mittels

Sanger-Sequenzierung gewonnenen Sequenz ist auszugsweise in Tabelle 10 dargestellt.

Tabelle 10: Charakterisierung des IGH-Rearrangements der Zelllinie OCI-LY3 mit dem FR1-Primerset durch IMGT/V-Quest. Für V-Gen und -Allel bestand gleiche Übereinstimmungen mit drei Einträgen der Referenzdatenbank. Die Sequenzlänge liegt außerhalb des mit dem Primerset erwarteten Bereiches.

V-Gen und Allel	D-Gen und Allel	J-Gen und Allel	CDR3-IMGT Länge (AA)	AA-Sequenz der CDR3	V-Region Identity (%)	Sequenz-Länge (nt)
IGHV4-34*01 F, or IGHV4-34*02 F or IGHV4-34*04 F	IGHD3-10*01 F	IGHJ4*02 F	21	ARGRTGDAEGDVA GLGYYYDF	79,7	451

Die V-Region Identity lag bei dieser Sequenz bei nur 79,7 %, was vor allem in einem Multiplex-PCR-Ansatz die Bindung des entsprechenden VH-Primers und damit die Amplifikation des IGH-Rearrangements beeinträchtigen könnte. Auffällig war ebenfalls die Sequenzlänge von 451 bp, die nicht im für ein IGH-Rearrangement mit dem Primerset FR1 erwarteten Bereich lag (310 – 356 bp) [36].

4.3.2 Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen

Die durch NGS mit der Illumina MiSeq-Plattform gewonnen Datensätze wurden mittels IMGT/HighV-Quest analysiert. Die Auswertung lieferte die Anzahl insgesamt mittels NGS gewonnener Sequenzen und die Anzahl der Sequenzen, die einem IMGT-Clonotype (AA) zugeordnet, d.h. als IGH-Rearrangement erkannt werden konnten. Des Weiteren ergab die Auswertung die Anzahl verschiedener IMGT-Clonotypes (AA) innerhalb des jeweiligen Datensatzes als Maß für die Vielfältigkeit des amplifizierten und sequenzierten IGH-Repertoires.

Zelllinie Granta-519: Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen

Es konnten in allen Ansätzen über 18.000 Sequenzen mittels NGS gewonnen werden, ausgenommen die Verdünnung 1:100 mit dem Primerset FR1; dieser Datensatz war nicht in den durch das Institut für Medizinische Genetik gelieferten Daten enthalten (Abbildung 12). Der Anteil einem IMGT-Clonotype (AA) zugeordneter Sequenzen, d.h. als IGH-Rearrangement erkannter Sequenzen an der Anzahl insgesamt gewonnener Sequenzen lag bei mindestens 82,4 % (Verdünnung 1:1.000, FR2) und bei maximal 97,5 % (Verdünnung 1:5, FR1). Der Anteil war bei gleicher Verdünnung in Ansätzen mit dem Primerset FR1 größer als in Ansätzen mit dem Primerset FR2. Das Primerset FR1 lieferte in Ansätzen gleicher Verdünnung mehr Sequenzen als das Primerset FR2. Ansätze mit höherem Ausgangsanteil an Zelllinien-DNA lieferten tendenziell eine größere Anzahl an Sequenzen als solche mit geringerem Ausgangsanteil.

Die Anzahl verschiedener in einem Datensatz vorhandener IMGT Clonotypes (AA) kann als Maß für die Vielfältigkeit des amplifizierten und sequenzierten IGH-Repertoires dienen. Für beide Zelllinien und beide Primersets war zu beobachten, dass in Ansätzen mit geringerem Ausgangsanteil an Lymphom-DNA tendenziell eine größere Anzahl verschiedener IMGT-Clonotypes (AA) nachweisbar war als in Ansätzen mit höherem Lymphom-DNA-Anteil. Vergleicht man die beiden Primersets, zeigt sich, dass in Ansätzen der gleichen Verdünnung einer Zelllinie mit dem Primerset FR1 weniger verschiedene IMGT-Clonotypes (AA) vorhanden waren als in Ansätzen der gleichen Zelllinie und Verdünnung, die mit dem Primerset FR2 amplifiziert wurden.

Ergebnisse

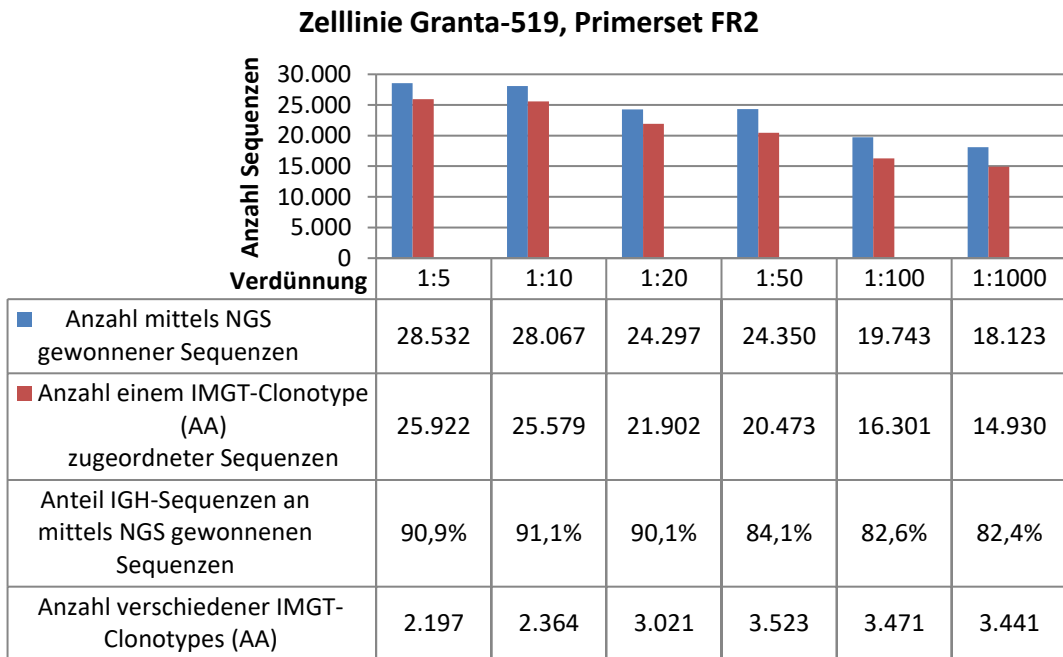
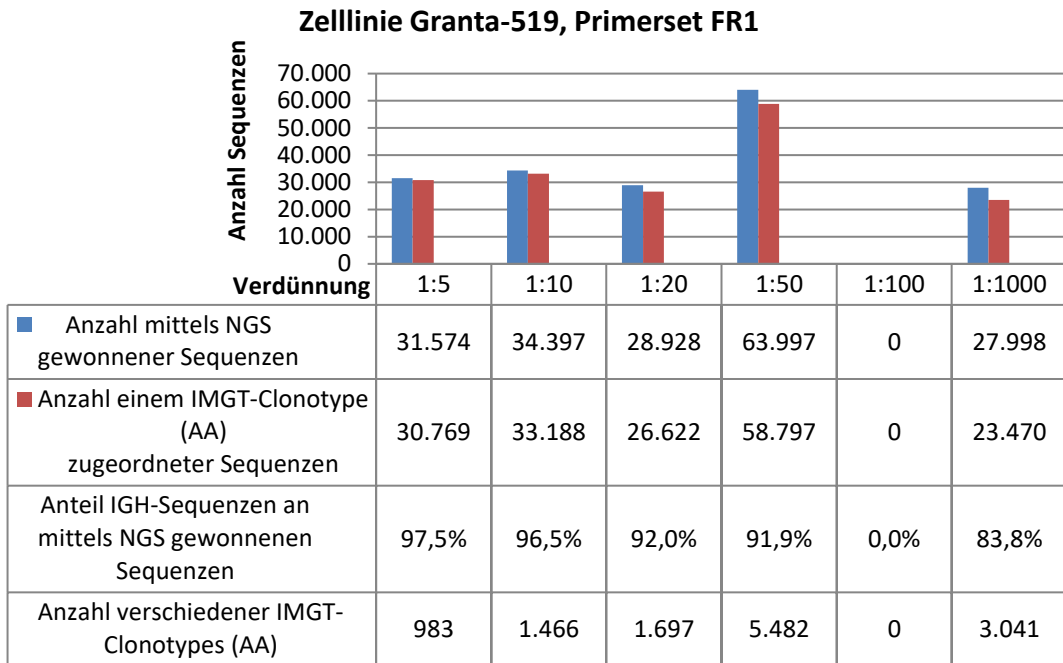


Abbildung 12: NGS-Analyse der IGH-Rearrangements der Zelllinie Granta-519 in verschiedenen Verdünnungsstufen. Anzahl mittels NGS gewonnener Sequenzen, Anzahl und Anteil einem IMGT-Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT-Clonotypes (AA) in Datensätzen verschiedener Verdünnungen der Zelllinien-DNA in PBMC-DNA. Jeweils mit VH-Primern für die FR1- und die FR2.

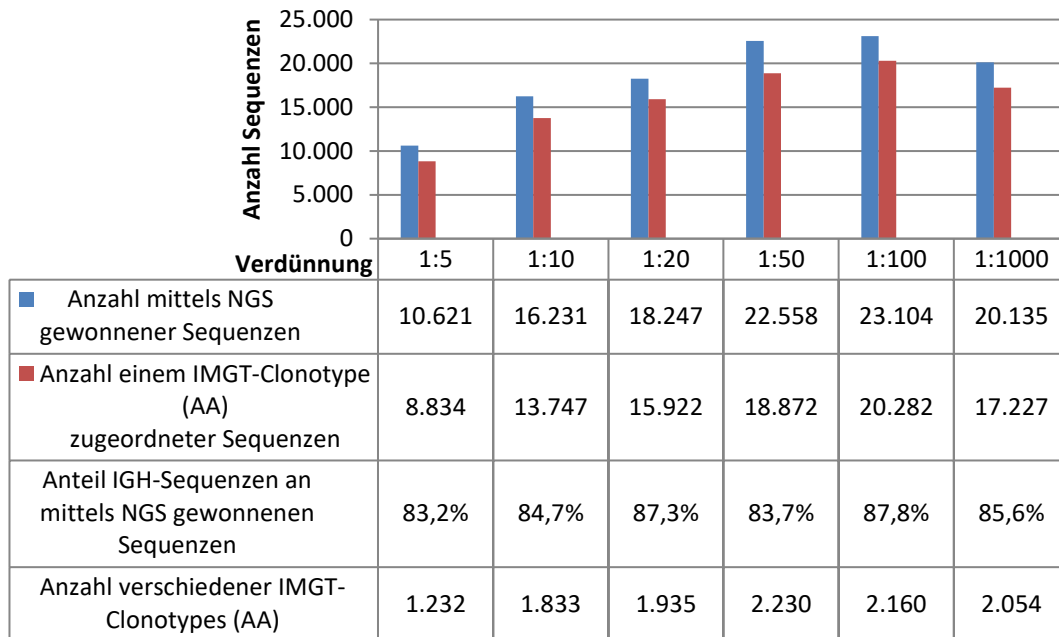
Zelllinie OCI-LY3: Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen

Es konnten in allen Ansätzen der Zelllinie OCI-LY3 über 10.000 Sequenzen mittels NGS gewonnen werden (Abbildung 13). Der Anteil einem IMGT-Clonotype (AA) zugeordneter, d.h. als IGH-Rearrangement erkannter Sequenzen an der Anzahl insgesamt gewonnener Sequenzen lag bei mindestens 79,1 % (Verdünnung 1:5, FR2) und bei maximal 87,8 % (Verdünnung 1:20, FR1). Auch hier war der Anteil bei gleicher Verdünnung in Ansätzen mit dem Primerset FR1 größer als in Ansätzen mit dem Primerset FR2. Das Primerset FR1 lieferte in Ansätzen gleicher Verdünnung mehr Sequenzen als das Primerset FR2, ausgenommen die Verdünnung 1:5. Tendenziell lieferten Ansätze mit geringerer Ausgangskonzentration an Zelllinien-DNA mehr Sequenzen als solche mit höherer Ausgangskonzentration.

Wie oben beschrieben war zu beobachten, dass auch für diese Zelllinie mit beiden Primersets in Ansätzen mit geringerem Ausgangsanteil an Lymphom-DNA tendenziell eine größere Anzahl verschiedener IMGT-Clonotypes (AA) nachweisbar war als in Ansätzen mit höherem Lymphom-DNA-Anteil. Beim Vergleich der beiden Primersets zeigt sich, dass in Ansätzen der gleichen Verdünnung der Zelllinie mit dem Primerset FR1 weniger verschiedene IMGT-Clonotypes (AA) vorhanden waren als in Ansätzen der gleichen Zelllinie und Verdünnung, die mit dem Primerset FR2 amplifiziert wurden.

Ergebnisse

Zelllinie OCI-LY3, Primerset FR1



Zelllinie OCI-LY3, Primerset FR2

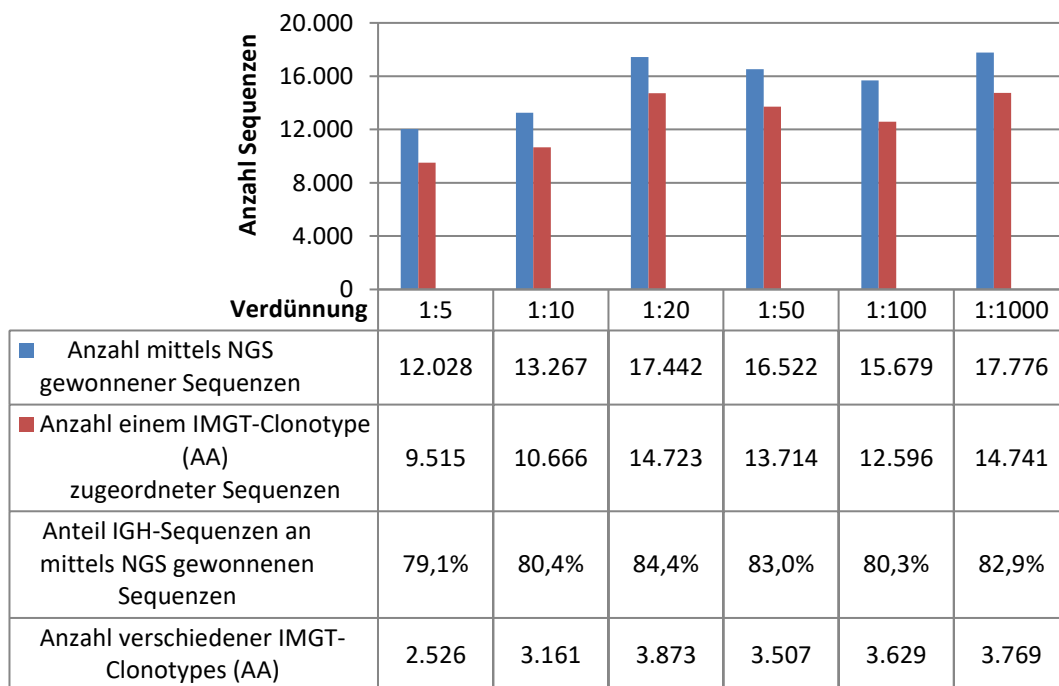


Abbildung 13: NGS-Analyse der IGH-Rearrangements der Zelllinie OCI-LY3 in verschiedenen Verdünnungsstufen. Anzahl mittels NGS gewonnener Sequenzen, Anzahl und Anteil einem IMGT-Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT-Clonotypes (AA) in Datensätzen verschiedener Verdünnungen der Zelllinien-DNA in PBMC-DNA. Jeweils mit VH-Primern für die FR1 und die FR2.

4.3.3 Anteil des IGH-Rearrangements der Lymphom-Zelllinie an der Gesamtzahl der IGH-Sequenzen

Die Analyse durch IMGT/HighV-Quest ordnete pro Datensatz alle als IGH-Rearrangement identifizierte Sequenzen einem IMGT-Clonotype (AA) zu und ordnete die ermittelten und charakterisierten Clonotypes (AA) nach der Anzahl ihnen zugeordneter Sequenzen (nach Sequenzstärke). Im Folgenden sollen die Ergebnisse zweier Datensätze (Verdünnung 1:5, FR1 sowie Verdünnung 1:1.000, FR2) der Zelllinie Granta-519 ausführlich dargestellt werden, um exemplarisch die durch IMGT/HighV-Quest gelieferten Parameter zu demonstrieren und aufzuzeigen, wie diese sich interpretieren lassen, um eine Aussage zum Vorliegen von Klonalität in NGS-Datensätzen zu liefern.

Zelllinie Granta-519: Exemplarische Auswertung des Verdünnungsansatzes 1:5, Primerset FR1

Dieser Verdünnungsansatz 1:5 der Zelllinie Granta-519, amplifiziert mit dem Primerset FR1 lieferte mittels NGS 31.574 Sequenzen, von denen durch Analyse mit IMGT/HighV-Quest 30.769 als IGH-Rearrangement identifiziert und einem IMGT-Clonotype (AA) zugeordnet werden konnten (Tabelle 11). Insgesamt 983 verschiedene Clonotypes (AA) wurden so definiert und nach der Anzahl ihnen zugeordneter Sequenzen von Rang 1 bis Rang 983 angeordnet. Jeder Rang bzw. Clonotype (AA) konnte nun mit den zuvor mittels Sanger-Sequenzierung und Analyse durch IMGT/V-Quest bestimmten Charakteristika des Lymphom-Rearrangements verglichen werden. So wurde deutlich, dass der auf Rang 1 liegende Clonotype (AA) mit V-, D- und J-Gen und -Allel sowie AA-Sequenz und -Länge der CDR3 dem mittels Sanger-Sequenzierung ermittelten IGH-Rearrangement der Zelllinie entspricht. Darüber hinaus wiesen einige Clonotypes (AA) starke Ähnlichkeit zu Rang 1 und dem Rearrangement der Lymphomzelllinie auf und konnten daher als Varianten des Lymphom-Rearrangements interpretiert werden. Hierzu zählen z.B. die Clonotypes (AA) auf den Rängen 2 bis 5. Clonotypes (AA), die keine Ähnlichkeit zu dem der Lymphomzelllinie aufwiesen, waren der durch Verdünnung in PMBC-DNA simulierten Hintergrundpopulation zuzuordnen, hierzu zählen z.B. die Clonotypes (AA) auf den Rängen 6 bis 10

(Tabelle 11). Die Clonotypes (AA) jenseits von Rang 10 sind nicht in Tabelle 11 dargestellt. Hierunter fanden sich hauptsächlich solche, die der Hintergrundpopulation zuzuordnen waren, aber auch solche, die Varianten des Lymphom-Rearrangements darstellten. Hieraus resultierte, dass der Anteil der IGH-Sequenzen, die den Lymphomzellen zuzuordnen waren, noch höher lag als der Wert, der aus der Betrachtung der zehn sequenzstärksten Clonotypes (AA) resultierte.

Aus den so durch IMGT/HighV-Quest gelieferten Daten ließ sich der Mindestanteil des IGH-Rearrangements der Lymphomzelllinie an der Gesamtzahl aller einem IGH-Rearrangement zugeordneter Sequenzen in diesem Verdünnungsansatz ermitteln (Abbildung 14).

Etwa 81,6 % (25.118 von 30.769) aller einem IGH-Rearrangement zugeordneter Sequenzen wurden dem unter Rang 1 aufgeführten Clonotype (AA) zugeordnet, der in seinem VDJ-Rearrangement sowie der Länge und AA-Sequenz der CDR3 dem mittels Sanger-Sequenzierung bestimmten IGH-Rearrangement der Lymphomzelllinie entsprach. Auf die Clonotypes (AA) auf den Rängen 2 bis 5, die Varianten des Lymphom-Rearrangements darstellten, entfielen zusammen 2,2 % (691 von 30.769; Abweichung zu Abbildung 14 durch Rundung bedingt). Auf Rang 6 bis 10 fanden sich Clonotypes (AA) der Hintergrundpopulation, denen jeweils maximal 0,2 % aller IGH-Sequenzen zugeordnet wurden. Etwa 15,1 % aller IGH-Sequenzen wurden Clonotypes (AA) der Ränge 11 bis 983 zugeordnet (in Tabelle 11 nicht dargestellt). Unter diesen Rängen fanden sich weitere IGH-Rearrangements der Hintergrundpopulation, sowie Varianten des IGH-Rearrangements der Lymphomzelllinie, denen jeweils weniger als 59 IGH-Sequenzen (1 bis 57 Sequenzen) zugeordnet wurden.

Ergebnisse

Tabelle 11: Darstellung der zehn Clonotypes (AA) mit den meisten zugeordneten Sequenzen im Ansatz der Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1 nach IMGT/HighV-Quest-Analyse. Der NGS-Datensatz umfasste insgesamt 31.574 Sequenzen, von denen 30.769 durch IMGT/HighV-Quest einem von 983 verschiedenen Clonotypes (AA) zugeordnet wurden. Die zehn Clonotypes (AA), denen die meisten Sequenzen zugeordnet wurden, sind mit der jeweiligen Anzahl zugeordneter Sequenzen und genauer Charakterisierung des IGH-Rearrangements aufgeführt. o.a.: „oder andere“ bei mehreren möglichen V-Genen und -Allelen.

Gesamtzahlen gewonnener Sequenzen und ermittelter IMGT-Clonotypes (AA) des Datensatzes

Anzahl mittels NGS gewonnener Sequenzen	31.574
Anzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	30.769
Anzahl verschiedener IMGT-Clonotypes (AA)	983

Top zehn Clonotypes (AA) mit den meisten zugeordneten Sequenzen

Rang	Anzahl zugeordneter Sequenzen	V-Gen und -Allel	D-Gen und -Allel	J-Gen und -Allel	CDR3-IMGT Länge (AA)	AA-Sequenz der CDR3	Sequenz-Länge (nt)
1	25.118	IGHV4-59*01 F	IGHD2-15*01 F, or IGHD3-10*01 F, or IGHD4-17*01 F, or IGHD5-18*01 F, or IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	324
2	250	IGHV4-39*07 F	IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	329
3	211	IGHV4-34*03 F	IGHD5-18*01 F, or IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	323
4	147	IGHV4-59*03 F	IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	323
5	83	IGHV4-61*05 F	IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	329
6	69	IGHV3-73*01 F, or IGHV3-73*02 F	-	IGHJ6*03 F	10	TSTYYYY MDV	323
7	67	IGHV4-59*01 F	IGHD6-13*01 F	IGHJ6*02 F	17	ARDGGSS WYGYYYG MDV	335
8	64	IGHV4-34*01 F, or IGHV4-34*02 F	IGHD1-7*01 F	IGHJ5*02 F	14	ARGTRN WNYIWF DP	326
9	60	IGHV3-23*01 F, o.a.	IGHD3-16*02 F	IGHJ4*02 F	8	AKGHRYF D	312
10	59	IGHV6-1*01 F, or IGHV6-1*02 F	IGHD6-19*01 F	IGHJ4*02 F	12	ARGGSG WYYFDY	330

Granta-519, Verdünnung 1:5, FR1: Verteilung der IGH-Sequenzen auf verschiedene Clonotypes (AA)

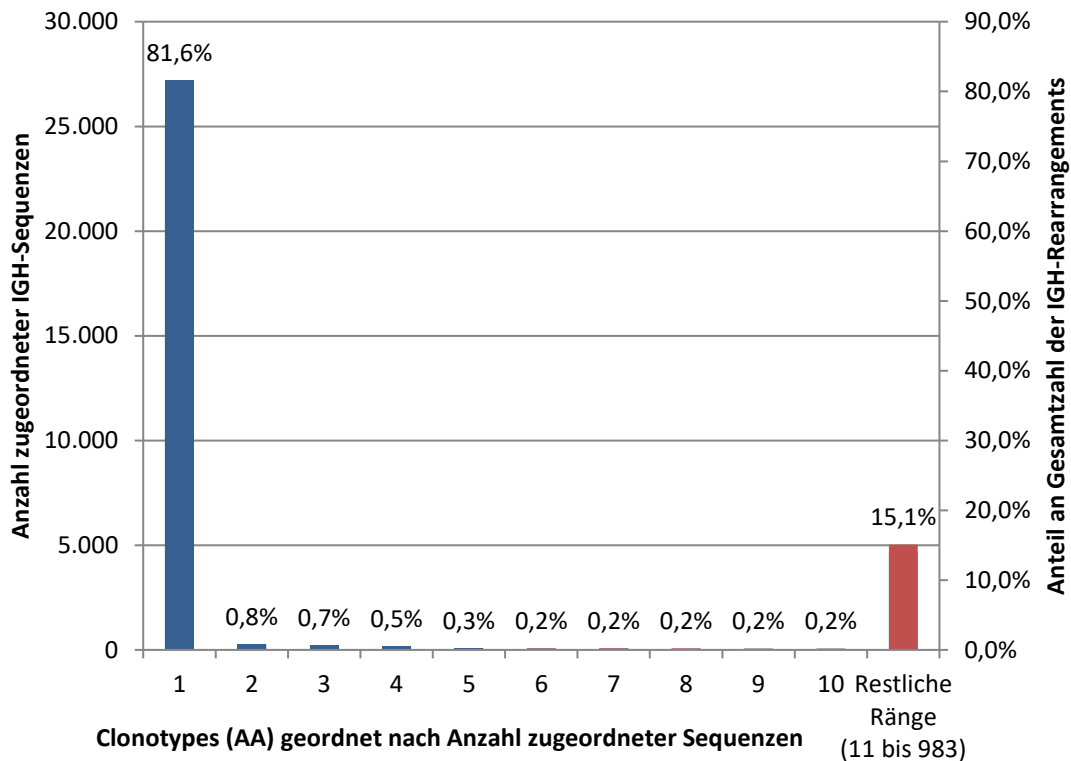


Abbildung 14: Verteilung der IGH-Sequenzen auf verschiedene Clonotypes (AA) in der Probe Granta-519, Verdünnung 1:5, Primerset FR1. Dem Clonotype (AA) auf Rang 1 wurden 81,6 % aller IGH-Sequenzen zugeordnet, er entspricht dem IGH-Rearrangement der Lymphomzelllinie (blau). Auf Rang 2 bis 5 liegen Clonotypes (AA), die Varianten des Rearrangements der Lymphomzelllinie darstellen (blau). Auf verschiedene Clonotypes (AA) der Hintergrundpopulation auf den Rängen 6 bis 10 entfallen jeweils 0,2 % aller Sequenzen. Den restlichen Rängen (11 bis 983) wurden 15,1 % der Sequenzen zugeordnet (Rang 11: 57 Sequenzen, Rang 983: 1 Sequenz); auch hierunter finden sich neben IGH-Rearrangements der Hintergrundpopulation solche, die der Lymphomzelllinie zuzuordnen sind.

Die Analyse durch IMGT/HighV-Quest ermöglichte des Weiteren eine Zuordnung der IGH-Sequenzen nach detektiertem V-Gen und –Allel (Abbildung 15). Auch hier zeigte sich die Überrepräsentation des Lymphom-Rearrangements: 26.195 der 30.769 IGH-Sequenzen (85,1 %) wiesen das V-Gen und –Allel IGHV4-59*01 der Lymphom-Zelllinie auf. V-Genen und –Allelen der Hintergrundpopulation wurden zwischen einer und maximal 467 IGH-Sequenzen zugeordnet.

Ergebnisse

Granta-519, Verdünnung 1:5 , FR1: Verteilung der IGH-Sequenzen nach V-Gen und -Allel

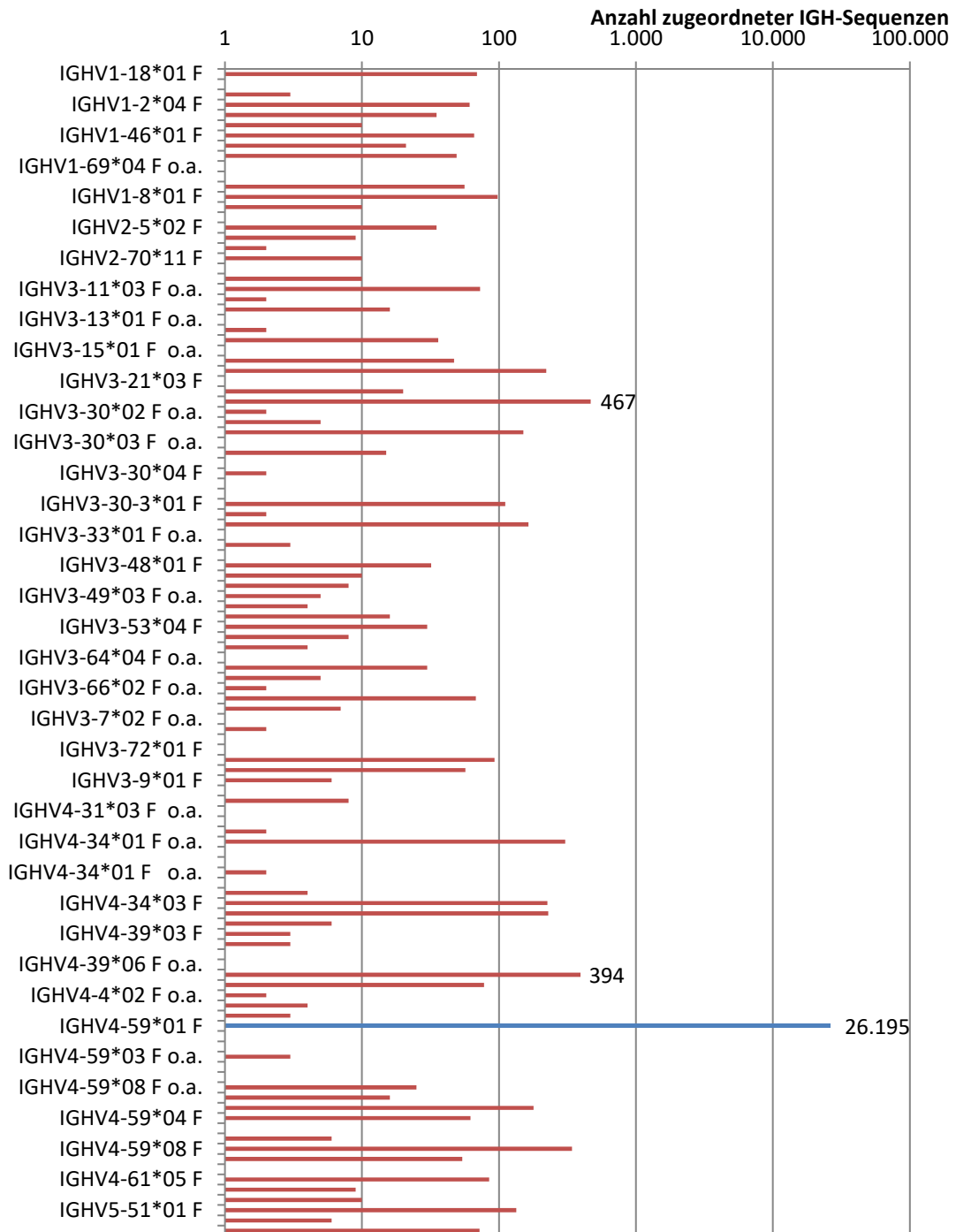


Abbildung 15: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe Granta-519, Verdünnung 1:5, Primerset FR1. (Logarithmische Skalierung.) Die Hintergrundpopulation setzte sich aus Rearrangements mit verschiedenen V-Genen und –Allelen zusammen, denen jeweils zwischen einer und 467 Sequenzen zugeordnet wurden. Dem V-Gen und –Allel des Lymphom-Rearrangements (IGHV-59*01) wurden 26.195 Sequenzen zugeordnet (in blau). Der zweit- und der dritthäufigsten Gen- und Allelkombination wurden 467 respektive 394 Sequenzen zugeordnet. Aus Übersichtsgründen sind nicht alle V-Gene und –Allele in der Achsenbeschriftung aufgeführt. o.a.: „oder andere“, bei mehreren möglichen V-Genen und –Allelen.

Die Zuordnung der sequenzierten und analysierten IGH-Amplifikate nach Sequenzlänge (Abbildung 16) folgt demselben Prinzip wie die etablierte kapillarelektrophoretische Auftrennung eines amplifizierten IGH-Repertoires im Rahmen der GeneScan-Analyse. In diesem Datensatz wurden 25.135 der 30.769 IGH-Sequenzen (81,7 %) einer Sequenzlänge von 324 nt zugeordnet, die der Länge des Lymphom-Rearrangements mit diesem Primerset entspricht. Die graphische Darstellung liefert hier einen klonalen Peak.

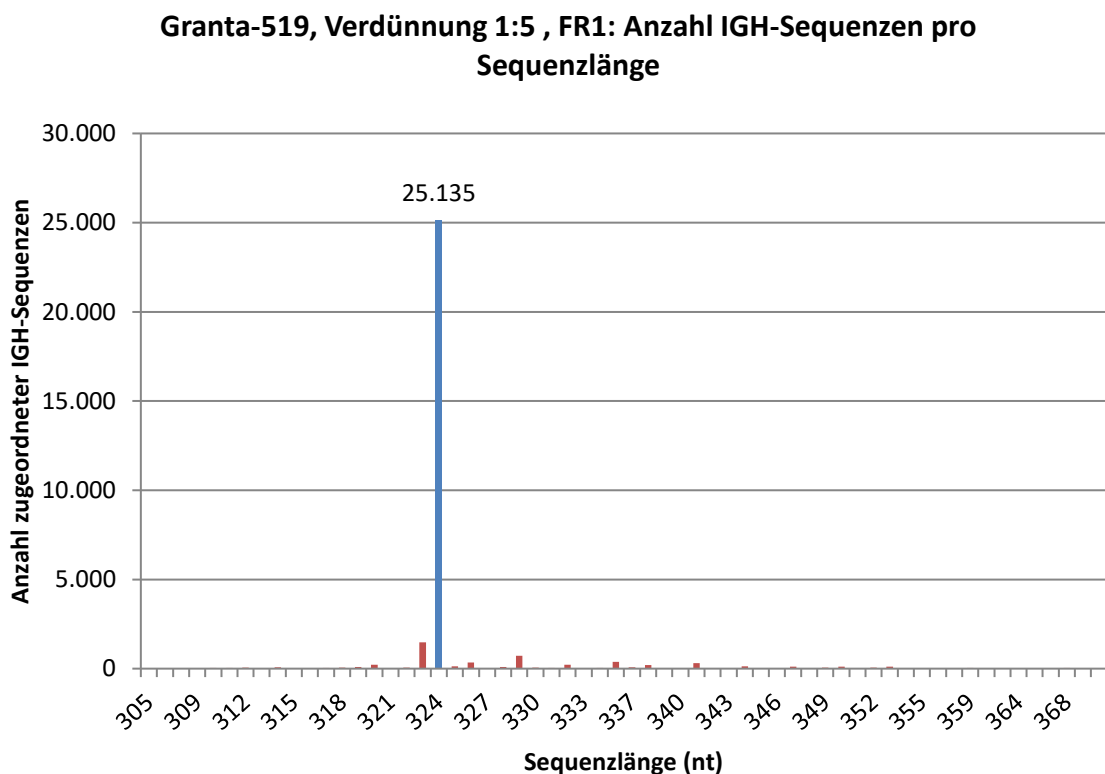


Abbildung 16: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in der Probe Granta-519, Verdünnung 1:5, Primerset FR1. Die Darstellung ist vergleichbar mit der einer kapillarelektrophoretischen Auftrennung von IGH-Amplifikaten im Rahmen der GeneScan-Analyse. Es zeigt sich analog ein monoklonaler Peak (blau) von Amplifikaten der Länge 324 nt, die der Länge des Lymphom-Rearrangements mit diesem Primerset entspricht.

Zelllinie Granta-519: Exemplarische Auswertung des Verdünnungsansatzes 1:1.000, Primerset FR2

Die analoge Auswertung eines Verdünnungsansatzes mit geringem Ausgangsanteil an Lymphom-DNA, des Datensatzes der Verdünnung 1:1.000

der Zelllinie Granta-519, amplifiziert mit dem Primerset FR2, ergab das folgende Bild: Rund 1,1 % der IGH-Sequenzen (170 von 18.123) wurden dem auf Rang 1 liegenden Clonotype (AA) der Lymphomzelllinie zugeordnet (Tabelle 12, Abbildung 17).

Tabelle 12: Darstellung der zehn Clontypes (AA) mit den meisten zugeordneten Sequenzen im Ansatz der Zelllinie Granta-519, Verdünnung 1:1.000, Primerset FR2. Von 18.123 analysierten Sequenzen des NGS-Datensatzes wurden 14.930 als IGH-Rearrangement identifiziert und einem von 3.441 verschiedenen Clontypes (AA) zugeordnet. Es sind wiederum die zehn sequenzstärksten Clontypes (AA) mit genauer Charakterisierung des jeweiligen IGH-Rearrangements und der Anzahl ihnen zugeordneter Sequenzen dargestellt.

Gesamtzahlen gewonnener Sequenzen und ermittelter IMGT-Clontypes (AA) des Datensatzes

Anzahl mittels NGS gewonnener Sequenzen	18.123
Anzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	14.930
Anzahl verschiedener IMGT-Clontypes (AA)	3.441

Top zehn Clontypes (AA) mit den meisten zugeordneten Sequenzen

Rang	Anzahl zugeordneter Sequenzen	V-Gen und -Allel	D-Gen und -Allel	J-Gen und -Allel	CDR3-IMGT Länge (AA)	AA-Sequenz der CDR3	Sequenz-Länge (nt)
1	170	IGHV4-59*01 F, or IGHV4-59*02 F or IGHV4-59*07 F or IGHV4-61*01 F or IGHV4-61*08 F	IGHD5-12*01 F, or IGHD5-18*01 F, or IGHD5-24*01 ORF	IGHJ4*02 F	13	ARVGYRS EYYFDY	260
2	133	IGHV2-5*02 F, or IGHV2-5*08 F	IGHD2-15*01 F, or IGHD2-2*02 F, or IGHD3-3*01 F	IGHJ4*02 F	12	AHRRLPY DTYDY	255
3	131	IGHV1-2*04 F	IGHD2-21*01 F, or IGHD3-3*01 F	IGHJ4*02 F	10	ARGDYDT VDY	254
4	126	IGHV1-8*01 F, or IGHV1-8*02 F	IGHD6-6*01 F	IGHJ4*02 F	13	ARGRIAA RSCWDY	261
5	110	IGHV1-3*01 F	IGHD3-10*01 F	IGHJ1*01 F	10	AREWAGE QQP	253
6	108	IGHV1-46*01 F, or IGHV1-46*02 F or IGHV1-46*03 F	IGHD3-10*01 F	IGHJ4*02 F	14	ARGYYGS GSPFDY	266
7	107	IGHV4-34*01 F, or IGHV4-34*02 F or IGHV4-34*08 F	IGHD2-2*01 F	IGHJ4*02 F	8	ARGQPPA R	243
8	103	IGHV1-69*05 F	IGHD1-1*01 F, or IGHD3-16*01 F	IGHJ6*03 F	16	GRDLGRT GRGPPYM AV	271
9	96	IGHV2-5*02 F, or IGHV2-5*08 F	IGHD3-22*01 F	IGHJ4*02 F	14	AHKSPYY DSSGVVY	261
10	93	IGHV2-70D*04 F	IGHD6-13*01 F	IGHJ3*02 F	17	ARFTPGM AAAGTDA FDI	270

Es fällt auf, dass mit dem Primerset FR2 gewonnene und analysierte Sequenzen zu einer größeren Anzahl verschiedener möglicher V-, D und J-Gene und Allele pro IMGT Clonotype (AA) führen, als mit dem Primerset FR1. So ergeben sich beispielsweise fünf verschiedene V-Gen- und Allelkombinationen für den Clonotype (AA) auf Rang 1. Dies zeigt an, dass die mit dem Primerset FR2 amplifizierten und analysierten Sequenzen zu einer größeren Anzahl von Genen und Allelen die gleiche Übereinstimmungsrate aufweisen als mit dem Primerset FR1. Im Gegensatz zu dem oben beschriebenen Ansatz der Verdünnung 1:5, fanden sich in diesem Datensatz keine Varianten des Lymphom-Rearrangements unter den zehn sequenzstärksten Clonotypes (AA). Dafür fanden sich auf den Rängen 2 bis 10 verschiedene Rearrangements der Hintergrundpopulation mit Anteilen von 0,9 % bis 0,6 %. Etwa 92,1 % aller IGH-Sequenzen wurden Clonotypes (AA) der Ränge 11 bis 3.441 zugeordnet (in Tabelle 12 nicht dargestellt). Unter diesen fanden sich wiederum sequenzschwache Varianten des Lymphom-Rearrangement, sowie Rearrangements der Hintergrundpopulation, denen jeweils zwischen 92 und einer Sequenz zugeordnet wurden.

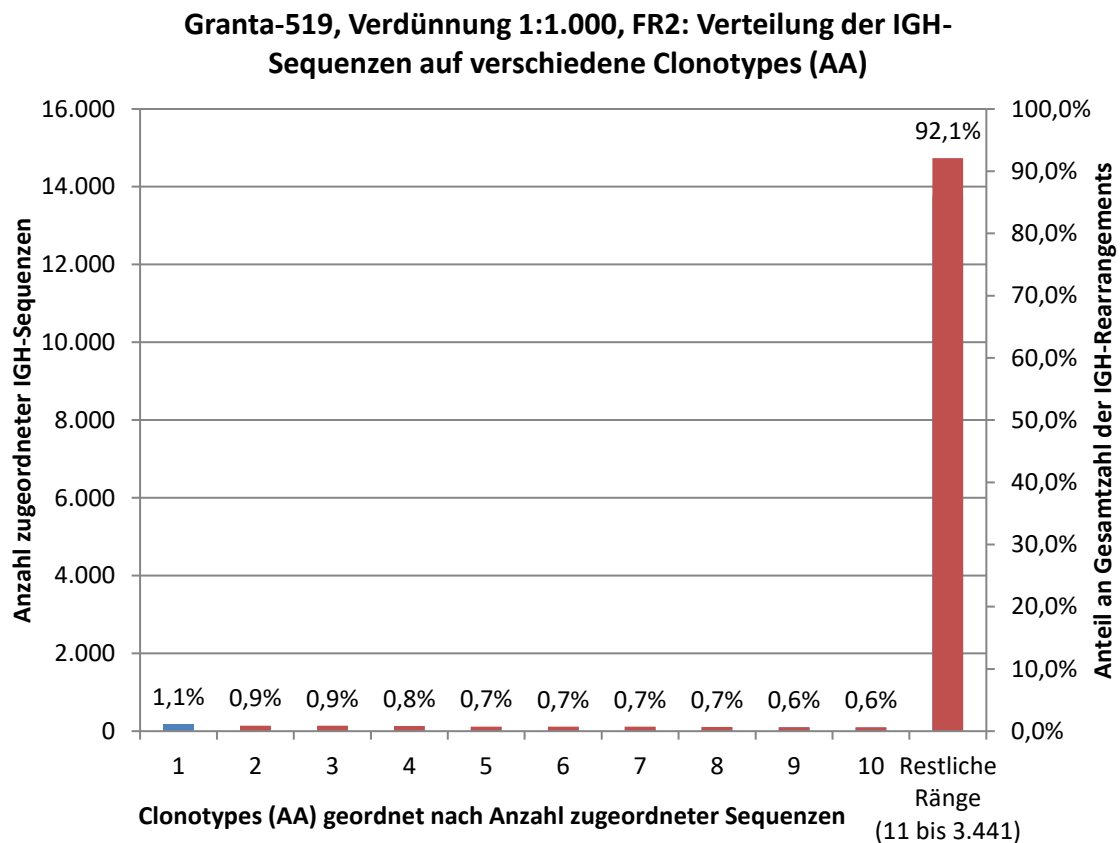


Abbildung 17: Verteilung der IGH-Sequenzen auf verschiedene Clonotypes (AA) in der Probe Granta-519, Verdünnung 1:1.000, Primerset FR2. Dem Clonotype (AA) auf Rang 1 wurden 1,1 % aller IGH-Sequenzen zugeordnet (blau). Wie zeigt, entspricht er dem Rearrangement der Lymphomzelllinie. Verschiedenen Rearrangements der Hintergrundpopulation auf den Rängen 2 bis 10 (rot) wurden 0,9 % bis 0,6 % aller IGH-Sequenzen zugeordnet. 92,1 % aller IGH-Sequenzen wurden Clonotypes (AA) auf den Rängen 11 bis 3.441 zugeordnet, die ebenfalls Rearrangements der Hintergrundpopulation darstellen.

Die Zuordnung der IGH-Sequenzen nach verwendetem V-Gen und –Allel zeigte eine gleichmäßigere Verteilung auf verschiedene V-Gene und –Allele der Hintergrundpopulation als in der Verdünnungsprobe 1:5 dieser Zelllinie (Abbildung 18). Dem V-Gen und –Allel des Lymphom-Rearrangements wurden 681 der 18.123 IGH-Sequenzen (3,8 %) zugeordnet. Dem in dieser Probe am häufigsten verwendeten V-Gen wurden 1.391 IGH-Sequenzen zugeordnet (IGHV3-23).

Ergebnisse

Granta-519, Verdünnung 1:1.000, FR1: Verteilung der IGH-Sequenzen nach V-Gen und -Allel

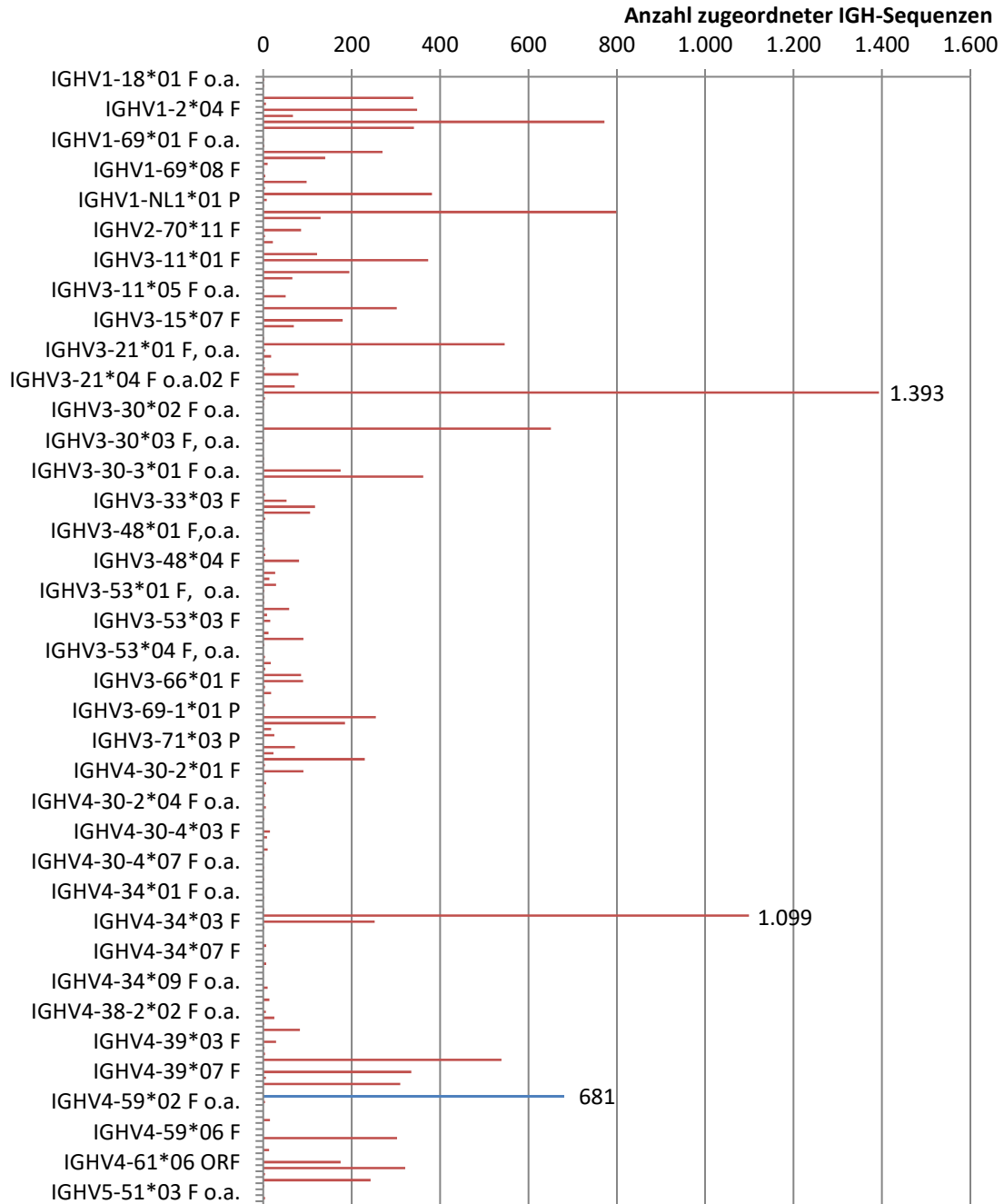


Abbildung 18: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe Granta-519, Verdünnung 1:1.000, Primerset FR2. In dieser Probe zeigte sich eine gleichmäßigere Verteilung der IGH-Sequenzen auf verschiedene V-Gene und –Allele als in der Verdünnungsprobe 1:5 dieser Zelllinie. Dem V-Gen und –Allel des Lymphoms wurden 681 IGH-Sequenzen zugeordnet (blau). Es lag damit hinter häufiger verwendeten V-Genen und –Allelen der Hintergrundpopulation (z.B. IGHV3-23*01 mit 1.393 zugeordneten Sequenzen). Aus Übersichtsgründen sind nicht alle V-Gene und –Allele in der Achsenbeschriftung aufgeführt. o.a.: „oder andere“, bei mehreren möglichen V-Genen und –Allelen.

In der graphischen Darstellung der Zuordnung analysierter IGH-Sequenzen nach Sequenzlänge bot sich ein polyklonales Verteilungsmuster (Abbildung 19). Der Sequenzlänge des Lymphom-Rearrangements mit diesem Primerset von 260 nt wurden 227 IGH-Sequenzen zugeordnet (1,3 %).

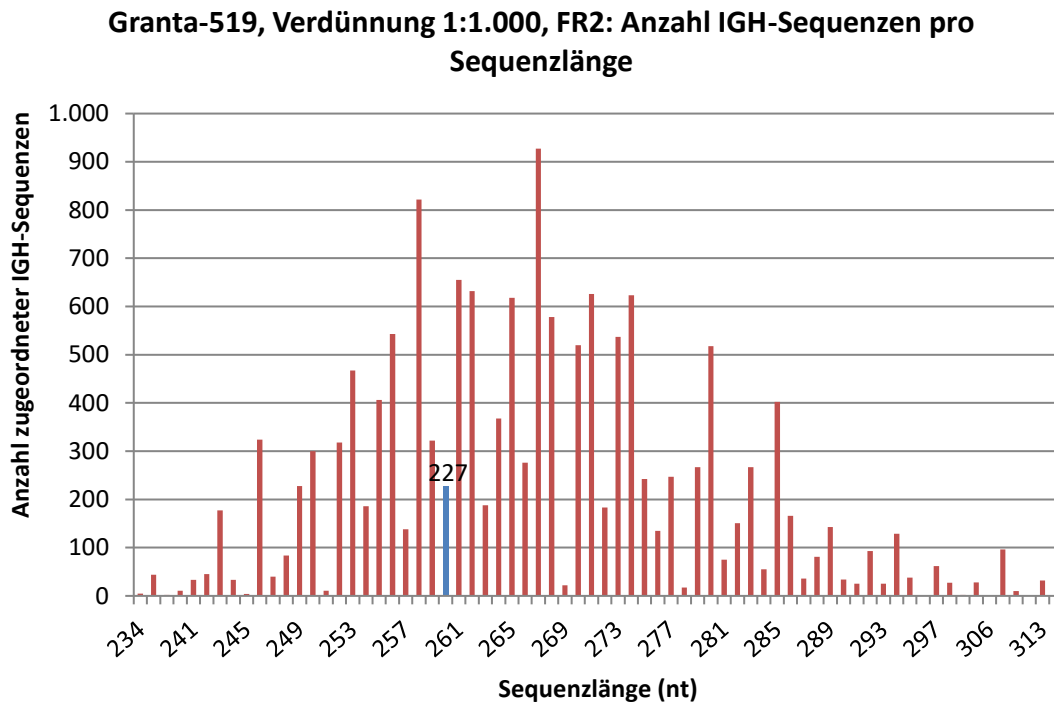


Abbildung 19: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in der Probe Granta-519, Verdünnung 1:1000, Primerset FR2. Es zeigt sich ein polyklonales Verteilungsmuster. Der Sequenzlänge des Lymphom-Rearrangements wurden 227 IGH-Sequenzen zugeordnet (blau).

Zelllinie Granta-519: Gesamtergebnisse der Verdünnungsproben

Insgesamt konnte in allen Verdünnungsansätzen der Zelllinie mit beiden Primersets das mittels Sanger-Sequenzierung bestimmte Lymphom-Rearrangement nachgewiesen werden (Tabelle 13). In allen Ansätzen fand sich das Lymphom-Rearrangement auf Rang 1 als sequenzstärkster Clonotype (AA) des jeweiligen IMGT/HighV-Quest-Datensatzes. Unter den zehn sequenzstärksten Clonotypes (AA) fanden sich außerdem Clonotypes (AA), deren IGH-Rearrangements Varianten des Lymphom-Rearrangements darstellen.

Ergebnisse

Tabelle 13: Übersicht der NGS-Ergebnisse der Verdünnungsproben Zelllinie Granta-519, Primersets FR1 und FR2 nach IMGT/HighV-Quest-Auswertung; Zahlen und Anteile der IGH-Sequenzen, die dem Lymphom-Rearrangement bzw. Rearrangements der Hintergrundpopulationen zugeordnet werden konnten. Für jeden Verdünnungsansatz wurden die zehn Clonotypes (AA) mit den meisten zugeordneten Sequenzen betrachtet und als Lymphom-Rearrangement (Sanger-Sequenz), Lymphom-Rearrangement (Variante) oder als Teil der Hintergrundpopulation gewertet. Aus der Anzahl der so zugeordneten Sequenzen ließ sich der Mindestanteil des Lymphom-Rearrangements und seiner Varianten an der Gesamtzahl aller als IGH-Rearrangement identifizierter Sequenzen bestimmen. Dieser Anteil war jeweils deutlich höher als der Ausgangsanteil der Lymphom-DNA in der Verdünnungsprobe. Auch in den Verdünnungen von 1:1.000 war das Lymphom-Rearrangement mit mindestens 2,3 % (FR1) bzw. mindestens 1,1 % (FR2) nachweisbar. Die Prozentwerte der Hintergrundpopulation geben den Gesamtanteil aller Clonotypes (AA) der Hintergrundpopulation unter den top zehn Clonotypes (AA) an. Der Prozentwert des sequenzstärksten einzelnen Clonotypes der Hintergrundpopulation ist jeweils in Klammern angegeben.

FR1	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Clonotypes (AA) der Ränge 1 bis 10: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen			Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
		Lymphom-Rearrangement (Sanger-Sequenz)	Lymphom-Rearrangement (Varianten)	Rearrangements der Hintergrundpopulation	
1:5	30.769 100,0%	25.118 81,6%	691 2,2%	319 1,0% (0,22)	4.641 15,1%
1:10	33.188 100,0%	23.710 71,4%	972 2,9%	605 1,8% (0,38)	7.901 23,8%
1:20	26.622 100,0%	14.603 54,9%	906 3,4%	615 2,3% (0,48)	10.498 39,4%
1:50	58.797 100,0%	16.998 28,9%	2.929 5,0%	523 0,9% (0,45)	38.347 65,2%
1:100	-	-	-	-	-
1:1.000	23.470 100,0%	544 2,3%	0 0,0%	2297 9,8% (1,38)	20.629 87,9%

FR2	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Clonotypes (AA) der Ränge 1 bis 10: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen			Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
		Lymphom-Rearrangement (Sanger-Sequenz)	Lymphom-Rearrangement (Varianten)	Rearrangements der Hintergrundpopulation	
1:5	25.922 100,0%	15.710 60,6%	1.125 4,3%	90 0,3% (0,30)	8.997 34,7%
1:10	25.579 100,0%	13.294 52,0%	113 0,4%	275 1,1% (0,37)	11.897 46,5%
1:20	21.902 100,0%	7.812 35,7%	870 4,0%	475 2,2% (0,63)	12.745 58,2%
1:50	20.473 100,0%	4.369 21,3%	511 2,5%	816 4,0% (0,64)	14.777 72,2%
1:100	16.301 100,0%	1.647 10,1%	302 1,9%	749 4,6% (0,76)	13.603 83,4%
1:1.000	14.930 100,0%	170 1,1%	0 0,0%	1.007 6,7% (0,89)	13.753 92,1%

Es zeigte sich ein Zusammenhang zwischen Ausgangsgehalt an Lymphom-DNA und dem Anteil des Lymphom-Rearrangements an der Gesamtzahl der als IGH-Rearrangement identifizierten Sequenzen: Je größer der Anteil an Lymphom-DNA in der Ausgangsprobe, desto größer war der Anteil des Lymphom-Rearrangements an der Gesamtzahl der gewonnenen IGH-Sequenzen. Es fiel auf, dass in allen Ansätzen der Anteil des Lymphom-Rearrangements deutlich höher war als der in der Verdünnungsprobe enthaltene Ausgangsanteil an Lymphom-DNA. Auch in der Verdünnung von 1:1.000 war nach NGS und Analyse mittels IMGT-High/V-Quest das Lymphom-Rearrangement deutlich nachweisbar. In den Ansätzen mit dem Primerset FR1 war der Anteil des Lymphom-Rearrangements an der Gesamtzahl der IGH-Sequenzen unter den zehn sequenzstärksten Clonotypes (AA) höher als in Ansätzen der gleichen Verdünnung mit dem Primerset FR2. Der Anteil der Sequenzen, die Clonotypes (AA) jenseits von Rang 10 zugeordnet wurden, war in Ansätzen mit dem Primerset FR2 größer.

Zelllinie OCI-LY3

In den NGS-Datensätzen der Zelllinie OCI-LY3 ließ sich das mittels Sanger-Sequenzierung ermittelte IGH-Rearrangement nicht nachweisen. In allen Verdünnungsansätzen der Zelllinie bot sich ein polyklonales Verteilungsmuster auf Ebene der Clonotypes (AA), der detektierten Sequenzlängen (exemplarisch Verdünnungsansätze 1:5 und 1:1.000, FR2, Abbildung 20) sowie der verwendeten V-Gene und –Allele (exemplarisch Verdünnungsansätze 1:5 und 1:1.000, FR2, Abbildung 21 und Abbildung 22) sowohl in Proben mit hohem als auch in solchen mit niedrigem Anteil an Lymphomzelllinien-DNA.

Ergebnisse

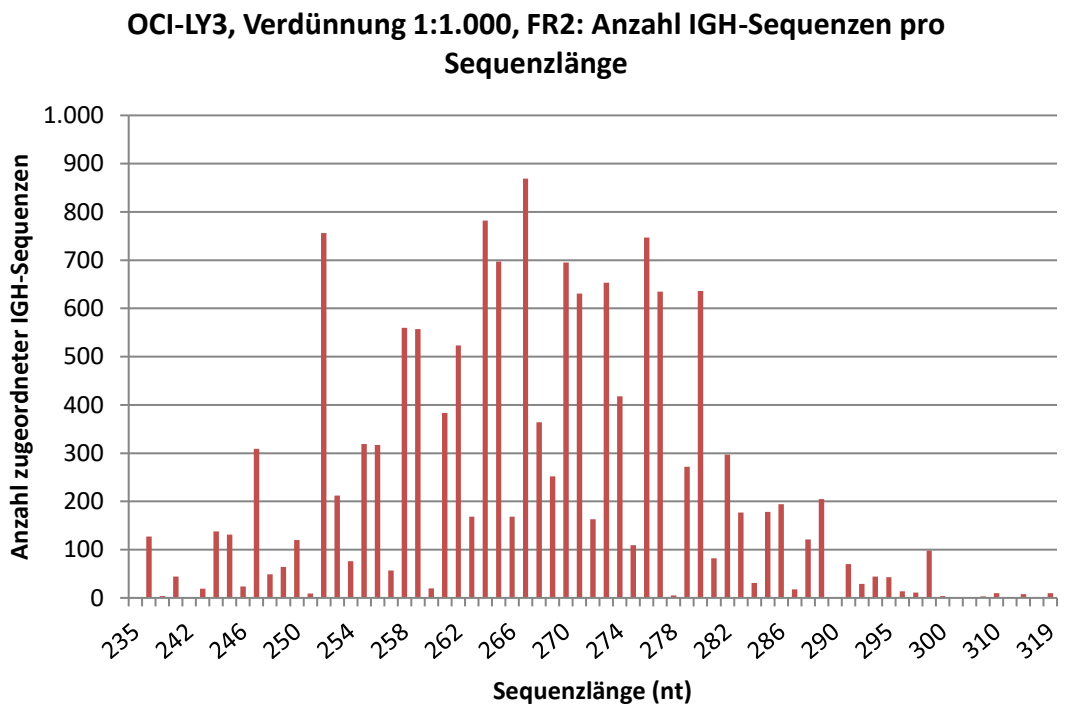
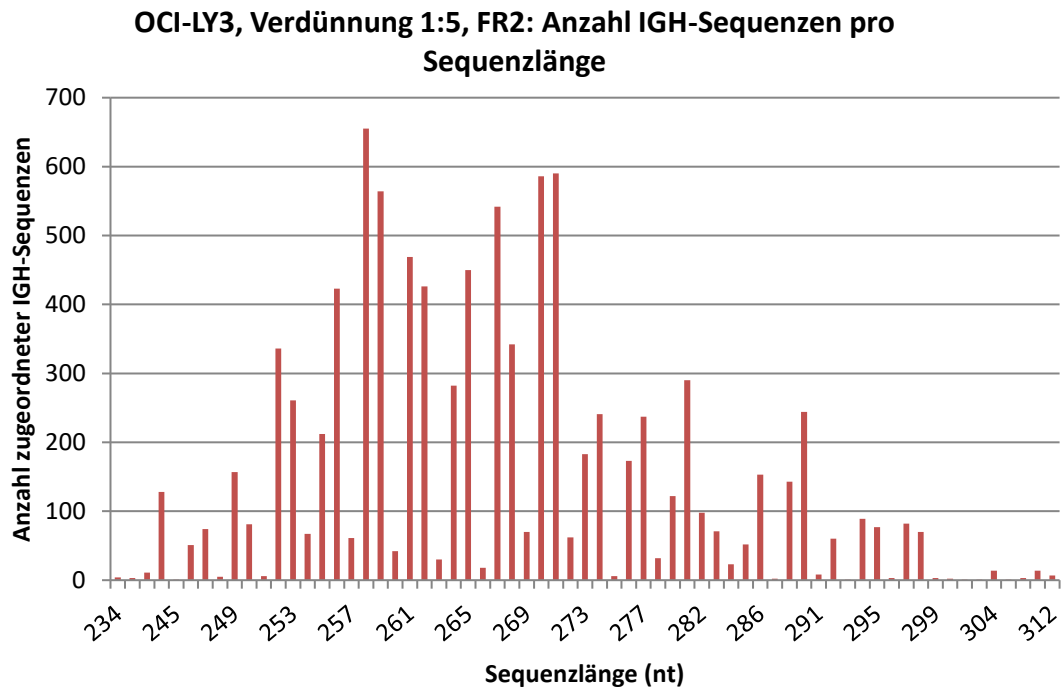


Abbildung 20: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in den Probe OCI-LY3, Verdünnung 1:5 sowie Verdünnung 1:1.000, Primerset FR2. Es zeigt sich in beiden Datensätzen ein polyklonales Verteilungsmuster.

OCI-LY3, Verdünnung 1:5, FR2: Verteilung der IGH-Sequenzen nach V-Gen und -Allel

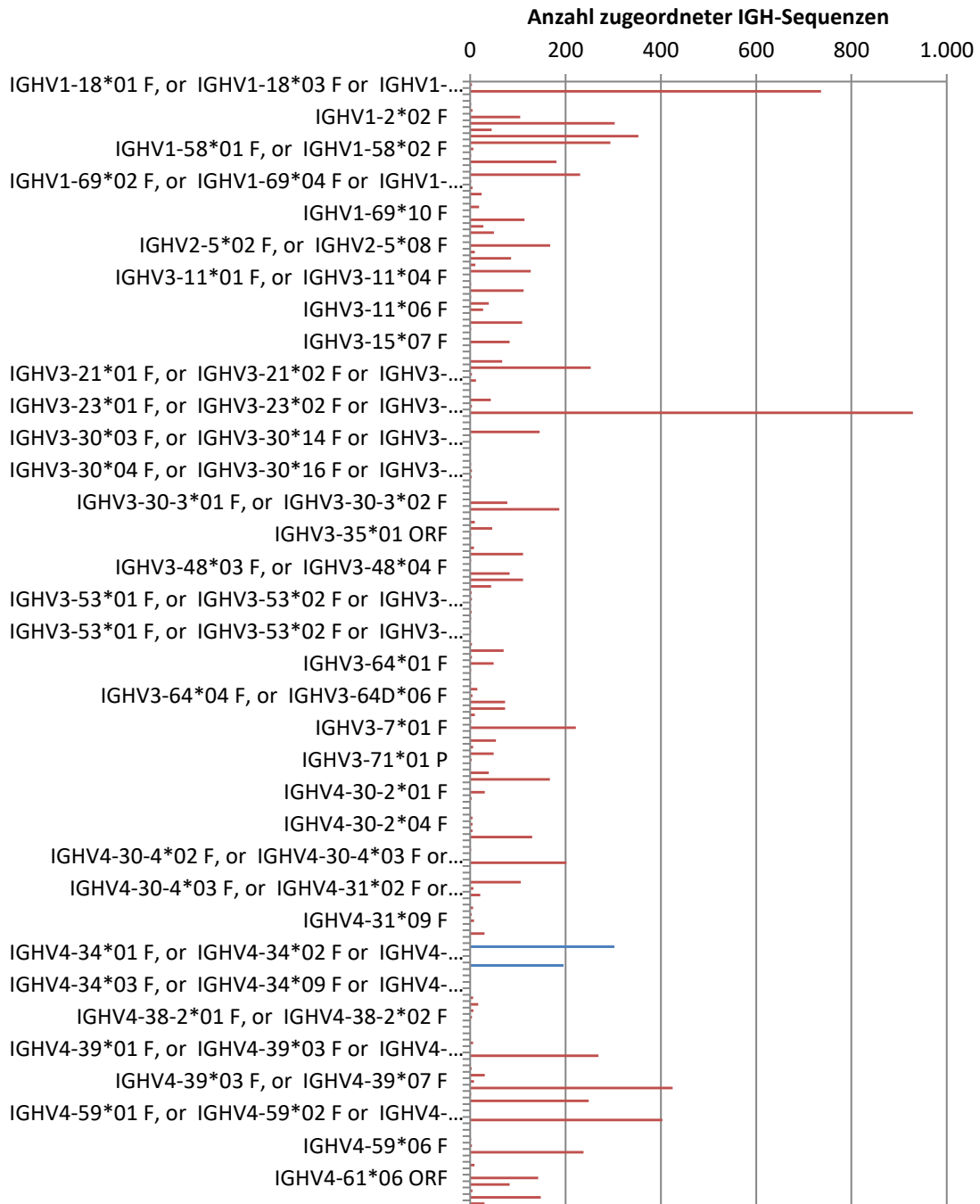


Abbildung 21: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und -Allel in der Probe OCI-LY3, Verdünnung 1:5, Primerset FR2. Es findet sich keine überproportionale Verwendung einzelner V-Gene oder -Allele, insbesondere nicht des V-Gens des mittels Sanger-Sequenzierung ermittelten IGH-Rearrangements der Lymphomzelllinie (IGHV4-34, blau). Aus Übersichtsgründen sind nicht alle V-Gene und -Allele in der Achsenbeschriftung aufgeführt.

OCI-LY3, Verdünnung 1:1.000, FR2: Verteilung der IGH-Sequenzen nach V-Gen und -Allel

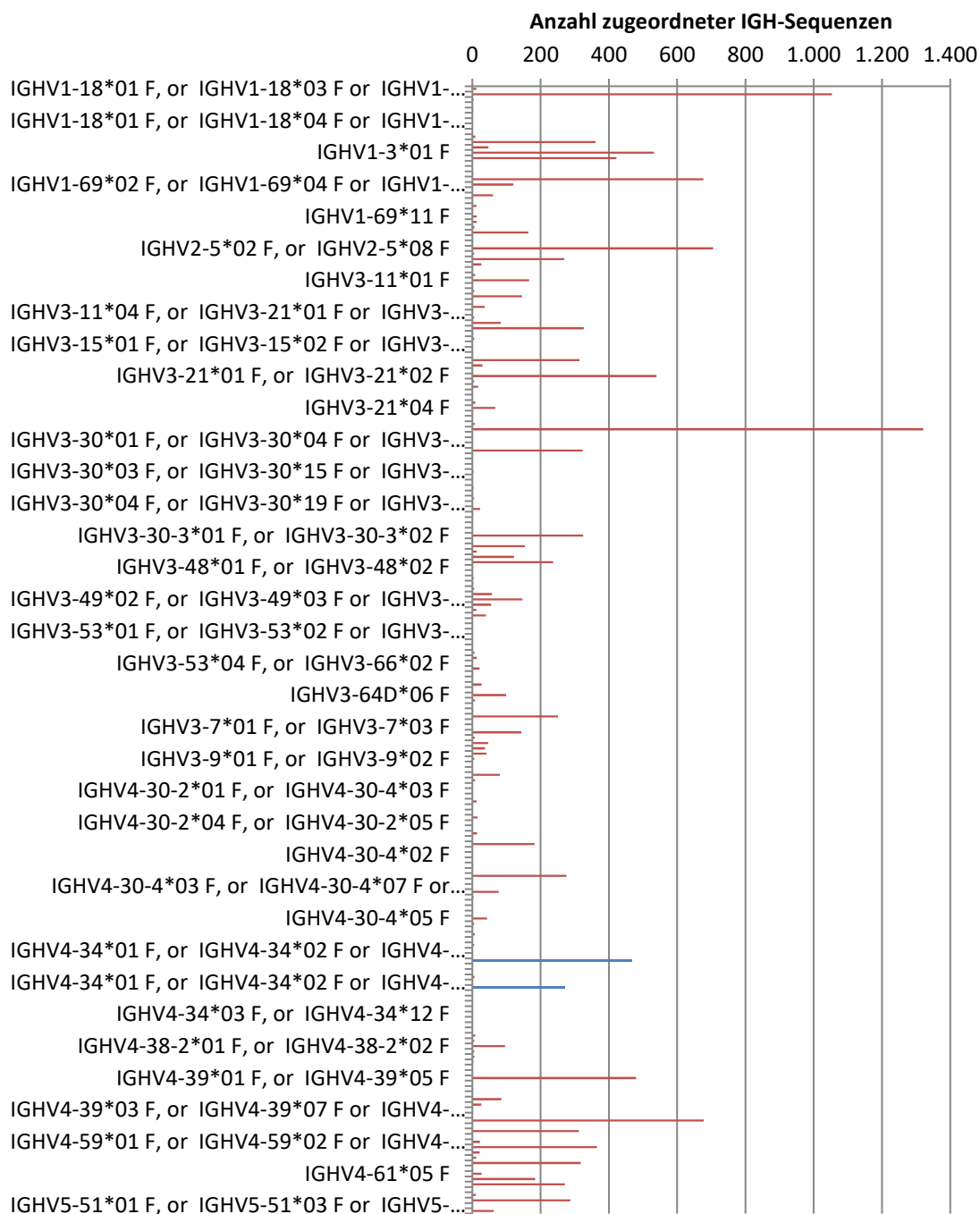


Abbildung 22: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und -Allel in der Probe OCI-LY3, Verdünnung 1:1.000, Primersset FR2. Es bietet sich ein polyklonales Bild ohne Überrepräsentation des V-Gens der Lymphomzelllinie (IGHV4-34, blau). Aus Übersichtsgründen sind nicht alle V-Gene und -Allele in der Achsenbeschriftung aufgeführt.

4.3.4 IGH-Rearrangements der lymphozytären Hintergrundpopulation

Durch die fehlende Nachweisebarkeit des Lymphom-Rearrangements in den Datensätzen der Zelllinie OCI-LY3, eigneten sich diese, um Charakteristika der im hier gewählten Vorgehen amplifizierten Hintergrundpopulation zu ermitteln. In den beiden exemplarisch dargestellten Datensätzen der Zelllinie OCI-LY3 war das häufigste verwendete V-Gen IGHV3-23, gefolgt von IGHV1-18 (Abbildung 21, Abbildung 22).

Den Clonotypes (AA) auf Rang 1 wurden folgende Anteile der Gesamtsequenzen zugeordnet: In der Verdünnungsprobe 1:5, im Datensatz FR1 1,41 %, im Datensatz FR2 0,97 %. In der Verdünnungsprobe 1:1.000, im Datensatz FR1 1,05 %, im Datensatz FR2 0,83 %. Die sequenzschwächsten Clonotypes (AA) waren mit einer Sequenz repräsentiert, entsprechend $< 0,01$ % der Gesamtsequenzen.

Auch Clonotypes (AA) der Hintergrundpopulation, die in Datensätzen der Zelllinie Granta-519 unter den top zehn sequenzstärksten Clonotypes (AA) nachgewiesen wurden, bewegten sich zwischen Werten von 0,22 bis 1,38 % zugeordneter Sequenzen (Tabelle 13).

4.4 NGS- und GeneScan-basierte Klonalitätsanalyse im Vergleich bei Patientenproben mit hohem Lymphomzellgehalt

Nach der Entwicklung eines Auswertungsansatzes der IMGT/HighV-Quest-Datensätze und der Ermittlung zentraler Parameter der Detektion klonaler Lymphozytenpopulationen mittels NGS, wurden in einem zweiten Schritt Patientenproben mit histologisch ermitteltem, hohem Lymphomzellgehalt untersucht, und die anhand der Zelllinienverdünnungsproben ermittelten Parameter der Klonalitätsanalyse angewandt. Die Ergebnisse wurden mit denen der etablierten GeneScan-Analyse zum Klonalitätsnachweis verglichen.

4.4.1 Sequenzierbarkeit nach Sanger und ermittelte IGH-Rearrangements

Zunächst wurde der Versuch unternommen, durch einzelne Amplifikation mit dem entsprechenden VH-Primer für die FR2-Region, durch den sich der klonale Peak der vorangegangenen GeneScan-Analyse reproduzieren ließ, möglichst selektiv das klonale IGH-Rearrangement der jeweiligen Patientenprobe zu amplifizieren und anschließend mittels Kettenabbruchsynthese zu sequenzieren.

Insgesamt konnte in sechs von 15 Patientenproben mit hohem Lymphomzellgehalt eine Sequenz mittels Kettenabbruchsynthese gewonnen werden, in den anderen neun Fällen gelang der Versuch aufgrund von Überlagerungen in der kapillarelektrophoretischen Analyse nicht (Tabelle 14).

Tabelle 14: Sequenzierbarkeit nach Sanger von Patientenproben mit hohem Lymphomzellgehalt. Für alle 15 Patientenproben ist aufgeführt, ob die Sanger-Sequenzierung des klonalen IGH-Rearrangements erfolgreich war, oder nicht. Sie gelang bei vier von fünf B-CLL-Fällen, zwei von fünf ABC DLBCL-Fällen und keinem der fünf FL-Fälle. Sowohl aus KM- als auch aus LK-Proben, sowie einer Probe der Glandula parotidea (Gl. parotidea) gelang die Sequenzierung. Auch bei niedriger in der Test-PCR amplifizierbarer maximaler Fragmentlänge gelang die Sanger-Sequenzierung (Probe S23, bis 200 bp).

Probe	Patient	Diagnose	Material	PCR zur Analyse der DNA-		Sanger-Sequenzierung erfolgreich?
				Amplifizierbarkeit	GeneScan	
S13	P01	CLL	KM	bis 200 bp	pseudoklonal	ja
S14	P02	CLL	KM	bis 200 bp	klonal	ja
S15	P03	CLL	LK	bis 200 bp	klonal	ja
S16	P04	CLL	Gl. parotidea	bis 300 bp	klonal	ja
S17	P05	CLL	LK	bis 300 bp	klonal	nein
S18	P06	FL	LK	bis 200 bp	polyklonal	nein
S19	P07	FL	LK	bis 300 bp	polyklonal	nein
S20	P08	FL	LK	über 300 bp	klonal	nein
S21	P09	FL	LK	bis 300 bp	klonal	nein
S22	P10	FL	KM	über 300 bp	polyklonal	nein
S23	P11	ABC DLBCL	LK	bis 200 bp	pseudoklonal	ja
S24	P12	ABC DLBCL	LK	bis 300 bp	pseudoklonal	nein
S25	P13	ABC DLBCL	LK	Über 300 bp	klonal	ja
S26	P14	ABC DLBCL	LK	über 300 bp	pseudoklonal	nein
S27	P15	ABC DLBCL	LK	über 300 bp	polyklonal	nein

Von den fünf untersuchten B-CLL-Fällen ließ sich in vier eine IGH-Sequenz bestimmen. In allen fünf untersuchten FL-Fällen misslang der Versuch. Von den

untersuchten DLBCL-Fällen lieferten zwei eine IGH-Sequenz. Die Sanger-Sequenzierung gelang sowohl aus verschiedenen Materialien (KM, LK, Glandula parotidea) als auch bei niedriger maximaler Größe der amplifizierbaren Fragmente in der PCR zur Analyse der DNA-Amplifizierbarkeit bis 200 bp.

Die gewonnenen IGH-Sequenzen wurden mittels IMGT/V-Quest analysiert (Tabelle 15).

Tabelle 15: Charakterisierung durch IMGT/V-Quest der durch Sanger-Sequenzierung ermittelten klonalen IGH-Sequenzen aus Patientenproben mit hohem Lymphomzellgehalt. Die V-Region-Identity derjenigen Rearrangements, die sich mittels Sanger-Sequenzierung ermitteln ließen lag bei mindestens 89 %.

Probe/ Diagnose		IMGT/V-Quest-Analyse					V-Region Identity (%)	Sequenz -Länge (nt)
		V-Gen und Allel	D-Gen und Allel	J-Gen und Allel	CDR3- IMGT Länge (AA)	AA-Sequenz der CDR3		
S13	CLL	IGHV4-34*01 F, or IGHV4-34*02 F or IGHV4-34*08 F	IGHD6- 19*01 F	IGHJ4*02 F	17	ARARRSGGA VAGTDFDY	91,7	285
S14	CLL	IGHV1-69*01 F, or IGHV1-69*12 F or IGHV1-69*13 (F) or IGHV1-69D*01 F	IGHD4- 17*01 F	IGHJ6*03 F	18	ARVTVTPTLY YYYYYMDV	99,4	291
S15	CLL	IGHV3-72*01 F	IGHD2-2*01 F	IGHJ6*02 F	20	VRAYCVTTSC SKSRYHGMD V	95	300
S16	CLL	IGHV1-3*01 F	IGHD6- 19*01 F	IGHJ3*01 F	18	ARGYRGDWS GYGNDAYDV	89	289
S23	ABC DLBCL	IGHV3-7*03 F	IGHD2-8*02 F	IGHJ5*02 F	17	ARDPRKDWG SEHNYFDP	94,8	287
S25	ABC DLBCL	IGHV1-2*04 F	IGHD1- 20*01 F	IGHJ6*02 F	22	ARGGKYNW NPLVGGDY YGLDV	94,19	298

4.4.2 Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen

Jede Probe wurde in einem Ansatz mit dem Primerset FR1 und einem mit dem Primerset FR2 amplifiziert und auf der Illumina MiSeq-Plattform sequenziert. Analog zur Auswertung der Zelllinienverdünnungsproben (siehe 4.3.2), lieferte die Analyse der Datensätze mittels IMGT/HighV-Quest die Gesamtzahl mittels NGS gewonnener Sequenzen, die Anzahl einem IMGT-Clonotype zugeordneter

Ergebnisse

und somit als IGH-Rearrangement erkannter Sequenzen, sowie die Anzahl verschiedener IMGT-Clonotypes (AA) pro jeweiligem Datensatz (Tabelle 16).

Tabelle 16: Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT-Clonotypes (AA) in Patientenproben mit hohem Lymphomzellgehalt nach IMGT/HighV-Quest-Analyse. * Fälle mit erfolgreicher Sanger-Sequenzierung des klonalen IGH-Rearrangements.

Probe	Diagnose	Primer-Set	Anzahl mittels NGS gewonnener Sequenzen	Anzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Anteil IGH-Sequenzen an mittels NGS gewonnenen Sequenzen	Anzahl verschiedener IMGT-Clonotypes (AA)
S13 *	CLL	FR1	1.5301	14.652	95,8%	206
		FR2	4.968	2.633	53,0%	302
S14 *	CLL	FR1	7.190	7.081	98,5%	177
		FR2	13.162	10.322	78,4%	34
S15 *	CLL	FR1	19.103	18.624	97,5%	394
		FR2	8.594	8.126	94,6%	264
S16 *	CLL	FR1	7.821	7.199	92,0%	199
		FR2	28.741	27.102	94,3%	1.458
S17	CLL	FR1	18.967	17.596	92,8%	341
		FR2	5.214	4.274	82,0%	564
S18	FL	FR1	3.427	2.616	76,3%	389
		FR2	10.446	7.406	70,9%	1.568
S19	FL	FR1	11.280	3.981	35,3%	125
		FR2	8.236	3.451	41,9%	410
S20	FL	FR1	15.308	12.056	78,8%	1.132
		FR2	10.189	7.619	74,8%	1.444
S21	FL	FR1	16.759	14.275	85,2%	1.633
		FR2	6.616	4.699	71,0%	1.291
S22	FL	FR1	18.501	14.938	80,7%	178
		FR2	7.756	5.040	65,0%	298
S23 *	ABC DLBCL	FR1	24.800	24.407	98,4%	277
		FR2	10.859	10.460	96,3%	475
S24	ABC DLBCL	FR1	21	9	42,9%	6
		FR2	12.153	7.736	63,7%	306
S25 *	ABC DLBCL	FR1	18.899	1.110	5,9%	66
		FR2	12.532	7.204	57,5%	517
S26	ABC DLBCL	FR1	20.032	2	0,0%	2
		FR2	7.897	6.867	87,0%	379
S27	ABC DLBCL	FR1	9.803	1.210	12,3%	35
		FR2	9.536	7269	76,2%	684

Mit dem Primerset FR1 wurden im Durchschnitt 13.814 Sequenzen gewonnen (Minimum 21, Maximum 24.800). Die Anzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen betrug im Schnitt 9.317 (Minimum 2, Maximum 24.407). Damit betrug der Anteil der als IGH-Rearrangement erkannten Sequenzen an der Gesamtzahl analysierter Sequenzen für das Primerset FR1 im Schnitt 58,4 % (Minimum gerundet 0,0 %, Maximum 98,5 %). Die Anzahl verschiedener IMGT-Clontypes (AA) pro Datensatz lag für das Primerset FR1 durchschnittlich bei 344 (Minimum 2, Maximum 1.633).

Mit dem Primerset FR2 wurden durchschnittlich 10.460 Sequenzen mittels NGS gewonnen und analysiert (Minimum 4.968, Maximum 28.741). Hiervon wurden im Schnitt 8.014 einem IMGT-Clonotype (AA) zugeordnet (Minimum 2.633, Maximum 27.102), entsprechend einem Anteil von durchschnittlich 73,4 % (Minimum 41,9 %, Maximum 96,3 %). Die Anzahl verschiedener IMGT-Clontypes (AA) pro Datensatz lag für das Primerset FR2 bei im Schnitt 666 (Minimum 34, Maximum 1.568).

Damit war die durchschnittliche Anzahl der gewonnenen und als IGH-Rearrangement einem Clonotype (AA) zugeordneter Sequenzen zwar größer in den Ansätzen mit dem Primerset FR1 als mit dem Primerset FR2. Allerdings war der Anteil der als IGH-Rearrangement erkannten Sequenzen an der Gesamtzahl mittels NGS gewonnener und analysierter Sequenzen für das Primerset FR1 kleiner als für das Primerset FR2. Analog zu den Ergebnissen der Zelllinienverdünnungsreihen, war die durchschnittliche Gesamtzahl an verschiedenen Clontypes (AA) pro Datensatz in den Ansätzen mit dem Primerset FR2 deutlich größer als mit dem Primerset FR1.

4.4.3 Anteile sequenzstarker IGH-Rearrangements an der Gesamtzahl der IGH-Sequenzen

Patientenproben mit gelungener Sanger-Sequenzierung des klonalen IGH-Rearrangements

Die IMGT/HighV-Quest-Analysen der Patientenproben, bei denen die Sanger-Sequenzierung des klonalen IGH-Rearrangements gelungen war, wurden

Ergebnisse

entsprechend des Vorgehens ausgewertet, das anhand der Zelllinienverdünnungsproben erarbeitet worden war (Tabelle 17).

Tabelle 17: Übersicht der NGS-Ergebnisse der Patientenproben mit hohem Lymphomzellgehalt bei denen das klonale Rearrangement zuvor mittels Sanger-Sequenzierung bestimmt werden konnte , Primersets FR1 und FR2 nach IMGT/HighV-Quest-Auswertung; Anzahlen und Anteile der IGH-Sequenzen, die dem Lymphom-Rearrangement, seinen Varianten oder Rearrangements der Hintergrundpopulation zugeordnet werden können. Es wurden die zehn sequenzstärksten Clonotypes (AA) pro Datensatz betrachtet und als der Sanger-Sequenzierung entsprechendes Lymphom-Rearrangement, als Variante oder als Teil der Hintergrundpopulation eingeordnet. Dann wurde der Anteil der jeweiligen Kategorien an der Gesamtzahl der IGH-Sequenzen des Datensatzes errechnet. In allen Proben bis auf eine entsprach der Clonotype (AA) auf Rang 1 dem Rearrangement der Sanger-Sequenzierung. Die Hintergrundpopulationen waren nur sehr schwach vertreten. Auch unter den Rängen jenseits von Rang 10 fanden sich in allen Datensätzen überwiegend Varianten des Lymphom-Rearrangements. Der Prozentwert des sequenzstärksten einzelnen Clonotypes der Hintergrundpopulation ist jeweils in Klammern angegeben.

Probe und Diagnose	Primer-Set	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Clonotypes (AA) der Ränge 1 bis 10: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen			Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
			Lymphom-Rearrangement (Sanger-Sequenz)	Lymphom-Rearrangement (Varianten)	Rearrangements der Hintergrundpopulation	
S13 CLL	FR1	14.652 100,0%	12.928 88,2%	952 6,5%	0 0,0%	772 5,3%
	FR2	2.633 100,0%	1.544 58,6%	527 20,0%	0 0,0%	562 21,3%
S14 CLL	FR1	7.081 100,0%	6.001 84,7%	616 8,7%	0 0,0%	464 6,6%
	FR2	10.322 100,0%	9.167 88,8%	1.080 10,5%	0 0,0%	75 0,7%
S15 CLL	FR1	18.624 100,0%	16.627 89,3%	521 2,8%	124 0,7% (0,7)	1.352 7,3%
	FR2	8.126 100,0%	6.723 82,7%	401 4,9%	165 2,0% (2,0)	837 10,3%
S16 CLL	FR1	7.199 100,0%	5.407 75,1%	0 0,0%	1.178 16,4% (3,5)	614 8,5%
	FR2	(27.102) (100,0%)	(15.109) * (55,7%)*	(0) (0,0%)	(5.928) (21,9% (5,2))	(6.065) (22,4%)
S23 ABC DLBCL	FR1	24.407 100,0%	22.905 93,8%	584 2,4%	0 0,0%	918 3,8%
	FR2	10.460 100,0%	7.903 75,6%	1.029 9,8%	0 0,0%	1.528 14,6%
S25 ABC DLBCL	FR1	1.110 100,0%	960 86,5%	64 5,8%	0 0,0%	86 7,7%
	FR2	7.204 100,0%	4.880 67,7%	719 10,0%	0 0,0%	1.605 22,3%

* Der Clonotype (AA) auf Rang 1 entsprach nicht dem mittels Sanger-Sequenzierung ermittelten Rearrangement

In allen Ansätzen bis auf einen (S16 FR2) entsprach der sequenzstärkste Clonotype (AA) auf Rang 1 dem zuvor mittels Sanger-Sequenzierung bestimmten klonalen IGH-Rearrangement. Diesem Clonotype (AA) wurden

zwischen 58,6 % und 93,8 % aller einem IMGT-Clonotype (AA) zugeordneter Sequenzen zugeordnet. Unter den betrachteten zehn sequenzstärksten Clonotypes (AA) fanden sich teilweise verschiedene Varianten dieses klonalen Rearrangements denen bis zu 20 % der IGH-Sequenzen zugeordnet wurden. Rearrangements der Hintergrundpopulation waren nur in zwei Proben, bzw. vier Ansätzen unter den zehn sequenzstärksten Clonotypes (AA) vertreten, mit in Summe 0,7 % bis 16,4 % (bzw. 21,9 %, S16 FR2) der IGH-Sequenzen. Einem einzelnen Clonotype (AA) der Hintergrundpopulation wurden maximal 3,5 % der Sequenzen zugeordnet (S16, FR1). Den Clonotypes (AA) der restlichen Ränge wurden 0,7 % bis 21,3 % (bzw. 22,4 %, S16 FR2) der IGH-Sequenzen zugeordnet; unter den hier charakterisierten Rearrangements fanden sich in allen Ansätzen zahlreiche weitere Varianten des klonalen Rearrangements mit jeweils geringer Anzahl zugeordneter Sequenzen. Hieraus folgt, dass die Zahl der dem Lymphom-Rearrangement zuzuordnenden IGH-Sequenzen noch über jener Zahl liegt, die aus der Betrachtung der zehn sequenzstärksten Clonotypes (AA) resultiert.

Im Ansatz der Patientenprobe S16 mit dem Primerset FR2 entsprach der Clonotype (AA) auf Rang 1 nicht dem mittels Sanger-Sequenzierung charakterisierten IGH-Rearrangement. Die Sequenz des Clonotypes (AA) auf Rang 1 war produktiv und hatte eine V-Region Identity von rund 97 %. Im Ansatz mit dem Primerset FR1 derselben Probe war das mittels Sanger-Sequenzierung bestimmte Rearrangement auf Rang 1 nachweisbar.

Insbesondere im Falle der CLL ist die V-Region Identity ermittelter IGH-Rearrangements von Interesse, da anhand derer eine Einteilung prognostischer Relevanz in mutierte und unmutierte Formen erfolgt. In Fällen mit erfolgreicher Sanger-Sequenzierung (S13 bis S16) ergaben sich mittels beider Methoden, NGS und Sanger-Sequenzierung, gleiche Befunde bezüglich des Mutationsstatus (Tabelle 18). Ausgehend vom etablierten Cut-Off-Wert von 98 % V-Region Identity [66] waren die Fälle S13, S15 und S16 mutiert, während Fall S14 unmutiert war. Mittels NGS ergaben sich dabei in allen Fällen eine höhere V-Region Identity als mittels Sanger-Sequenzierung; diese Abweichung betrug bis zu 2,2 Prozentpunkte im Fall S15. Wie oben beschrieben wies der Datensatz

FR2 der Probe S16 ein abweichendes IGH-Rearrangement auf und wurde bei Verdacht auf einen methodischen Fehler von der Beurteilung ausgenommen. Die übrigen Fälle zeigten dabei zwischen den Analysen mit dem FR1- und dem FR2-Primerset übereinstimmende V-Gen- und –Allel-Zuordnungen, wobei die auf dem Primerset FR2 basierenden Analysen in den Fällen S13 und S14 mehr mögliche Allele lieferten als die NGS-Datensätze FR1.

Tabelle 18: V-Region Identity untersuchter CLL-Fälle mit hohem Lymphomzellgehalt in Sanger-Sequenzierung und Clontypes (AA) auf Rang 1 der NGS-Datensätze. Die Werte der V-Region Identity liefern übereinstimmende Befunde bezüglich der Einteilung in mutierte und unmutierte CLL in den Fällen mit gelungener Sanger-Sequenzierung. Die NGS-Datensätze der Probe S17 lieferten keinen Klonalitätsnachweis. Der Datensatz FR2 der Probe S16 wurde bei Verdacht auf methodische Fehler ausgeklammert.

CLL-Probe	Sanger-Sequenz (FR2)		Sequenzstärkster Clonotype (AA) Datensatz FR1		Sequenzstärkster Clonotype (AA) Datensatz FR2	
	V-Gen und -Allel	V-Region Identity (%)	V-Gen und -Allel	V-Region Identity (%)	V-Gen und -Allel	V-Region Identity (%)
S13	IGHV4-34*01 F, or IGHV4-34*02 F or IGHV4-34*08 F	91,7	IGHV4-34*01 F, or IGHV4-34*02 F	91,8	IGHV4-34*01 F, or IGHV4-34*02 F or IGHV4-34*08 F	93,8
S14	IGHV1-69*01 F, or IGHV1-69*12 F or IGHV1-69*13(F) or IGHV1-69D*01F	99,4	IGHV1-69*13 (F)	100	IGHV1-69*01 F, or IGHV1-69*12 F or IGHV1-69*13 (F) or IGHV1-69D*01 F	100
S15	IGHV3-72*01 F	95	IGHV3-72*01 F	97,2	IGHV3-72*01 F	96,2
S16	IGHV1-3*01 F	89	IGHV1-3*01 F	90,1	(IGHV3-72*01 F)	(97,2)
S17	Sanger-Sequenzierung misslungen		IGHV3-30*02 F, or IGHV3-30-5*02 F	90,7	IGHV1-2*04 F	91,3

Patientenproben ohne gelungene Sanger-Sequenzierung des klonalen IGH-Rearrangements

Für Patientenproben, bei denen die Sanger-Sequenzierung des klonalen IGH-Rearrangements nicht gelungen war, wurde die Verteilung der als IGH-Rearrangement erkannten Sequenzen auf Clontypes (AA) und nach Sequenzlänge betrachtet und mit den anhand der Zelllinienverdünnungsproben erarbeiteten Merkmalen von Mono- bzw. Polyklonalität in IMGT/HighV-Quest-Datensätzen verglichen. Im Folgenden werden die NGS-Befunde der einzelnen Proben kurz dargestellt.

Die CLL-Probe S17 zeigte ein Verteilungsmuster der IGH-Sequenzen mit schwachem Gefälle innerhalb der zehn sequenzstärksten Clontypes (AA) und

ohne starke Überrepräsentation eines einzelnen IGH-Rearrangements (Abbildung 23). Auch fanden sich in den Datensätzen FR1 und FR2 jeweils andere Rearrangements auf den Rängen 1 bis 10, sodass sich in den NGS-Datensätzen kein Hinweis auf Klonalität ergab. Den Clonotypes (AA) auf Rang 1 wurden jeweils 38,9 % (Datensatz FR1) und 32,5 % (Datensatz FR2) der Sequenzen zugeordnet.

Probe S17, CLL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 17.596, FR2 – 4.274

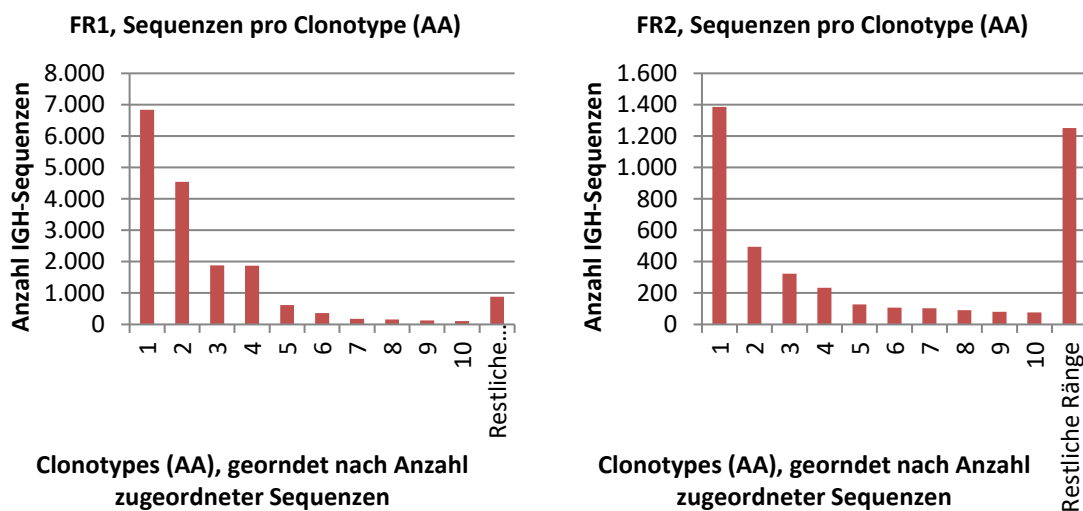


Abbildung 23: Probe S17, CLL; Anzahl zugeordneter Sequenzen pro Clonotype (AA). Es besteht ein schwaches Gefälle zwischen den Clonotypes (AA) der Ränge 1 bis 10, die in den Datensätzen FR1 und FR2 jeweils von anderen IGH-Rearrangements belegt werden.

Die FL-Probe S18 zeigte ein polyklonales Verteilungsmuster, sowohl in der Zuordnung der IGH-Sequenzen nach Sequenzlänge als auch nach Clonotypes (AA) in den Datensätzen FR1 und FR2 (Abbildung 24). In der Zuordnung nach Clonotypes (AA) wurden den zehn Sequenzstärksten Clonotypes (AA) nur kleine Anteile der Gesamtsequenzzahl zugeordnet, wohingegen der Großteil der Sequenzen den Rängen jenseits von Rang 10 zugeordnet wurde.

Die FL-Probe S19 zeigte kein klonales Verteilungsmuster mit einem schwachen Gefälle innerhalb der zehn sequenzstärksten Clonotypes (AA) ohne starke Überrepräsentation eines einzelnen IGH-Rearrangements (Abbildung 25). Es fanden sich in den beiden Datensätzen FR1 und FR2 jeweils andere IGH-Rearrangements auf den jeweiligen Rängen.

Probe S18, FL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 2.616, FR2 – 7.406

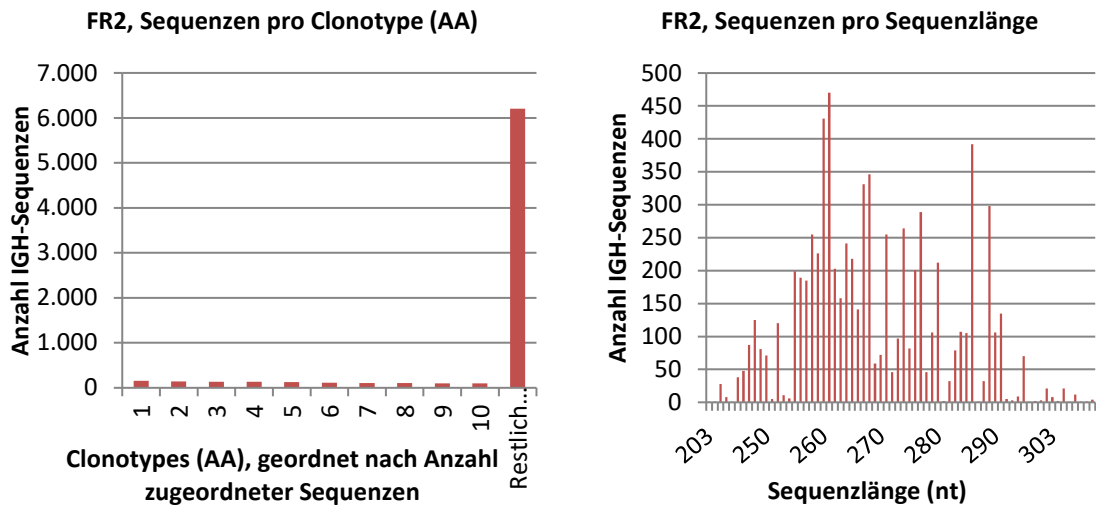


Abbildung 24: Probe S18, FL, Datensatz FR2; Anzahl zugeordneter Sequenzen pro Sequenzlänge und pro Clonotype (AA). In beiden Darstellungen des Datensatzes mit dem Primerset FR2 zeigt sich ein polyklonales Verteilungsmuster der IGH-Sequenzen.

Probe S19, FL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 3.981, FR2 – 3.451

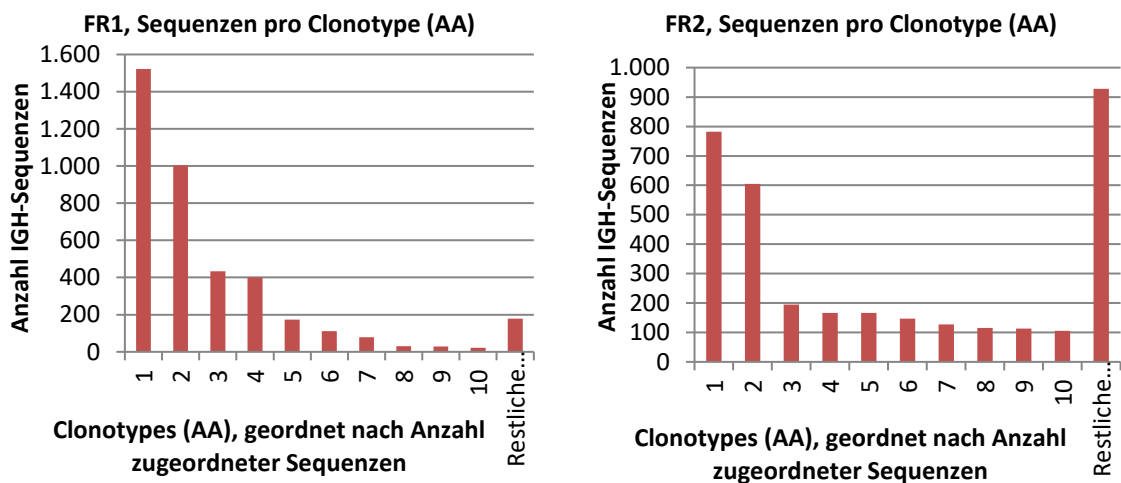


Abbildung 25: Probe S19, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA). Es besteht ein schwaches Gefälle zwischen den Sequenzzahlen, die den Clonotypes (AA) der Ränge 1 bis 10 zugeordnet wurden. Sie werden in den Datensätzen FR1 und FR2 jeweils von anderen IGH-Rearrangements belegt.

Die FL-Probe S20 zeigte ein polyklonales Verteilungsmuster, sowohl in der Zuordnung der IGH-Sequenzen nach Sequenzlänge als auch nach Clonotypes (AA) in den Datensätzen FR1 und FR2 (Abbildung 26). In der Zuordnung nach Clonotypes (AA) wurden den zehn Sequenzstärksten Clonotypes (AA) nur kleine Anteile der Gesamtsequenzzahl zugeordnet, wohingegen der Großteil der Sequenzen den Rängen jenseits von Rang 10 zugeordnet wurde. Auf Rang 2 fand sich in beiden Datensätzen das gleiche IGH-Rearrangement.

Probe S20, FL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 12.056, FR2 – 7.619

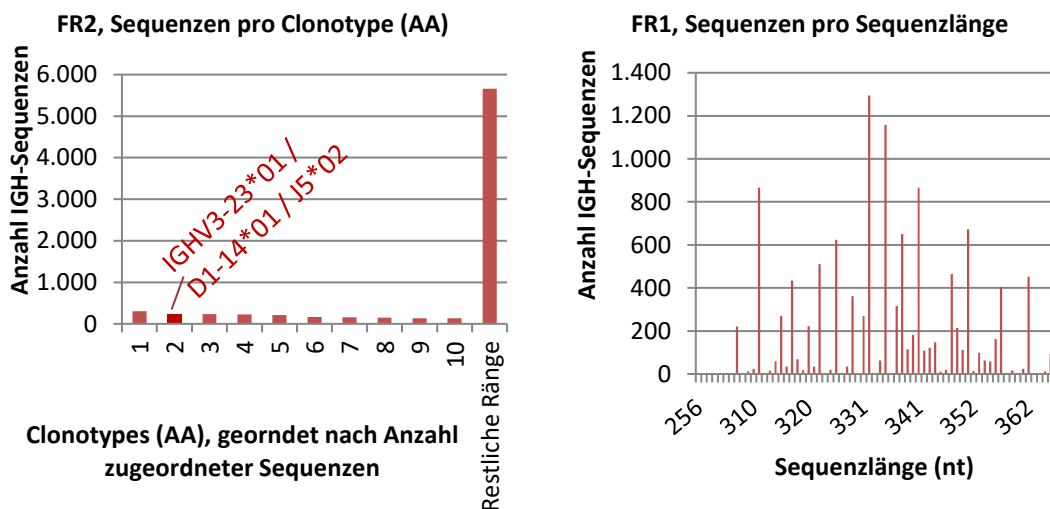


Abbildung 26: ProbeS20, FL; Anzahl zugeordneter Sequenzen pro Sequenzlänge, Datensatz FR1, sowie Anzahl zugeordneter Sequenzen pro Clonotype, Datensatz FR2. Das in Rot hervorgehobene IGH-Rearrangement fand sich in beiden Datensätzen auf Rang 2.

In der FL-Probe S21 fand sich in beiden Datensätzen in der Zuordnung der IGH-Sequenzen nach Clonotype (AA) auf Rang 1 das gleiche Rearrangement, dem deutlich mehr Sequenzen zugeordnet wurden als dem Clonotype (AA) auf Rang 2. Der Großteil der Sequenzen wurde Clonotypes (AA) jenseits von Rang 10 zugeordnet (Abbildung 27). In der Zuordnung nach Sequenzlänge zeigte sich in beiden Datensätzen ein Peak auf Höhe der Sequenzlänge des Clonotypes (AA) auf Rang 1 vor einer polyklonalen Hintergrundpopulation.

Probe S21, FL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 14.275, FR2 – 4.699

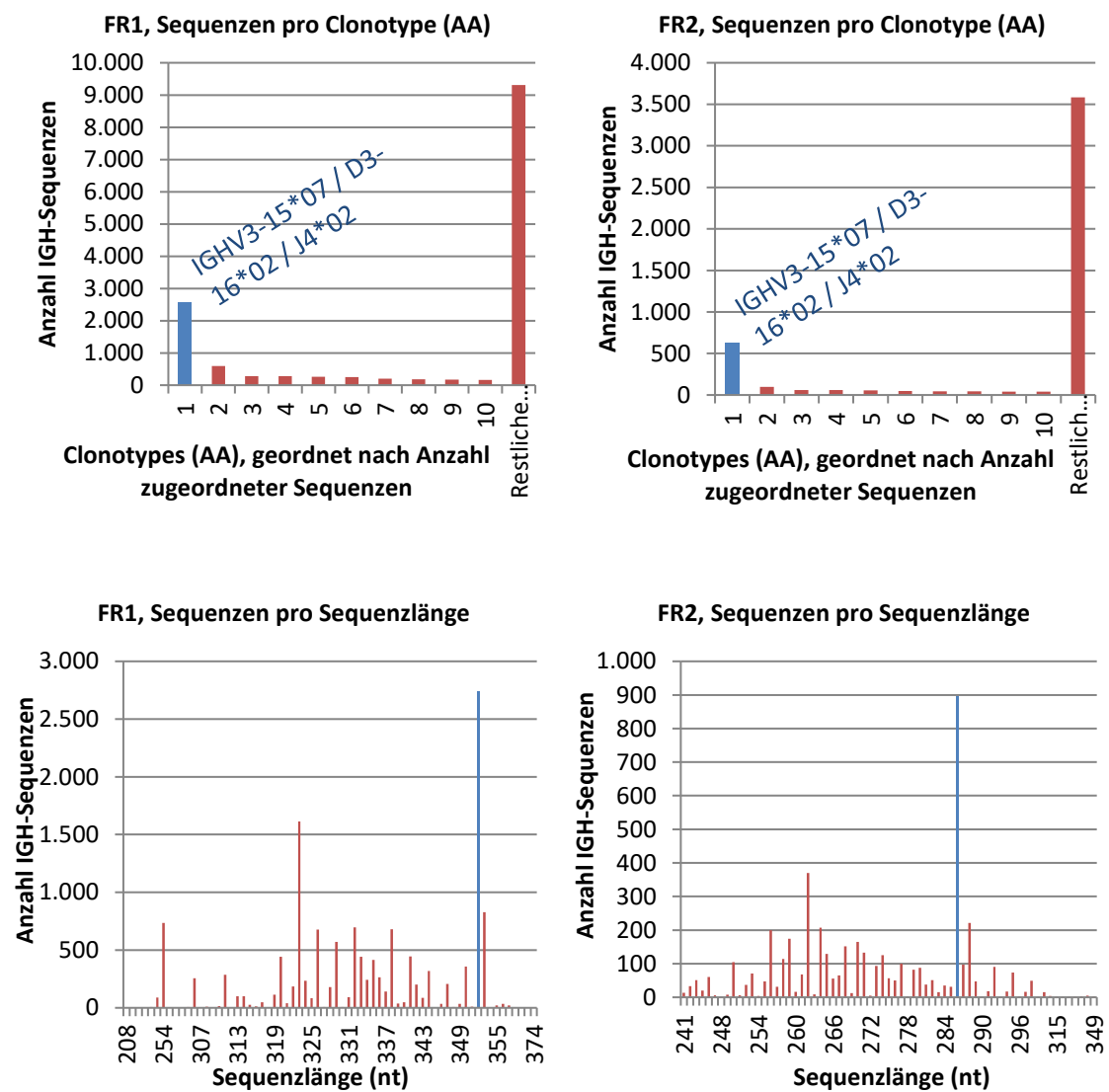


Abbildung 27: Probe S21, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) und pro Sequenzlänge. In der Zuordnung der IGH-Sequenzen nach Clonotype (AA) zeigt sich auf Rang 1 (blau) das gleiche IGH-Rearrangement mit deutlicher Absetzung von den Clonotypes (AA) ab Rang 2 (rot). In der Zuordnung nach Sequenzlänge zeigt sich ein Peak der Sequenzlänge des Clonotypes (AA) auf Rang 1 (blau) vor einem polyklonalen Hintergrund (rot).

In den Datensätzen der FL-Probe S22 fand sich in der Zuordnung nach Clonotypes (AA) auf Rang 1 das gleiche IGH-Rearrangement, dem jeweils der Großteil der Sequenzen zugeordnet wurden (Abbildung 28). Unter den zehn sequenzstärksten Clonotypes (AA) fanden sich außerdem Varianten des Rearrangements auf Rang 1.

Probe S22, FL

Gesamtzahl einem Clonotype (AA) zugeordneter
Sequenzen: FR1 – 14.938, FR2 – 5.040

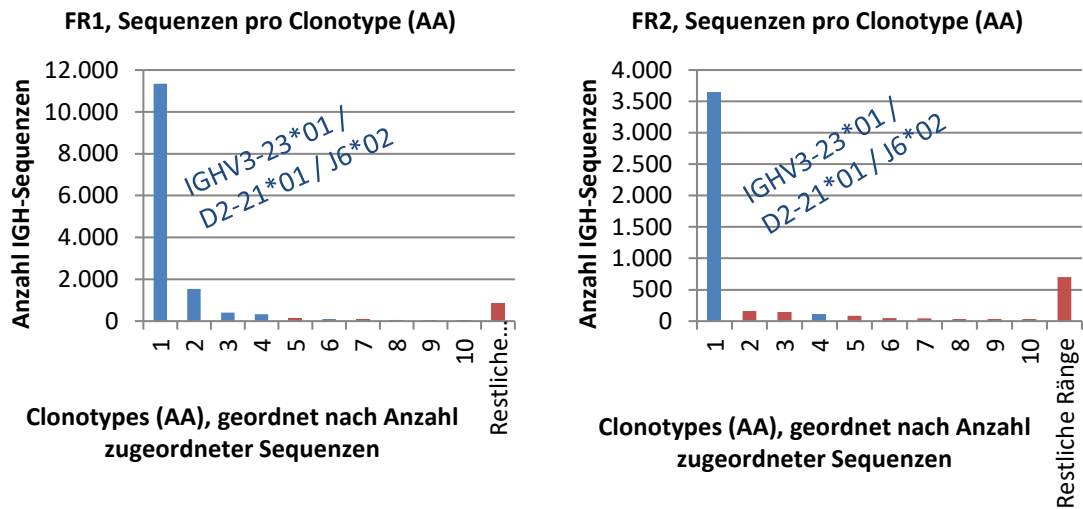


Abbildung 28: Probe S22, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) für die Datensätze FR1 und FR2. In beiden Datensätzen zeigen sich ein prominentes IGH-Rearrangement und seine Varianten (blau) bei schwacher Hintergrundpopulation (rot).

Der Datensatz FR2 der ABC-DLBCL-Probe S24 zeigte ein monoklonales Verteilungsmuster der Sequenzen bei Zuordnung nach Sequenzlänge, sowie eine deutliche Überrepräsentation des Clonotypes (AA) auf Rang 1 (Abbildung 29). Die Clonotypes (AA) der Ränge 3 bis 10 stellten Varianten des Rearrangements auf Rang 1 dar. Der Datensatz FR1 enthielt nur neun IGH-Sequenzen, unter denen sich das prominente Rearrangement aus dem Datensatz FR2 nicht befand.

Im Datensatz FR2 der ABC-DLBCL-Probe S26 wurde der Großteil der IGH-Sequenzen dem Clonotype (AA) auf Rang 1 sowie seinen Varianten zugeordnet; in der Zuordnung nach Sequenzlänge im Datensatz FR2 zeigte sich ein monoklonales Verteilungsmuster (Abbildung 30). Der Datensatz FR1 lieferte nur zwei IGH-Sequenzen.

Probe S24, ABC DLBCL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 9, FR2 – 7.736

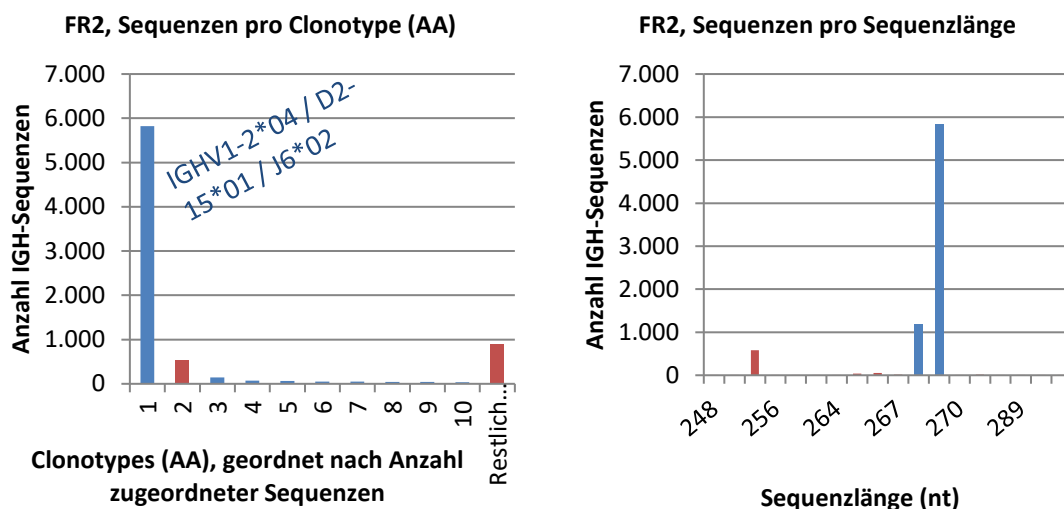


Abbildung 29: Probe S24, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) sowie pro Sequenzlänge, Datensatz FR2. Es zeigt sich eine Überrepräsentation des IGH-Rearrangements auf Rang 1 und seiner Varianten (blau), sowohl in der Zuordnung nach Clonotype (AA) als auch nach Sequenzlänge.

Probe S26, ABC DLBCL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 2, FR2 – 6.867

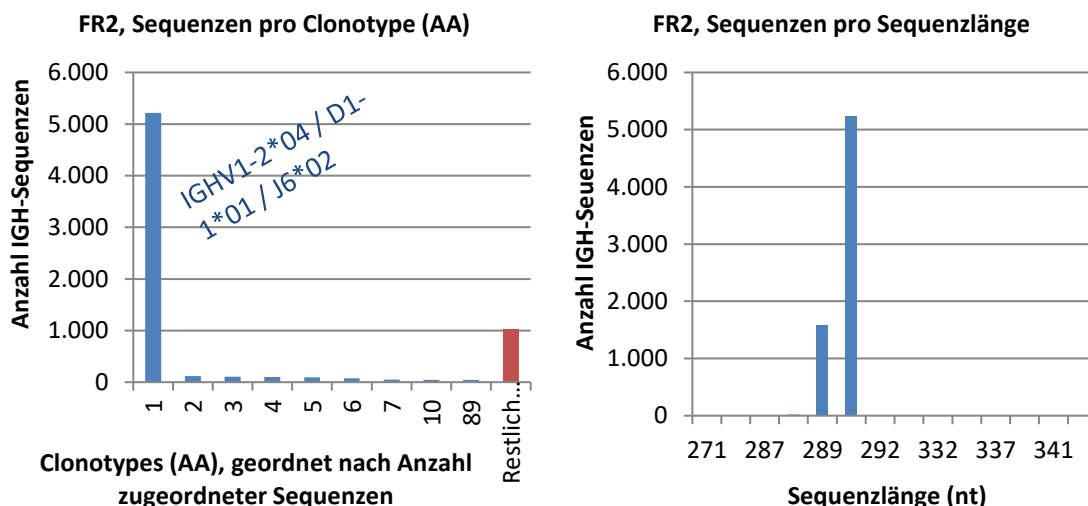


Abbildung 30: Probe S26, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Sequenzlänge sowie pro Clonotype (AA), Datensatz FR2. Es zeigt sich eine Überrepräsentation des IGH-Rearrangements auf Rang 1 und seiner Varianten (blau), sowohl in der Zuordnung nach Clonotype (AA) als auch nach Sequenzlänge.

Die ABC-DLBCL-Probe S27 zeigte im Datensatz FR1 eine Überrepräsentation des Clonotypes (AA) auf Rang 1 und seiner Varianten, im Datensatz FR2 hingegen ein schwaches Gefälle zwischen den Clonotypes (AA) der Ränge 1 bis 10 (Abbildung 31). Das prominente Rearrangement aus dem Datensatz FR1 fand sich nicht unter ihnen. Auch die Zuordnung nach Sequenzlänge lieferte ein klonales Muster für den Datensatz FR1 und ein polyklonales Muster für den Datensatz FR2.

Probe S27, ABC DLBCL

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 1.210, FR2 – 7.269

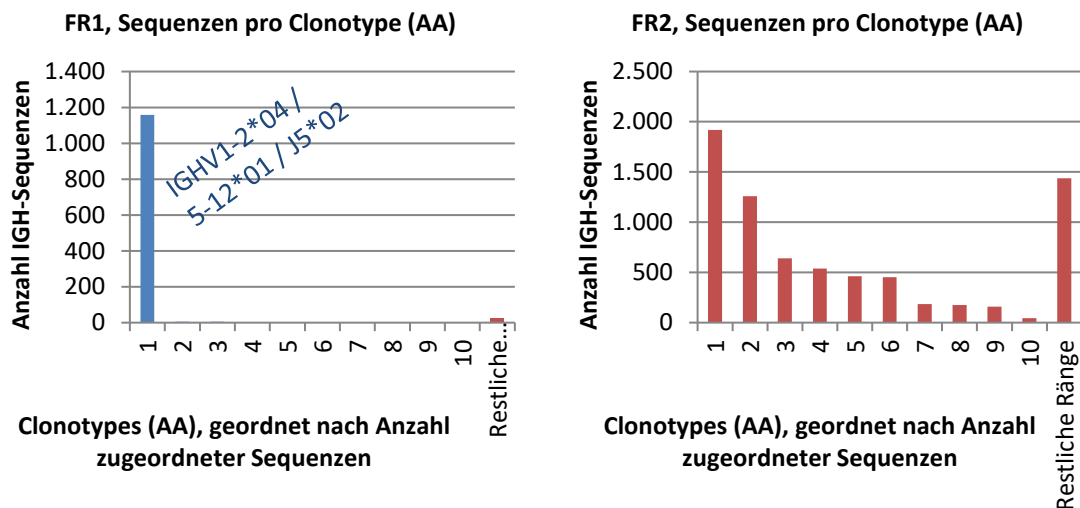


Abbildung 31: Probe S27, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Clonotype (AA). Im Datensatz FR1 zeigen sich ein überrepräsentiertes IGH-Rearrangement und seine Varianten (blau) vor einer schwachen Hintergrundpopulation (rot). Im Datensatz FR2 ergibt sich eine gleichmäßigere Verteilung auf die Clonotypes (AA), ohne Repräsentation des prominenten Rearrangements aus dem Datensatz FR1.

4.4.4 Vergleich der NGS-Daten mit Befunden der GeneScan-Analyse

Die Patientenproben mit hohem Lymphomzellgehalt wurden neben der Sanger-Sequenzierung des klonalen Rearrangements und der Untersuchung mittels NGS auch einer GeneScan-Analyse unterzogen und die Ergebnisse der drei Modalitäten miteinander verglichen (Tabelle 19).

Ergebnisse

Tabelle 19: Überblick über die Ergebnisse von GeneScan-Analyse und Klonalitätsbefund in NGS-Datensätzen sowie die Sequenzierbarkeit nach Sanger in Patientenproben mit hohem Lymphomzellgehalt. k.a.: kaum amplifiziert und daher nicht verwertbar (siehe auch Tabelle 16).

Probe	Diagnose	Sanger-Sequenzierung		NGS	
		erfolgreich?	GeneScan	Datensatz FR1	Datensatz FR2
S13	CLL	ja	pseudoklonal	klonal	klonal
S14	CLL	ja	klonal	klonal	klonal
S15	CLL	ja	klonal	klonal	klonal
S16	CLL	ja	klonal	klonal	(klonal)
S17	CLL	nein	klonal	polyklonal	polyklonal
S18	FL	nein	polyklonal	polyklonal	polyklonal
S19	FL	nein	polyklonal	polyklonal	polyklonal
S20	FL	nein	klonal	polyklonal	polyklonal
S21	FL	nein	klonal	klonal	klonal
S22	FL	nein	polyklonal	klonal	klonal
S23	ABC DLBCL	ja	pseudoklonal	klonal	klonal
S24	ABC DLBCL	nein	pseudoklonal	k.a.	klonal
S25	ABC DLBCL	ja	klonal	klonal	klonal
S26	ABC DLBCL	nein	pseudoklonal	k.a.	klonal
S27	ABC DLBCL	nein	pseudoklonal	klonal	polyklonal

Vier der Proben mit gelungener Sanger-Sequenzierung des klonalen Rearrangements wiesen sowohl in der GeneScan-Analyse als auch in den NGS-Datensätzen ein klonales Bild auf (S14, S15, S16, S25), zwei wiesen ein pseudoklonales Bild im GeneScan bei klonalem Bild in den NGS-Datensätzen auf (S13, S25).

Die Proben ohne gelungene Sanger-Sequenzierung des klonalen Rearrangements präsentierten sich insgesamt heterogen. Eine Probe stellte sich mittels GeneScan sowie mittels NGS klonal dar (S21), zwei Proben präsentierten sich im GeneScan sowie in beiden NGS-Datensätzen polyklonal (S18, S19).

In den übrigen sechs Proben ergaben sich Diskrepanzen zwischen GeneScan-Analyse und NGS-Befund. Zwei Proben ergaben ein klonales Bandenmuster mittels GeneScan, aber ein polyklonales Muster mittels NGS (S17, S20). Eine Probe präsentierte sich in der GeneScan-Analyse polyklonal, in beiden NGS-Datensätzen hingegen klonal (S22). Zwei Proben ergaben ein pseudoklonales Muster mittels GeneScan bei klonalem Bild in den NGS-Datensätzen FR2 (S24,

S26), wobei die Datensätze FR1 in beiden Fällen kaum Sequenzen aufwiesen (sechs bzw. zwei Sequenzen), sodass diese nicht auswertbar waren. Eine weitere Probe präsentierte sich in der GeneScan-Analyse ebenfalls pseudoklonal, wies im Datensatz FR1 jedoch ein klonales, im Datensatz FR2 ein polyklonales Muster auf (S27).

4.5 NGS-basierte Klonalitätsuntersuchung bei Verlaufspuben

Bei der Anwendung von NGS zur Klonalitätsanalyse von Verlaufspuben lag ein besonderes Augenmerk auf der Detektierbarkeit klonaler IGH-Rearrangements und damit maligner Lymphozytenpopulationen im zeitlichen Verlauf. Auch wurde untersucht, ob sich Veränderungen von IGH-Rearrangements auf Aminosäure- und Nukleotidebene durch SHM oder Artefakte des Amplifikations- und Sequenzierprozesses darstellen und nachverfolgen lassen.

4.5.1 Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen

Auch für dieses Probenkollektiv wurden zunächst die Anzahlen mittels NGS gewonnener Sequenzen, die Anzahlen der als IGH-Rearrangement erkannten Sequenzen, sowie die Anzahl verschiedener in einer Probe detektierter IMGT Clonotypes (AA) betrachtet (Tabelle 20).

Drei Proben (S32, S33 und S41) hatten mit dem Primerset FR1 bereits gelelektrophoretisch keine Bande geliefert, sodass die entsprechenden Ansätze nicht sequenziert und nicht für die Ermittlung der folgenden Größen berücksichtigt wurden.

Ergebnisse

Tabelle 20: Anzahl mittels NGS gewonnener und einem IMGT Clonotype (AA) zugeordneter Sequenzen sowie Anzahl verschiedener IMGT Clontypes (AA) in Verlaufsproben.

Patient	Probe	Diagnose	Primer set	Anzahl mittels NGS gewonnener Sequenzen	Anzahl einem IMGT Clonotype (AA) zugeordneter Sequenzen	Anteil IGH-Sequenzen an mittels NGS gewonnenen Sequenzen	Anzahl verschiedener IMGT Clontypes (AA)
P16	1 (S28)	NHL	FR1	5.332	2.041	38,3%	92
			FR2	34.542	31.109	90,1%	1.664
	2 (S29)	DLBCL	FR1	17.856	17.287	96,8%	242
			FR2	34.625	33.594	97,0%	1.563
	3 (S30)	DLBCL	FR1	28.872	17.658	61,2%	241
			FR2	29.365	24.022	81,8%	1.227
P17	1 (S31)	GC DLBCL	FR1	24.727	14.912	60,3%	76
			FR2	15.500	71	0,5%	19
	2 (S32)	GC DLBCL	FR1	-	-	-	-
			FR2	23.140	21.257	91,9%	1.190
	3 (S33)	GC DLBCL	FR1	-	-	-	-
			FR2	3.809	19	0,5%	11
P18	1 (S34)	CLL	FR1	20.995	19.968	95,1%	435
			FR2	32.082	24.615	76,7%	150
	2 (S35)	CLL	FR1	21.750	21.400	98,4%	475
			FR2	35.064	27.402	78,1%	84
	3 (S36)	CLL	FR1	20.191	19.495	96,6%	493
			FR2	32.698	25.938	79,3%	83
P19	1 (S37)	FL	FR1	23.814	22.928	96,3%	497
			FR2	34.219	31.380	91,7%	1.376
	2 (S38)	FL	FR1	22.600	21.815	96,5%	484
			FR2	38.395	35.725	93,0%	1.777
	3 (S39)	FL	FR1	33.334	30.534	91,6%	740
			FR2	31.189	26.178	83,9%	1.943
P20	1 (S40)	FL	FR1	28.297	17.718	62,6%	338
			FR2	30.971	19.798	63,9%	1.333
	2 (S41)	FL	FR1	-	-	-	-
			FR2	3.509	3.310	94,3%	324
	3 (S42)	FL	FR1	37.594	36.185	96,3%	731
			FR2	32.893	30.052	91,4%	2.037
P21	1 (S43)	FL	FR1	42.143	32.644	77,5%	303
			FR2	30.421	19.924	65,5%	1.232
	2 (S44)	FL	FR1	38.836	38.099	98,1%	564
			FR2	32.752	10.482	32,0%	497
	3 (S45)	FL	FR1	43.534	6.046	13,9%	238
			FR2	30.581	5.127	16,8%	136

Mit dem Primerset FR1 wurden im Durchschnitt 27.325 Sequenzen gewonnen (Minimum 5.332, Maximum 43.534). Die Anzahl einem IMGT Clonotype zugeordneter Sequenzen betrug im Schnitt 21.249 (Minimum 2.041, Maximum 38.099). Damit betrug der Anteil der als IGH-Rearrangement erkannten Sequenzen an der Anzahl analysierter Sequenzen durchschnittlich 78,6 % (Minimum 13,9 %, Maximum 98,4 %). Die Anzahl verschiedener IMGT Clonotypes (AA) pro Datensatz lag im Schnitt bei 397 (Minimum 76, Maximum 740).

Das Primerset FR2 lieferte durchschnittlich 31.686 Sequenzen (Minimum 3.509, Maximum 38.395) von denen im Schnitt 20.556 als IGH-Rearrangement erkannt wurden (Minimum 19, Maximum 35.725). Dies entsprach einem Anteil von durchschnittlich 69,5 % (Minimum 0,5 %, Maximum 97 %). Die mittlere Anzahl verschiedener IMGT Clonotypes (AA) pro Datensatz betrug 1.008 (Minimum 11, Maximum 2.037).

Damit lieferte das Primerset FR2 in diesem Probenkollektiv im Durchschnitt mehr Sequenzen als das Primerset FR1, für das allerdings der Anteil der als IGH-Rearrangement erkannten Sequenzen größer war. Dieser Vergleich ist jedoch aufgrund der drei nicht amplifizierten Ansätze mit dem Primerset FR1 nur eingeschränkt aussagekräftig. Wie bereits bei den Zelllinienverdünnungsproben, sowie den Patientenproben mit hohem Lymphomzellgehalt beobachtet, wurden in Ansätzen mit dem Primerset FR2 mehr verschiedene IMGT Clonotypes (AA) definiert als in Ansätzen mit dem Primerset FR1.

4.5.2 Anteile der sequenzstärksten IGH-Rearrangements an der Gesamtzahl der IGH-Sequenzen und Detektierbarkeit im zeitlichen Verlauf

Nun wurden wiederum die zehn sequenzstärksten Clonotypes (AA) innerhalb der Datensätze jedes Patienten betrachtet, ihre Anteile an der Gesamtzahl der IGH-Sequenzen ermittelt und untersucht, ob ein IGH-Rearrangement oder ihm zuzuordnende Varianten probenübergreifend in mehreren Datensätzen repräsentiert waren. Aus Gründen der Übersicht wurden im Folgenden ein

Rearrangement „X“ und seine Varianten als „Population X“ zusammengefasst. Als Population wurde dabei eine Gruppe von Clonotypes (AA) betrachtet, die eine hohe Ähnlichkeiten zueinander aufwiesen, die klonale Verwandtschaft nahelegten. War eine Rearrangement-Population in einem der Datensätze als sequenzstärkster Clonotype (AA) auf Rang 1 vertreten, so wurden alle übrigen Datensätze auf Rearrangements dieser Population hin untersucht und ihr Anteil an der Gesamtzahl der detektierten IGH-Rearrangements in jedem Datensatz in Prozent in den Ergebnistabellen angegeben.

In den Proben der Patienten 16, 18 und 20 ließen sich auf diese Art Rearrangements einer Population ermitteln, die in allen drei Proben unter den zehn sequenzstärksten Clonotypes (AA) zu finden waren und daher als Rearrangements und Rearrangement-Varianten des bekannten Lymphoms interpretiert werden konnten. Im Fall von Patient 19 war das Rearrangement der Population A in der dritten Probe nicht mehr unter den zehn sequenzstärksten Clonotypes (AA) vertreten, aber dennoch mit wenigen zugeordneten Sequenzen in dem Datensatz enthalten. Die Verlaufspalten der Patienten 17 und 21 lieferten kein eindeutiges Bild im Hinblick auf Klonalität und es gelang nicht der Nachweis einer in allen Proben repräsentierten Rearrangement-Population. Im Folgenden werden die Ergebnisse der einzelnen Patienten dargestellt.

Patient 16: DLBCL

In den Datensätzen von Patient 16 (Tabelle 21) ließ sich eine Rearrangement-Population ausmachen, die in allen Datensätzen unter den zehn sequenzstärksten Clonotypes (AA) vertreten war, und in vier davon den Clonotype (AA) auf Rang 1 stellte (Probe 1 – FR2, Probe 2 – FR2 sowie Probe 3 – FR1 und FR2). In den Datensätzen des Primersets FR1 der Proben 1 und 2 fand sich jeweils ein anderes IGH-Rearrangement auf Rang 1, das jedoch weder in den Datensätzen der jeweiligen Probe mit dem Primerset FR2 vertreten war, noch in einem Datensatz einer der anderen Proben.

Ergebnisse

Tabelle 21: Übersicht der NGS-Ergebnisse der Verlaufspuren des Patienten 16 nach IMGT/HighV-Quest-Auswertung; Anzahl und Anteile der IGH-Sequenzen, die den sequenzstärksten Rearrangement-Populationen bzw. der Hintergrundpopulation zugeordnet wurden. Die top 10 Clonotypes (AA) des jeweiligen Datensatzes wurden auf Repräsentation in anderen Datensätzen des Patienten überprüft. Populationen, die in mindestens einem Datensatz den sequenzstärksten Clonotype (AA) stellten, wurden in der Tabelle separat dargestellt. Nur Population A (blau) ist in allen Datensätzen repräsentiert und stellt in vier Datensätzen den sequenzstärksten Clonotype (AA). Unter „Referenz-Clonotype (AA)“ ist für jede aufgeführte Population ein ihr zugehöriger Clonotype (AA) dargestellt, alle übrigen Clonotypes der Population wiesen zu diesem große Ähnlichkeit auf.

Probe	Primerset	Gesamtzahl einem IMGT- Clonotype (AA) zugeordneter Sequenzen	Sequenzstarke Rearrangement-Populationen unter den top 10 Clonotypes (AA) pro Datensatz: Anzahl zugeordneter IGH- Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen				Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
			Population A	Population B	Population C	Hintergrund- population	
1	FR1	2.041 100,0%	495 24,3%	1.438 70,5%	0 0,0%	9 0,4%	99 4,9%
	FR2	31.109 100,0%	17.938 57,7%	0 0,0%	0 0,0%	6.475 20,8%	6696 21,5%
2	FR1	17.287 100,0%	4.681 27,1%	0 0,0%	11.891 68,8%	0 0,0%	715 4,1%
	FR2	33.594 100,0%	27.464 81,8%	0 0,0%	0 0,0%	0 0,0%	6130 18,2%
3	FR1	17.658 100,0%	4.909 27,8%	0 0,0%	0 0,0%	12.100 68,5%	649 3,7%
	FR2	24.022 100,0%	17.728 73,8%	0 0,0%	0 0,0%	1.532 6,4%	4762 19,8%
Referenz-Clonotype (AA)		V-Gen und - Allel	IGHV3-7*01 F	IGHV4-31*03 F	IGHV4-34*01 F		
		D-Gen und - Allel	IGHD2-15*01 F	IGHD1-26*01 F	IGHD6-13*01 F		
		J-Gen und - Allel	IGHJ5*02 F	IGHJ4*02 F	IGHJ4*02 F		
		CDR3-IMGT Länge	15 AA	14 AA	17 AA		
		AA-Sequenz der CDR3	ASLDPNIVVDN WFDP	ARDTFYGGATG FDY	ARARRSGGAV AGTDFDY		

Patient 17: GCB DLBCL

In den Verlaufspuren von Patient 17 ließ sich kein Rearrangement darstellen, das in allen Proben sequenzstark vertreten war. Im Datensatz FR1 der Probe zum Zeitpunkt 1 (S31) lag ein prominentes Rearrangement vor, das allerdings in keinem anderen Datensatz des Patienten repräsentiert war. Das gleiche galt für ein sequenzstarkes Rearrangement im Datensatz FR2 der Probe zum Zeitpunkt 2 (S32). In den Datensätzen FR2 der Proben zum Zeitpunkt 1 (S31) und Zeitpunkt 3 (S33) waren nur jeweils 0,5 % der mittels NGS gewonnenen Sequenzen als IGH-Rearrangement erkannt worden (entsprechend 19

Sequenzen für S31 – FR2 und elf Sequenzen für S33 – FR2, Tabelle 20). Auch hier fanden sich keine miteinander übereinstimmenden Clonotypes (AA). Die Proben zum Zeitpunkt 2 (S32) und 3 (S33) hatten in der Multiplex-IGH-PCR mit dem Primerset FR1 keine Bande ergeben, weswegen keine entsprechenden NGS-Datensätze zur Analyse vorlagen.

Patient 18: CLL

In den Datensätzen der Verlaufspuben von Patient 18 (Tabelle 22) ließ sich eine Rearrangement-Population nachweisen, die in allen Datensätzen unter den zehn sequenzstärksten Clonotypes (AA) vertreten war und den Clonotype (AA) auf Rang 1 stellte. Ihr wurden mindestens 89,0 % aller als IGH-Rearrangement erkannten Sequenzen innerhalb jedes Datensatzes zugeordnet.

Patient 19: FL

In den Datensätzen der Verlaufspuben von Patient 19 (Tabelle 23) ließ sich keine Rearrangement-Population darstellen, die in allen Datensätzen unter den zehn sequenzstärksten Clonotypes (AA) vertreten war. Jedoch fand sich eine Population, die in den Datensätzen der Proben 1 und 2 den Clonotype (AA) auf Rang 1 stellte, und die im Datensatz mit dem Primerset FR1 der Probe 3 jenseits der zehn sequenzstärksten Clonotypes (AA) mit 7 zugeordneten Sequenzen repräsentiert war. Die beiden Populationen, die in den Datensätzen der Probe 3 jeweils den sequenzstärksten Clonotype (AA) stellten, waren prozentual eher sequenzschwach und fanden sich nicht in den Datensätzen der Proben 1 und 2 wieder.

Ergebnisse

Tabelle 22: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 18 nach IMGT/HighV-Quest-Auswertung; Anzahl und Anteile der IGH-Sequenzen, die den sequenzstärksten Rearrangement-Populationen bzw. der Hintergrundpopulation zugeordnet wurden. Die top 10 Clonotypes (AA) des jeweiligen Datensatzes wurden auf Repräsentation in anderen Datensätzen des Patienten überprüft. Population A stellt in allen Datensätzen den sequenzstärksten Clonotype (AA). Unter „Referenz-Clonotype (AA)“ ist für die aufgeführte Population ein ihr zugehöriger Clonotype (AA) dargestellt, alle übrigen Clonotypes der Population wiesen zu diesem große Ähnlichkeit auf.

Probe	Primer-Set	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Sequenzstarke Rearrangement-Populationen unter den top 10 Clonotypes (AA) pro Datensatz: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen		Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
			Population A	Hintergrundpopulation	
1	FR1	19.968 100,0%	18.566 93,0%	0 0,0%	1.402 7,0%
	FR2	24.615 100,0%	23.844 96,9%	285 1,2%	486 2,0%
2	FR1	21.400 100,0%	19.782 92,4%	0 0,0%	1.618 7,6%
	FR2	27.402 100,0%	26.845 98,0%	222 0,8%	335 1,2%
3	FR1	19.495 100,0%	17.347 89,0%	504 2,6%	1.644 8,4%
	FR2	25.938 100,0%	25.533 98,4%	0 0,0%	405 1,6%
Referenz-Clonotype (AA)		V-Gen und -Allel	IGHV1-69*13 (F)		
		D-Gen und -Allel	IGHD3-3*01 F		
		J-Gen und -Allel	IGHJ6*02 F		
		CDR3-IMGT Länge	26 AA		
		AA-Sequenz der CDR3	ASPSIGRDYDFWSGYYPNYYY YGMDV		

Ergebnisse

Tabelle 23: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 19 nach IMGT/HighV-Quest-Auswertung; Anzahl und Anteile der IGH-Sequenzen, die den sequenzstärksten Rearrangement-Populationen bzw. der Hintergrundpopulation wurden. Die top 10 Clonotypes (AA) des jeweiligen Datensatzes wurden auf Repräsentation in anderen Datensätzen des Patienten überprüft. Populationen, die in mindestens einem Datensatz den sequenzstärksten Clonotype (AA) stellten, wurden in der Tabelle separat dargestellt. Nur Population A (blau) ist in allen Proben repräsentiert. In den Datensätzen der Proben 1 und 2 stellt sie den sequenzstärksten Clonotype (AA). In Probe 3 ist sie nicht unter den zehn sequenzstärksten Clonotypes, sondern nur sequenzschwach („low level“) vertreten. Unter „Referenz-Clonotype (AA)“ ist für jede aufgeführte Population ein ihr zugehöriger Clonotype (AA) dargestellt, alle übrigen Clonotypes der Population wiesen zu diesem große Ähnlichkeit auf.

Probe	Primer-Set	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Sequenzstarke Rearrangement-Populationen unter den top 10 Clonotypes (AA) pro Datensatz: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen			Hintergrundpopulation	Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
			Population A	Population B	Population C		
1	FR1	22.928 100,0%	17.450 76,1%	0 0,0%	0 0,0%	3.839 16,7%	1.639 7,1%
	FR2	31.380 100,0%	24.914 79,4%	0 0,0%	0 0,0%	388 1,2%	6.078 19,4%
2	FR1	21.815 100,0%	14.342 65,7%	0 0,0%	0 0,0%	6.098 28,0%	1.375 6,3%
	FR2	35.725 100,0%	25.361 71,0%	0 0,0%	0 0,0%	2.382 6,7%	7.982 22,3%
3	FR1	30.534 100,0%	„low level“ (7)	5.391 17,7%	0 0,0%	18.555 60,8%	6.588 21,6%
	FR2	26.178 100,0%	0 0,0%	0 0,0%	3.618 13,8%	13.244 50,6%	9.316 35,6%
Referenz-Clonotype (AA)	V-Gen und - Allel		IGHV1-69*13 (F)	IGHV3-48*02 F	IGHV3-15*07 F		
	D-Gen und - Allel		IGHD3-22*01 F	IGHD1-1*01 F	IGHD3-22*01 F		
	J-Gen und - Allel		IGHJ3*02 F	IGHJ4*02 F	IGHJ4*02 F		
	CDR3-IMGT Länge		17 AA	10 AA	20 AA		
	AA-Sequenz der CDR3		VRAVLNSNGFFSA AFDI	ARSLLGPLDY	TTSLSIGSPHLSY HYSPPDY		

Patient 20: FL

In den Datensätzen der Verlaufsproben von Patient 20 (Tabelle 24) ließ sich eine Rearrangement-Population darstellen, die sich in allen Datensätzen unter den zehn sequenzstärksten Clonotypes (AA) wiederfand. In beiden Datensätzen der Probe 3 stellte diese Population den sequenzstärksten Clonotype (AA). Weitere Populationen stellten in nur je einem Datensatz den sequenzstärksten Clonotype (AA) und fanden sich in keinem anderen Datensatz unter den top zehn

Clonotypes (AA) wieder. Für Probe 2 liegt nur der Datensatz FR2 vor, da mit dem Primerset FR1 keine Amplifikation gelungen war.

Tabelle 24: Übersicht der NGS-Ergebnisse der Verlaufspalten des Patienten 20 nach IMGT/HighV-Quest-Auswertung; Anzahl und Anteile der IGH-Sequenzen, die den sequenzstärksten Rearrangement-Populationen bzw. der zugeordnet wurden. Die top 10 Clonotypes (AA) des jeweiligen Datensatzes wurden auf Repräsentation in anderen Datensätzen des Patienten überprüft. Populationen, die in mindestens einem Datensatz den sequenzstärksten Clonotype (AA) stellten, wurden in der Tabelle separat dargestellt. Nur Population A (blau) ist in allen Datensätzen repräsentiert und stellt in den Datensätzen der Probe 3 jeweils den sequenzstärksten Clonotype (AA). Unter „Referenz-Clonotype (AA)“ ist für jede aufgeführte Population ein ihr zugehöriger Clonotype (AA) dargestellt, alle übrigen Clonotypes der Population wiesen zu diesem große Ähnlichkeit auf.

Probe	Primerset	Gesamtzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen	Sequenzstarke Rearrangement-Populationen unter den top 10 Clonotypes (AA) pro Datensatz: Anzahl zugeordneter IGH-Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen					Restliche Ränge: Anzahl zugeordneter Sequenzen und Anteil an Gesamtzahl der IGH-Sequenzen
			Population A	Population B	Population C	Population D	Hintergrundpopulation	
1	FR1	17.718 100,0%	1.539 8,7%	6.353 35,9%	0 0,0%	0 0,0%	9.020 50,9%	806 4,5%
	FR2	19.798 100,0%	731 3,7%	0 0,0%	4.838 24,4%	0 0,0%	10.038 50,7%	4.191 21,2%
2	FR1	- -	- -	- -	- -	- -	- -	- -
	FR2	3.310 100,0%	28 0,8%	0 0,0%	0 0,0%	2.410 72,8%	267 8,1%	605 18,3%
3	FR1	36.185 100,0%	27.579 76,2%	0 0,0%	0 0,0%	0 0,0%	3.835 10,6%	4.771 13,2%
	FR2	30.052 100,0%	16.170 53,8%	0 0,0%	0 0,0%	0 0,0%	4.964 16,5%	8.918 29,7%
Referenz-Clonotype (AA)	V-Gen und - Allel	IGHV3-30*03 F	IGHV5-51*01 F	IGHV3-30-3*01 F	IGHV4-38-2*01 F			
	D-Gen und - Allel	IGHD2-2*01 F	IGHD2-2*01 F	IGHD6-13*01 F	IGHD1-14*01 ORF			
	J-Gen und - Allel	IGHJ4*02 F	IGHJ4*02 F	IGHJ6*02 F	IGHJ4*02 F			
	CDR3-IMGT Länge	20 AA	9 AA	18 AA	9 AA			
	AA-Sequenz der CDR3	AKDGGSDF AGATPQNT LFAY	ARLAVAIDY	ARDHGGIA AAGSYYG MDV	AAGPWAL DF			

Patient 21: FL

In den Verlaufspalten von Patient 21 ließ sich kein prominentes Rearrangement darstellen, das in allen Datensätzen repräsentiert war. Die Proben zu den Zeitpunkten 1 (S43) und 2 (S44) wiesen mit beiden Primersets ein schwaches

Gefälle zwischen den zehn sequenzstärksten Clonotypes (AA) auf. In den Datensätzen der Probe 3 (S45) wurde jeweils ein großer Anteil der Sequenzen dem Clonotype (AA) auf Rang 1 zugeordnet (FR1: 76,1 %, FR2: 82,5 %). Dieser wies allerdings im Datensatz FR1 ein anderes IGH-Rearrangement auf als im Datensatz FR2. Die Betrachtung der mittels NGS gewonnenen und einem IGH-Rearrangement zugeordneten Sequenzen der Probe 3 (S45) hatte bereits gezeigt, dass im Datensatz FR1 nur 13,9 %, im Datensatz FR2 nur 16,8 % der gewonnenen Sequenzen als IGH-Rearrangement erkannt worden waren.

4.5.3 Varianten klonaler IGH-Rearrangements punktuell und im zeitlichen Verlauf

Wie bereits anhand der Zelllinienverdünnungsproben und Patientenproben mit hohem Lymphomzellgehalt im Ansatz gezeigt, ermöglicht NGS eines amplifizierten IGH-Repertoires mit anschließender Analyse durch IMGT/HighV-Quest die Charakterisierung und Unterscheidung verschiedener Varianten eines klonalen IGH-Rearrangements innerhalb einer Rearrangement-Population, die durch SHM oder aber als Artefakt der Fixierung, des Amplifikations- oder Sequenzierprozesses entstanden sein können. Die Möglichkeit, in einer Rearrangement-Population verschiedene Varianten insbesondere der CDR3 auf Aminosäure- und auf Nukleotidebene darzustellen, innerhalb einer Probe sowie probenübergreifend, soll im Folgenden exemplarisch an den Verlaufsproben zweier Patienten veranschaulicht werden. Dabei sollen ein Fall mit hoher V-Region Identity des klonalen Rearrangements von nahezu 100 % (Patient 18) und einer mit geringerer V-Region Identity des klonalen Rearrangements von rund 90 % (Patient 19) dargestellt werden, in der Annahme, dass in unterschiedlichem Maße SHM in den Populationen vorhanden sind. Im Vergleich hierzu soll die Darstellung der Varianten in einer Probe der Zelllinie Granta-519 erfolgen.

Patient 18: CLL, hohe V-Region Identity

In den CLL-Verlaufspuren von Patient 18 konnte ein prominentes IGH-Rearrangement dargestellt werden, das in allen Datensätzen den Clonotype (AA) auf Rang 1 stellte, sowie zahlreiche Varianten des Rearrangements, die sich unter den übrigen Rängen der zehn sequenzstärksten Clonotypes (AA) fanden (siehe Tabelle 22). Dabei glichen sich die verschiedenen Clonotypes (AA), die in Tabelle 22 als „Population A“ zusammengefasst wurden, in ihrem VDJ-Rearrangement und besaßen stark ähnliche CDR3s, die jedoch an einzelnen Positionen Unterschiede in der Aminosäuresequenz aufwiesen. Der Wert der V-Region Identity der Varianten des klonalen Rearrangements betrug dabei zwischen 99 und 100 %. Tabelle 25 zeigt alle Clonotypes (AA) der klonalen Rearrangement-Population, die sich in einem der drei Datensätze mit dem Primerset FR1 unter den zehn sequenzstärksten Clonotypes (AA) befanden, mit der genauen AA-Sequenz ihrer jeweiligen CDR3 und der Anzahl jeweils zugeordneter Sequenzen.

Unter den zehn jeweils sequenzstärksten Clonotypes (AA) der drei Datensätze fanden sich 16 verschiedene Varianten der CDR3 der klonalen Population, die sich jeweils in einer Aminosäure der CDR3 voneinander unterschieden. Jede der Varianten, die zu einem Zeitpunkt einen der zehn sequenzstärksten Clonotypes (AA) stellte, war auch zu beiden anderen Zeitpunkten innerhalb der Probe repräsentiert. Der auf Rang 1 liegende Clonotype (AA) wies dabei stets dieselbe AA-Sequenz der CDR3 auf.

Da die Definition eines Clonotype (AA) im Bereich der CDR3 auf Aminosäureebene erfolgt, ergeben sich wiederum verschiedene Varianten eines Clonotypes (AA) im Bereich der CDR3 auf Nukleotidebene, wenn ein Basenaustausch nicht zur Kodierung einer anderen Aminosäure führt. Tabelle 26 zeigt die CDR3 des in allen drei Datensätzen mit dem Primerset FR1 sequenzstärksten Clonotypes (AA) auf Aminosäureebene sowie die entsprechenden Varianten auf Nukleotidebene, des Datensatzes FR1 der Probe 1 (V-Region Identity 100 %).

Ergebnisse

Tabelle 25: Patient 18, CLL, Primerset FR1. Verschiedene Varianten in der CDR3 des klonalen IGH-Rearrangements und Anzahlen zugeordneter Sequenzen im zeitlichen Verlauf. Alle Varianten der klonalen Rearrangement-Population, die in einem der Datensätze unter den zehn sequenzstärksten Clonotypes (AA) waren, wurden auf Repräsentation in den jeweils anderen Datensätzen überprüft und mit ihrer CDR3 auf AA-Ebene und der Anzahl zugeordneter Sequenzen pro Datensatz aufgeführt. Unterschiede zum sequenzstärksten Clonotype (AA) der Rearrangement-Population (in Blau) auf AA-Ebene sind in Rot dargestellt.

Charakterisierung Clonotype (AA)		Anzahl zugeordneter Sequenzen in Probe...		
VDJ-Rearrangement	AA-Sequenz der CDR3	1	2	3
	ASPSIGRDYDFWSGYYPNYYYYGMDV	17.868	18.995	16.579
	ASPSIGRDYDFWSGYYPNYYY_GMDV	311	332	290
	ARPSIGRDYDFWSGYYPNYYYYGMDV	94	65	71
	ASPSIGRDYDFWSGY*PNYYYYGMDV	64	56	58
	ASPSIGRDYDFWGGYYPNYYYYGMDV	61	36	39
	AGPSIGRDYDFWSGYYPNYYYYGMDV	50	55	44
IGHV1-69*13	ASPSIGRDYDLWSGYYPNYYYYGMDV	32	46	27
	ASPSIGRDYDFWSGYYPNYYYCGMDV	31	25	23
IGHD3-3*01	ASPSVGRDYDFWSGYYPNYYYYGMDV	28	34	29
	ASPGIGRDYDFWSGYYPNYYYYGMDV	27	33	21
IGHJ6*02	ASPSIGRDYDFWSGYYPNYYYYGMDI	26	10	56
	ASPSIGRDYDFWSGYYPNYYYYDMDV	14	9	57
	ASPSIGRDYDFWSGYYPNYYYYGMNV	10	95	59
	ASPSIGRDYDFWSGYYPNYYYYGIDV	10	19	51
	TSPSIGRDYDFWSGYYPNYYYYGMDV	8	57	7
	ASPSIGRDYNFWSGYYPNYYYYGMDV	5	4	126

Tabelle 26: Patient 18, CLL, Probe 1, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) der klonalen Rearrangement-Population auf Nukleotidebene. V-Region Identity des Clonotypes (AA) 100 %. Der sequenzstärkste Clonotype (AA) in allen drei Verlaufsproben des Patienten ist charakterisiert durch sein VDJ-Rearrangement sowie die Sequenz seiner CDR3 auf Aminosäureebene (vergleiche Tabelle 25, in Blau). Die diesem Clonotype (AA) zugeordneten Sequenzen weisen in ihrer CDR3-Region teilweise Unterschiede auf Nukleotidebene auf, die sich nicht auf die AA-Sequenz auswirken. Abweichungen von der sequenzstärksten Nukleotidsequenz sind in Rot dargestellt.

Aminosäuresequenz der CDR3	Anzahl zugeordneter Sequenzen
ASPSIGRDYDFWSGYYPNYYYYGMDV	17.868
Verschiedene Varianten dieser CDR3 auf Nukleotidebene	Anzahl zugeordneter Sequenzen
gcgagtctagcatcgggcgtgattacgatttttgagtggttattatccgaactactactactacggatggacgtc	16.961
gcgagtctagcatcgggcgtgattatgatttttgagtggttattatccgaactactactactacggatggacgtc	410
gcgagtctagcatcgggcgtgattacgatttttgagtggttattatccaactactactactacggatggacgtc	107
gcgagccctagcatcgggcgtgattacgatttttgagtggttattatccgaactactactactacggatggacgtc	31
gcaagtcctagcatcgggcgtgattacgatttttgagtggttattatccgaactactactactacggatggacgtc	28
gcgagtctagcatcgggcgtgattacgatttttgagtggttattatccgaactactactactacggatggacgtc	28

Ergebnisse

g cg agtc cc gagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	24
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggat g tc	24
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggt ct attatccgaactactactactacggtatggacgtc	23
g cg agtc cc tagcatcggg cg tgattacgatttttggag cg ggttattatccgaactactactactacggtatggacgtc	21
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactacta tt acggtatggacgtc	21
g cg agtc cc tag at cggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	21
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggt ct atccgaactactactactacggtatggacgtc	17
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactacta tg ggtatggacgtc	14
g cg agtc cc tagcatt ggg cg tg attacgatttttggagtggttattatccgaactactactactacggtatggacgtc	14
g cg agtc cc tagcatcggg cg g g attacgatttttggagtggttattatccgaactactactactacggtatggacgtc	13
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacg tt	12
g cg agtc cc tagcatcggg cg tg ct acgatttttggagtggttattatccgaactactactactacggtatggacgtc	10
g cg agtc cc gagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	9
g cg agtc cc tagcatcgg ac g tg attacgatttttggagtggttattatccgaactactactactacggtatggacgtc	9
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttatta cc gaactactactactacggtatggacgtc	8
g cg agtc cc tagcatcggg cg g cg attacgatttttggagtggttattatccgaactactactactacggtatggacgtc	7
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccga at tactactactactacggtatggacgtc	7
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaacta tt tactactactactacggtatggacgtc	6
g cg agtc cc tagcatcggg cg tgattacga ct tttggagtggttattatccgaactactactactacggtatggacgtc	5
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactacta tt tactacggtatggacgtc	5
g ct agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	4
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggt at tattatccgaactactactactacggtatggacgtc	3
g cg agtc cc tagcatcgg ct gattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	3
g cg agtc cc agcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacgtc	2
g cg agtc cc tagcatcggg cg tgattacgatt ct ggagtggttattatccgaactactactactacggtatggacgtc	2
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggt g tattatccgaactactactactacggtatggacgtc	2
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacg g atggacgtc	2
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacg tg	2
g cg ag cc gagcatcggg cg tgattatgatttttggagtggttattatccgaactactactactacggtatggacgtc	1
g cg ag cc tagcatcggg cg g g attacgatttttggagtggttattatccgaactactactactacggtatggacgtc	1
g cg ag cc tagcatcggg cg tgattatgatttttggagtggttattatccgaactactactactacggtatggacgtc	1
g cg agtc cc gagcatcggg cg tgattatgatttttggagtggttattatccgaactactactactacggtatggacgtc	1
g cg agtc cc tagcatcggg cg cgattatgatttttggagtggttattatccgaactactactactacggtatggacgtc	1
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatcc ca actactactactacggtatggacgtc	1
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacg ga tggacgtc	1
g cg agtc cc tagcatcggg cg tgattacgatttttggagtggttattatccgaactactactactacggtatggacg ta	1
g cg agtc cc tagcatcggg cg tgattatgatt ct ggagtggttattatccgaactactactactacggtatggacgtc	1
g cg agtc cc tagcatcggg cg tgattatgatttttggag cg ggttattatccgaactactactactacggtatggacgtc	1
g cg agtc cc tagcatcggg cg tgattatgatttttggagtggttattatccgaactactactactacggtatggacg tt	1
g cg agtc cc tagcatt ggg cg tg attacgatt ct ggagtggttattatccgaactactactactacggtatggacgtc	1

Dem Clonotype (AA) mit der in Tabelle 26 angegebenen AA-Sequenz der CDR3 wurden in dieser Probe 17.868 Sequenzen zugeordnet. Diese Sequenzen umfassten 48 verschiedene Varianten der CDR3 auf Nukleotidebene, die alle für dieselbe Aminosäuresequenz kodierten. Dabei waren jeweils nur eine oder zwei Basen geändert im Vergleich zu derjenigen Nukleotidsequenz der CDR3 mit den meisten zugeordneten Sequenzen. Von den insgesamt 55 Basenaustauschen im Vergleich zur sequenzstärksten Variante entfielen 19 (34,5 %) auf einen Austausch von Cytosin mit Thymin (C > T), 17 (30,9 %) auf einen Austausch von Thymin mit Cytosin (T > C) und 7 (12,5 %) auf einen Austausch von Thymin mit Guanin (T > G). Auf weitere Austauschkonstellationen entfielen jeweils weniger als 6 %.

Patient 19: FL, geringe V-Region Identity

In den Proben von Patient 19 konnte eine IGH-Rearrangement-Population dargestellt werden, die in den Datensätzen der Proben 1 und 2 den Clonotype (AA) auf Rang 1 stellte und im Datensatz FR1 der Probe 3 mit wenigen Sequenzen nachweisbar war (siehe Tabelle 23). Dabei setzte die Population sich aus verschiedenen Clonotypes (AA) zusammen, die sich untereinander vor allem an einzelnen Stellen der Aminosäuresequenz ihrer CDR3 unterschieden (Tabelle 27). Die Clonotypes (AA) der klonalen Population wiesen dabei schwankende Werte in der V-Region Identity auf (bezogen auf die Datensätze FR1): In Probe 1 lag der Wert für alle betrachteten Clonotypes (AA) der Population bei 90 bis 91 %. In Probe 2 wiesen die zwei sequenzstärksten Clonotypes (AA) der Population einen Wert von ca. 99 % auf, während die restlichen Clonotypes (AA) der Population wiederum Werte um die 90 % präsentierten. Die in Probe 3 detektierbaren Clonotypes (AA) der Population lagen ebenfalls bei etwa 90 % V-Region Identity.

Unter den zehn sequenzstärksten Clonotypes (AA) der Datensätze FR1 fanden sich fünf verschiedene Varianten der CDR3 auf Aminosäureebene, von denen sich zwei auch im Datensatz der Probe 3 fanden. Der auf Rang 1 liegende Clonotype (AA) wies dabei in den Datensätzen der Proben 1 und 2 dieselbe

Aminosäuresequenz auf; diese war auch in der dritten Probe nachzuweisen. Die verschiedenen Varianten der CDR3 unterschieden sich jeweils in einer Aminosäure.

Tabelle 27: Patient 19, FL, Primerset FR1. Verschiedene Varianten in der CDR3 des klonalen IGH-Rearrangements und Anzahlen zugeordneter Sequenzen im zeitlichen Verlauf. Alle Varianten der klonalen Rearrangement-Population, die in einem der Datensätze unter den zehn sequenzstärksten Clonotypes (AA) waren, wurden auf Repräsentation in den jeweils anderen Datensätzen überprüft und mit ihrer CDR3 auf AA-Ebene und der Anzahl zugeordneter Sequenzen pro Datensatz aufgeführt. Unterschiede zum sequenzstärksten Clonotype (AA) der Rearrangement-Population (in Blau) auf AA-Ebene sind in Rot dargestellt.

Charakterisierung Clonotype (AA)		Anzahl zugeordneter Sequenzen in Probe...		
V-D-J-Rearrangement	AA-Sequenz der CDR3	1	2	3
	VRAVLNSNGFFSAAFDI	16.941	9.292	4
IGHV1-69*13	VRAVLNSN D FFSAAFDI	258	46	0
IGHD3-22*01	VRAVLNSNGFFSAA FNI	269	5	0
IGHJ3*02	VRAVLN T NGFFSAAFDI	18	5.014	2
	VRAVLN T NGFFS V AFDI	0	108	0

Die Sequenz der CDR3 des sequenzstärksten Clonotypes (AA) in Probe 1 (V-Region Identity ca. 90 %) wies wiederum verschiedene Varianten auf Nukleotidebene auf (Tabelle 28).

Tabelle 28: Patient 19, FL, Probe 1, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) der klonalen Rearrangement-Population auf Nukleotidebene. V-Region Identity des Clonotypes (AA) ca. 90 %. Der sequenzstärkste Clonotype (AA) in allen drei Verlaufsproben des Patienten ist charakterisiert durch sein VDJ-Rearrangement sowie die Sequenz seiner CDR3 auf Aminosäureebene (vergleiche Tabelle 27/Tabelle 25, in Blau). Die diesem Clonotype (AA) zugeordneten Sequenzen weisen in ihrer CDR3 teilweise Unterschiede auf Nukleotidebene auf, die sich nicht auf die AA-Sequenz auswirken. Abweichungen von der sequenzstärksten Nukleotidsequenz sind in Rot dargestellt.

Aminosäuresequenz der CDR	Anzahl zugeordneter Sequenzen
VRAVLNSNGFFSAAFDI	16.562
Verschiedene Varianten dieser CDR3 auf Nukleotidebene	Anzahl zugeordneter Sequenzen
gtcagagcagtcctaaatagtaatggTTTTTTtagtgcggctttgatata	16.115
gtcagagcagtcctaaatagtaatggTTTTTTtagtgcggc ct ttgatata	99
gtcagagcagtcctaaatagtaatggTTTTTTtagtgc ag ctttgatata	62
gtcag gg cagtcctaaatagtaatggTTTTTTtagtgcggctttgatata	37
gtcagagcagtcctaaatagtaatggTTTTTTtag cg cgctttgatata	28
gtcagagcagtc ct aaatagtaatggTTTTTTtagtgcggctttgatata	27
gtcagagcagtcctaaatagtaatgg ct TTTTTTtagtgcggctttgatata	26

Ergebnisse

gtcagagcagtcctaaatagtaa cg gttttttagtg cg gctttgatata	18
gtcagagcagtcctaaatagta atgg ttcttagtg cg gctttgatata	16
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgacata	15
gtcagag cg gtcctaaatagta atgg tttttagtg cg gctttgatata	12
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	11
gt tagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	11
gtcagagcagtcctaaatagta atgg tttttagtg ct gctttgatata	9
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	8
gtcagagcagtcctaaatag ca atggtttttagtg cg gctttgatata	7
gtcagagcag tt taaatagta atgg tttttagtg cg gctttgatata	7
gt aagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	6
gtcagagcagtcctaa ac agta atgg tttttagtg cg gctttgatata	6
gtcagagcagtcctaaatagta atgg tttttagtg cg g ct ttgatata	6
gtcagagcagtc tg aatagta atgg tttttagtg cg gctttgatata	6
gtcagagcagtcctaaatagta atgg tttttagtg cg gctt cg atata	5
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgatat c	5
gtcagagcagtcctaaatagta atgg attttagtg cg gctttgatata	3
gtcagagcag ta taaatagta atgg tttttagtg cg gctttgatata	2
gtcagagcag gt taaatagta atgg tttttagtg cg gctttgatata	2
gtcagag cg gtcctaaatagta atgg tttttagtg cg gctttgatata	2
gtcagag ct gtcctaaatagta atgg tttttagtg cg gctttgatata	2
gtcagagcag ct ctaaatagta atgg tttttagtg cg gctttgatata	1
gtcagagcagtcctaaatag ca atggtttttagtg cg gctttgatata	1
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	1
gtcagagcag tt ctaaatagta atgg tttttagtg cg gctttgatata	1
gtcagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	1
gtcagag gg cagtcctaaatagta atgg tttttagtg cg gctttgatata	1
gt cg gagcagtcctaaatagta atgg tttttagtg cg gctttgatata	1
gt gagagcagtcctaaatagta atgg tttttagtg cg gctttgatata	1

Dem Clonotype (AA) mit der in Tabelle 28 dargestellten Aminosäuresequenz wurden im Datensatz FR1 der Probe 1 16.562 Sequenzen zugeordnet. Hierunter fanden sich wiederum 37 verschiedene Varianten auf Nukleotidebene, die sich an bis zu zwei Stellen von der Nukleotidsequenz mit den meisten zugeordneten Sequenzen unterschieden. Von den insgesamt 40 Basenaustauschen im Vergleich zur sequenzstärksten Variante entfielen 16 (40,0 %) auf einen Austausch von Thymin mit Cytosin (T > C), 4 (10 %) auf einen Austausch von Adenin mit Guanin (A >G) und jeweils 3 (7,5 %) auf einen Austausch von Cytosin

mit Thymin (C > T) bzw. von Guanin mit Adenin (G >A). Auf übrige Austauschkonstellationen entfielen jeweils maximal 5 %.

Zelllinie Granta-519

Auch in den Datensätzen der Verdünnungsproben der Lymphomzelllinie Granta-519 waren jeweils ein prominentes IGH-Rearrangement, das dem der vorangegangenen Sanger-Sequenzierung entsprach, und verschiedene Varianten mit großer Ähnlichkeit vorhanden. Im Folgenden soll ein detaillierterer Blick auf die Unterschiede zwischen einigen sequenzstarken Varianten der Datensätze FR1 und FR2 der Verdünnung 1:5 (S01) erfolgen (Tabelle 29, Tabelle 30).

Im Datensatz FR1 fanden sich sieben verschiedene Varianten des klonalen Rearrangements unter den 20 sequenzstärksten Clonotypes (AA). Sie wiesen alle die gleiche Aminosäuresequenz der CDR3 auf. Für alle dargestellten Clonotypes wurde die IGHV-Genfamilie IGHV4 ermittelt, während die Gene und Allele sich teils unterschieden. Die V-Region Identity lag dabei jeweils bei 99,2 oder 100 %. Alle Varianten wiesen zudem eine Übereinstimmung im D-Gen und –Allel auf, während in zwei Clonotypes (AA) auch die Möglichkeit abweichender D-Gene und –Allele ermittelt wurde. Allen Clonotypes wurde dasselbe J-Gen und –Allel zugeordnet.

Im Datensatz FR2 fanden sich 15 verschiedene Varianten des Lymphom-Rearrangements unter den 20 sequenzstärksten Clonotypes (AA). Sieben hiervon wiesen dieselbe Aminosäuresequenz der CDR3 auf, während die übrigen acht CDR3s mit jeweils einer geänderten Aminosäure bzw. einem Stopcodon aufwiesen. Auch in diesem Datensatz wurde für alle Clonotypes (AA) die IGHV-Genfamilie IGHV4, innerhalb dieser Familie aber zahlreiche mögliche Gene und Allele ermittelt. Die V-Region Identity der dargestellten Clonotypes (AA) lag zwischen 98,9 und 100 %. Elf der 15 Clonotypes (AA) wiesen unter anderem die IGHD-Gen und –Allel-Kombination IGHD5-24*01 auf, die der der Clonotypes (AA) im Datensatz FR1 entspricht. In allen dargestellten Clonotypes (AA) wurde dasselbe J-Gen und –Allel ermittelt.

Ergebnisse

Tabelle 29: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1, Sequenzstarke Clonotypes (AA) der Lymphomzelllinie. Es wurden die 20 sequenzstärksten Clonotypes (AA) betrachtet und diejenigen aufgeführt, die der Lymphomzelllinie zuzuordnen waren. Es wurde für jeden so selektierten Clonotype (AA) der Lymphomzelllinie eine Spalte angelegt. Diese beinhaltet die jeweilige Aminosäuresequenz der CDR3 in der Spaltenbeschriftung sowie alle durch IMGT/HighV-Quest ermittelten möglichen V-, D- und J-Gene und –Allele als Kreuz in der jeweiligen Zeile der Spalte markiert. Die Anzahl dem jeweiligen Clonotype (AA) zugeordneter Sequenzen ist am Spaltenende ebenfalls angegeben. Im Datensatz FR1 wiesen alle Clonotypes (AA) der Lymphompopulation die gleiche Aminosäuresequenz der CDR3 auf. Die V-Region Identity der dargestellten Clonotypes (AA) betrug über 99 %.

Datensatz FR1: Varianten des Lymphom- Rearrangements		Aminosäuresequenz der CDR3						
		ARVGYRSEYFDY						
V-Gen und -Allel	IGHV4-34*01						x	
	IGHV4-34*02						x	
	IGHV4-34*03			x				
	IGHV4-39*07		x					
	IGHV4-59*01	x						
	IGHV4-59*03					x		
	IGHV4-61*05						x	
	IGHV4-61*01							x
D-Gen und -Allel	IGHD2-15*01	x						
	IGHD3-10*01	x						
	IGHD4-17*01	x						
	IGHD5-18*01	x		x				
	IGHD5-24*01	x	x	x	x	x	x	x
J-Gen und - Allel	IGHJ4*02	x	x	x	x	x	x	x
V-Region Identity (%)		100	100	100	100	99,2	100	99,2
Anzahl zugeordneter Sequenzen		25118	250	211	147	83	51	49

Ergebnisse

Tabelle 30: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR2, Sequenzstarke Clonotypes (AA) der Lymphomzelllinie. Es wurden die 20 sequenzstärksten Clonotypes (AA) betrachtet und diejenigen aufgeführt, die der Lymphomzelllinie zuzuordnen waren. Es wurde für jeden so selektierten Clonotype (AA) der Lymphomzelllinie eine Spalte angelegt. Diese beinhaltet die jeweilige Aminosäuresequenz der CDR3 in der Spaltenbeschriftung sowie alle durch IMGT/HighV-Quest ermittelten möglichen V-, D- und J-Gene und –Allele als Kreuz in der jeweiligen Zeile der Spalte markiert. Die Anzahl dem jeweiligen Clonotype (AA) zugeordneter Sequenzen ist am Spaltenende ebenfalls angegeben. Im Datensatz FR2 fanden sich neun verschiedene Aminosäuresequenzen der CDR3 unter den Clonotypes (AA) der Lymphompopulation. Die V-Region Identity betrug mindestens 98,9 %.

Datensatz FR2: Varianten des Lymphom- Rearrangements		Aminosäuresequenz der CDR3																
		ARVGRSEYFDY							ASVGRSEYFDY	ARVGRSEYFDY	ARVGYR*EYFDY	SRVGRSEYFDY	ARVG*RSEYFDY	ARGGRSEYFDY	ERVGRSEYFDY	ARVCYRSEYFDY		
V-Gen und -Allel	IGHV4-4*02								x									
	IGHV4-4*03								x									
	IGHV4-4*04								x									
	IGHV4-4*05								x									
	IGHV4-30-4*03							x										
	IGHV4-34*03			x														
	IGHV4-39*07		x															
	IGHV4-59*01	x								x	x	x	x	x	x	x	x	x
	IGHV4-59*02	x								x	x	x	x	x	x	x	x	x
	IGHV4-59*07	x								x	x	x	x	x	x	x	x	x
	IGHV4-61*01	x								x	x	x	x	x	x	x	x	x
	IGHV4-61*05							x										
	IGHV4-61*06					x												
	IGHV4-61*08	x								x	x	x	x	x	x	x	x	x
D-Gen und -Allel	IGHD1-26*01										x							
	IGHD2-8*01									x					x			x
	IGHD2-15*01	x				x								x				
	IGHD3-10*01	x			x	x	x			x	x			x			x	x
	IGHD3-16*01	x																
	IGHD4-17*01										x							
	IGHD5-12*01	x					x											
	IGHD5-18*01	x			x	x	x			x	x	x	x			x	x	x
	IGHD5-24*01	x	x		x	x	x			x	x		x				x	
	IGHD6-13*01											x						
	IGHD6-19*01														x			
IGHD6-25*01															x			
J-Gen & Allel	IGHJ4*02	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
V-Region Identity (%)		100	99,4	100	99,4	98,9	98,9	99,4	100	99,4	99,4	98,9	99,4	98,9	99,4	99,4	100	
Anzahl zugeordneter Sequenzen		15710	178	168	106	63	62	62	252	146	100	83	82	75	73	65		

Ergebnisse

Die Betrachtung der CDR3 des sequenzstärksten Clonotypes (AA) des Datensatzes FR1 zeigt wiederum verschiedene Varianten der Nukleotidsequenz (Tabelle 31). Der betrachtete Clonotype (AA) wies eine V-Region Identity von 100% auf.

Tabelle 31: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) des Lymphom-Rearrangements auf Nukleotidebene. V-Region-Identity des Clonotypes (AA) 100 %. Der sequenzstärkste Clonotype (AA) der Probe entsprach dem mittels Sanger-Sequenzierung ermittelten Lymphom-Rearrangement. Die Tabelle zeigt unterschiedliche Varianten der Nukleotidsequenz der CDR3 dieses Clonotypes (AA) bei gleicher Aminosäuresequenz, sowie die Anzahl der jeweiligen Variante zugeordneter Sequenzen. Abweichungen von der sequenzstärksten Nukleotidsequenz sind in Rot dargestellt.

Aminosäuresequenz der CDR3	Anzahl zugeordneter Sequenzen
ARVGYRSEYFDY	25.118
Verschiedene Varianten dieser CDR3 auf Nukleotidebene	Anzahl zugeordneter Sequenzen
gcgagagttggctacaggtcggagtactactttgactac	24.950
gcgagagtgggctacaggtcggagtactactttgactac	20
gcgagagtgggctacaggtcggagtactactttgactac	17
gcgagggttggctacaggtcggagtactactttgactac	16
gcgagagttggctacaggtcggagtactacttcgactac	15
gcgagagttggctacaggtcagagtactactttgactac	13
gcaagagttggctacaggtcggagtactactttgactac	12
gctagagttggctacaggtcggagtactactttgactac	11
gcgagagttggatacaggtcggagtactactttgactac	7
gcgagagttggctacagatcggagtactactttgactac	7
gcgagagttggctacaggtcggagtactactttgattac	7
gcgagagttggttacaggtcggagtactactttgactac	6
gcgagagttggctacaggtcggagtactactttgactat	5
gcgagagttggctatagggtcggagtactactttgactac	5
gcgagagttggctacaggtcggagtactactttgactac	5
gcgagagttggctacaggtcggataactactttgactac	4
gcgagagttggctacaggtcggagtactactttgactac	4
gcgagagttggctacaggtctgagtactactttgactac	4
gcgagagttgggtacaggtcggagtactactttgactac	4
gcgagagttggctacaggtcggagtactactttgactac	2
gcgagagttggctacaggtcggagttactttgactac	2
gccagagttggctacaggtcggagtactactttgactac	1
gcgagagtggctacaggtcggagtactactttgactac	1

Dem sequenzstärksten Clonotype (AA) im Datensatz FR1 wurden 25.118 Sequenzen zugeordnet, die 33 verschiedene Nukleotidsequenzen bei gleicher

Aminosäuresequenz der CDR3 aufwiesen. Der sequenzstärksten Nukleotidvariante wurden dabei 24.950 Sequenzen zugeordnet, die übrigen Varianten unterschieden sich in jeweils einem Nukleotid. Von den insgesamt 22 Basenaustauschen im Vergleich zur sequenzstärksten Variante entfielen 6 (27,3 %) auf einen Austausch von Cytosin mit Thymin (C > T), 4 (18,2 %) auf einen Austausch von Guanin mit Adenin (G > A) und jeweils 2 (9,1 %) auf einen Austausch von Thymin mit Cytosin (T > C), von Guanin mit Cytosin (G > C) bzw. von Guanin mit Thymin (G > T). Auf übrige Austauschkonstellationen entfielen jeweils weniger als 5 %.

4.6 NGS-basierte Klonalitätsanalyse von Glaskörperaspiraten mit geringem Lymphomzellgehalt

Um die Anwendbarkeit von NGS-basierter Klonalitätsanalyse an Proben mit geringem Lymphomzellgehalt zu testen, wurden drei Glaskörperaspirate von intraokulären Lymphomen untersucht. Sie wurden anhand ihres Bandenmusters in der GeneScan-Analyse ausgewählt, sodass je eine Probe mit klonalem, mit pseudoklonalem und mit polyklonalem Bandenmuster untersucht wurden.

4.6.1 Anzahl mittels NGS gewonnener und einem IGH-Rearrangement zugeordneter Sequenzen

Zunächst wurden die Anzahlen mittels NGS gewonnener und einem IMGT Clonotype (AA) zugeordneter Sequenzen, sowie die Anzahl verschiedener IMGT Clonotypes (AA) pro Datensatz betrachtet (Tabelle 32).

Die Glaskörperprobe S46 mit monoklonalem Bandenmuster in der GeneScan-Analyse lieferte mit dem Primerset FR1 mittels NGS 7.283 Sequenzen, von denen 6.665 als IGH-Rearrangements eingeordnet wurden, entsprechend 91,5 %. Mit dem Primerset FR2 wurden 7.484 Sequenzen gewonnen, von denen 5.032 als IGH-Rearrangement erkannt wurden, entsprechend 67,2 %. Der

Datensatz FR1 enthielt 73, der Datensatz FR2 198 verschiedene Clonotypes (AA).

Tabelle 32: Anzahl mittels NGS gewonnener und einem IMGT Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT Clonotypes (AA) in Glaskörperproben mit geringem Lymphomzellgehalt.

Probe	GeneScan	Primerset	Anzahl mittels NGS gewonnener Sequenzen	Anzahl einem IMGT Clonotype (AA) zugeordneter Sequenzen	Anteil IGH-Sequenzen an mittels NGS gewonnenen Sequenzen	Anzahl verschiedener IMGT Clonotypes (AA)
S46	klonal	FR1	7.283	6.665	91,5%	73
		FR2	7.484	5.032	67,2%	198
S47	pseudoklonal	FR1	-	-	-	-
		FR2	208	20	9,6%	17
S48	polyklonal	FR1	52	10	19,2%	6
		FR2	188	3	1,6%	3

Die Probe S47, die sich in der GeneScan-Analyse als pseudoklonal dargestellt hatte, hatte mit dem Primerset FR1 keine Bande geliefert und wurde daher nur im Ansatz mit dem Primerset FR2 sequenziert. Hier wurden mittels NGS 208 Sequenzen gewonnen, von denen 20 als IGH-Rearrangement erkannt wurden, entsprechend 9,6 %. Der Datensatz enthielt 17 verschiedene Clonotypes (AA).

Die Glaskörperprobe S48 hatte sich in der GeneScan-Analyse als polyklonal präsentiert. Der Ansatz mit dem Primerset FR1 lieferte 52 Sequenzen, von denen zehn als IGH-Rearrangement erkannt wurden, entsprechend 19,2 %. Der Datensatz enthielt sechs verschiedene Clonotypes (AA). Mit dem Primerset FR2 wurden 188 Sequenzen mittels NGS erzeugt, davon drei, die als IGH-Rearrangement eingeordnet wurden, entsprechend 1,6 %. Der Datensatz enthielt drei verschiedene IMGT Clonotypes (AA).

4.6.2 Beurteilung des nachweisbaren IGH-Repertoires und Vergleich mit Ergebnissen der GeneScan-Analyse

In der Probe S46 zeigte sich in der Zuordnung der Sequenzen nach Clonotype (AA) im Datensatz FR1 eine klare Überrepräsentation des Clonotypes (AA) auf

Rang 1 und seiner Varianten bei kaum vorhandener Hintergrundpopulation (Abbildung 32). Im Datensatz FR2 war ebenfalls ein Großteil der Sequenzen der Ränge 2 bis 10 diesem Rearrangement zuzuordnen. Auf Rang 1 lag ein anderes Rearrangement, das mit einer Sequenzlänge von 361 nt nicht im Erwartungsbereich für eine IGH-Sequenz bei Amplifikation mit diesem Primerset lag. Der mittels NGS gewonnene Befund über die Verteilung der analysierten Sequenzen auf Clonotypes (AA) ist mit dem Monoklonalitätsbefund der GeneScan-Analyse vereinbar.

Probe S46,

GeneScan-Analyse monoklonal

Gesamtzahl einem Clonotype (AA) zugeordneter Sequenzen: FR1 – 6.665, FR2 – 5.032

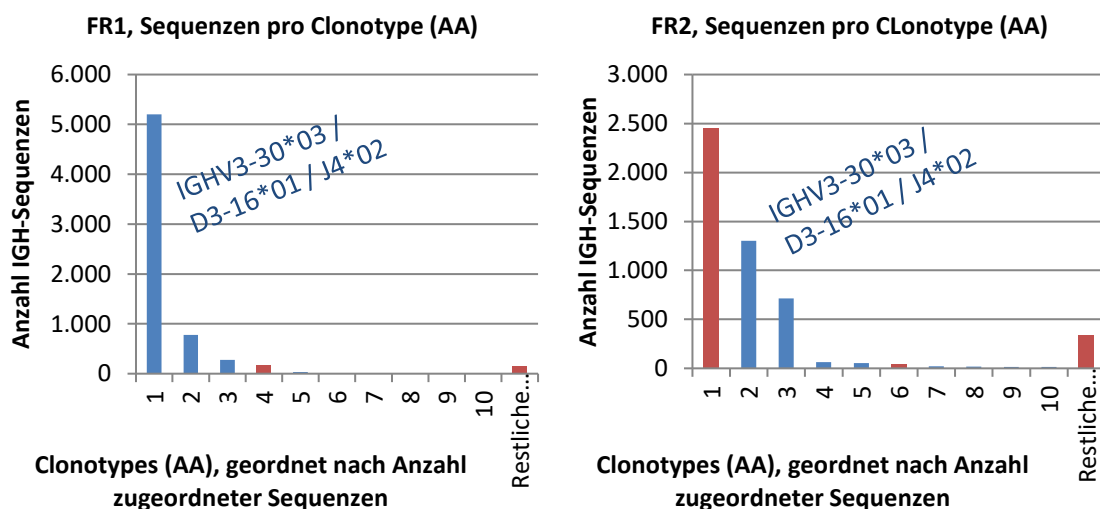


Abbildung 32: Glaskörperprobe S46, Anzahl zugeordneter Sequenzen pro Clonotype (AA). In beiden Datensätzen fanden sich ein IGH-Rearrangement und seine Varianten, denen ein Großteil der IGH-Sequenzen zugeordnet wurde (blau). Im Datensatz FR2 lag auf Rang 1 ein anderes Produkt, das mit einer Sequenzlänge von 361 nt nicht in dem für ein Amplifikat mit diesem Primerset erwarteten Bereich lag.

In der Probe S47 mit pseudoklonalem, und der Probe S48 mit polyklonalem Bandenmuster in der GeneScan-Analyse, zeigte sich mittels NGS und IMGT/HighV-Quest-Analyse das jeweils stark eingeschränkte IGH-Repertoire der Proben: Aufgrund der geringen Anzahlen erzeugter Sequenzen, des jeweils kleinen Anteils an Sequenzen, die als IGH-Rearrangement erkannt wurden und der wenigen verschiedenen IMGT Clonotypes (AA) pro Datensatz, konnten auch mittels NGS keine Rückschlüsse auf das Vorliegen von Klonalität in diesen beiden Glaskörperproben getroffen werden.

5. Diskussion

In verschiedenen Studien der letzten Jahre wurden Ansätze erprobt, die die Klonalitätsdiagnostik in malignen Lymphoproliferationen, sowohl der B- als auch der T-Zellreihe, durch den Einsatz des NGS verbessern sollten. Verschiedene Herangehensweisen wurden gewählt, unter anderem bezüglich der Library-Erstellung, der verwendeten NGS-Plattform, des bioinformatischen Auswertungsansatzes sowie des untersuchten Probenkollektives und -materials.

Diskutiert und verglichen werden sollen zunächst Charakteristika und Kriterien zur Definition klonaler IGH-Rearrangements in NGS-Datensätzen dieser und anderer Arbeiten zu diesem Thema. Da das hier angewandte Verfahren im Amplifikationsschritt modifizierte BIOMED-2 VH- und JH-Primer verwendet, muss im Hinblick auf die verschiedenen Fragestellungen zum Einsatz von NGS stets auch die Performance dieser und alternativer Primersets diskutiert werden. Die Bedeutung der Detektierbarkeit von Varianten klonaler Rearrangements in NGS-Datensätzen hinsichtlich der Beurteilbarkeit des Mutationsstatus und klonaler Evolution bedarf der Erörterung vor dem Hintergrund der problematischen Abgrenzbarkeit zu methodisch bedingten Artefakten. Schließlich wird auf Möglichkeiten NGS-basierter Diagnostik zum Nachweis kleiner Zahlen klonaler Zellen in dieser und anderen Arbeiten eingegangen.

5.1 Anhaltspunkte für Klonalität in NGS-Datensätzen und Vergleich mit der GeneScan-Analyse

5.1.1 Charakteristika klonaler IGH-Rearrangements in NGS-Datensätzen

Die mittels NGS untersuchten Verdünnungsproben von Lymphom-DNA sowie die Patientenproben mit hohem Lymphomzellgehalt sollten zunächst dazu dienen, Parameter der Klonalitätsanalyse in NGS-Datensätzen zu identifizieren und einen Interpretationsansatz der NGS-Datensätze im Hinblick auf Klonalität zu erarbeiten.

Im Rahmen dieser Arbeit, wurde es als Anhaltspunkt für Klonalität in NGS-Datensätzen betrachtet, wenn ein Großteil der als IGH-Rearrangement erkannten Sequenzen demselben IMGT-Clonotype (AA) zugeordnet wurde. Boyd et al. veröffentlichten 2009 einen Ansatz, in dem sie rearrangierte IGH-Loki in Blut- und Gewebeproben mit BIOMED-2 Primern für die FR2 amplifizierten und mittels Pyrosequencing sequenzierten [105]. Die Interpretation der generierten Daten im Hinblick auf Klonalität erfolgte, indem auf einem zweidimensionalen Gitter die in der Probe detektierten VH-Gensegmente gegen die JH-Gensegmente aufgetragen wurden. Ein farbiger Punkt an der Stelle des Gitternetzes zeigte das verwendete VH- und JH-Gensegment eines VDJ-Rearrangements in der Probe an. Die Größe und farbliche Wärme des Punktes zeigte wiederum an, welcher Anteil aller Sequenzen in der Probe die Angezeigten VH- und JH-Gensegmente verwendete. Die so untersuchten CLL-, FL- und PTLID-Fälle („posttransplant lymphoproliferative disease“), sowie Blutproben gesunder Spender entsprachen in ihren NGS-Befunden denen der zuvor durchgeführten GeneScan-Analyse. Im Gegensatz zum Vorgehen in dieser Arbeit, wurden Sequenzen, in denen V-, D- oder J-Segmente oder Verbindungsregionen nicht eindeutig zugeordnet werden konnten, vor der Erstellung der Plots herausgefiltert. Auch wurden Rearrangements nur nach der Kombination ihrer VH- und JH-Gensegmente unterschieden, wohingegen in dieser Arbeit eine Differenzierung der detektierten Rearrangements auf Gen- und Allelebene erfolgte und sowohl D-Gensegment als auch die CDR3-Region jeder IGH-Sequenz analysiert wurden.

Schumacher et al. untersuchten 2014 an Lymphoproliferationen der T-Zell-Reihe eine Möglichkeit der T-Zell-Rezeptor- γ -basierten Klonalitätstestung mittels NGS und verglichen ihre Ergebnisse mit denen der GeneScan-Analyse [106]. Ihre Arbeit wies Ähnlichkeiten mit der vorliegenden v.a. in der Analyse und Interpretation der NGS-Datensätze auf. Das PCR-amplifizierte und Barcode-markierte TCRG-Repertoire jeder Probe wurde sequenziert und die gewonnenen Sequenzen durch den Sortieralgorithmus „TCRDriver“ analysiert. Es erfolgte eine Einteilung in Cluster basierend auf Übereinstimmungen von V- und J-Segment sowie Sequenz der Verbindungsregion. Die Cluster wurden nach der Anzahl

zugeordneter Sequenzen sortiert. Ebenfalls übereinstimmend mit der vorliegenden Arbeit wurden derart generierte und analysierte NGS-Daten ausgewertet, indem wie in der konventionellen GeneScan-Analyse prominente Cluster mit der polyklonalen Hintergrundpopulation verglichen wurden: Als Kriterium für Klonalität wurde etabliert, dass das sequenzstärkste Cluster bzw. die zwei sequenzstärksten Cluster (im Falle eines biallelischen Rearrangements) mindestens vierfach mehr Sequenzen enthalten mussten, als Cluster der Hintergrundpopulation. Basierend auf dieser Herangehensweise der Auswertung wiesen 19 von 19 untersuchten polyklonalen, sowie 24 von 27 untersuchten klonalen Fällen übereinstimmende Ergebnisse in Kapillarelektrophorese- und NGS-basierter Klonalitätsuntersuchung auf.

Das Kriterium, dass ein klonales Cluster mindestens vierfach sequenzstärker als Cluster der Hintergrundpopulation sein muss, wurde auch von Amato et al. 2016 in einer Studie angewandt, in der der Klonalitätsnachweis mittels NGS an endemischen Burkitt-Lymphomen und damit an Neoplasien der B-Zell-Reihe geführt wurde. Auch hier präsentierten sich alle 37 untersuchten Fälle anhand dieses Kriteriums klonal [104].

Im Rahmen dieser Studie wurde kein quantitatives Kriterium für Klonalität in NGS-Datensätzen etabliert. Alle hier qualitativ als klonal gewerteten Fälle sind allerdings, unter dem Gesichtspunkt betrachtet, dass ein klonales Cluster mindestens vierfach sequenzstärker als der polyklonale Hintergrund sein muss, ebenfalls als klonal einzustufen und die Ergebnisse damit mit jenen der Literatur vereinbar. Die Datensätze der Verdünnungsproben der Zelllinie Granta-519 sind nach diesem Kriterium ebenfalls klonal, mit Ausnahme der höchsten Verdünnung von 1:1.000. Der Datensatz der Verdünnung 1:100 mit dem Primerset FR1 der Zelllinie Granta-519 war nicht in den durch das Institut für Medizinische Genetik gelieferten Daten enthalten. Hier steht in Frage, ob fälschlicherweise derselbe Barcode-Primer wie für die Verdünnung 1:50 verwendet wurde, was ebenfalls die große Anzahl in diesem Ansatz gewonnener Sequenzen erklären würde.

5.1.2 Charakteristika der amplifizierten Hintergrundpopulation

Bei der Etablierung des hier angewandten NGS-Verfahrens zum Klonalitätsnachweis war das Verhalten von Rearrangements der polyklonalen Hintergrundpopulation relevant. Wie in diesem Kapitel diskutiert, fand die Identifikation klonaler Rearrangements in dieser und in anderen Arbeiten in quantitativer Abhängigkeit von Rearrangements der mutmaßlichen Hintergrundpopulation statt. In Verdünnungsproben der Zelllinien Granta-519 in PBMC-DNA wurden in dieser Arbeit einzelnen Clonotypes (AA) der Hintergrundpopulation $< 0,01$ % bis maximal 1,38 % aller Sequenzen zugeordnet. Dies steht in Einklang mit anderen Arbeiten, in denen Maximalanteile von 1,5 bis 2,0 % einzelner Clonotypes normaler B-Zellpopulationen ermittelt wurden [33, 104, 127]. Mehrfach Anwendung findet außerdem ein höherer Cut-Off-Wert von 5 % zugeordneter Sequenzen zur Abgrenzung normaler von klonal expandierten Clonotypes [109, 110, 128]. Das häufigste verwendete V-Gen der Hintergrundpopulation war in exemplarisch dargestellten Datensätzen der Zelllinie OCI-LY3 IGHV3-23, vereinbar mit bestehenden Untersuchungen [33, 36, 104].

Für das Verhalten von polyklonalen Populationen aus FFPE-Material in NGS-Datensätzen sind bisher nur spärlich Daten veröffentlicht. In Datensätzen der Patientenproben mit hohem Lymphomzellgehalt und mit gelungener Sanger-Sequenzierung, waren sequenzstärkere Clonotypes (AA) der Hintergrundpopulation mit bis zu 3,5 % zugeordneten Sequenzen vertreten. In Patientenproben ohne gelungenen Nachweis einer Lymphompopulation entfielen bis zu 38,9 % (S17, FR1) der Sequenzen auf Clonotypes (AA) der mutmaßlichen Hintergrundpopulation. Es ist also davon auszugehen, dass Faktoren wie eingeschränkte DNA-Qualität und –Menge aus B-Lymphozyten auch in NGS-Datensätzen zu pseudo- und oligoklonalen Verteilungsmustern der Clonotypes (AA) führen können. Besonders die Verlaufsproben der Patienten 16 und 20 machen dabei auf die Möglichkeit falsch positiver Befunde mittels NGS aufmerksam: In fünf der zusammengekommen elf Datensätze dieser Patienten waren Clonotypes (AA) auf Rang 1, die mutmaßlich nicht der zu verfolgenden

Lymphompopulation angehörten. Hiervon wurde ausgegangen, da diese Populationen in jeweils nur einem Datensatz und nicht im weiteren zeitlichen Verlauf nachweisbar waren, im Gegensatz zu den jeweils als Lymphompopulation interpretierten Populationen. Dennoch wären in drei der diskutierten Datensätze (Patient 16, Zeitpunkt 1 – FR1, Zeitpunkt 2 – FR1 sowie Patient 20, Zeitpunkt 2 – FR2), für sich betrachtet und nicht in Zusammenschau mit den weiteren Datensätzen der jeweiligen Verlaufsproben, der Clonotype (AA) auf Rang 1 anhand der an den Zelllinienverdünnungsproben erarbeiteten und in der Literatur etablierten Kriterien als klonales Rearrangement einzustufen. Dass keine der detektierten pseudoklonalen Populationen sich mit dem jeweils anderen Primerset oder in einer weiteren Verlaufsprobe des jeweiligen Patienten reproduzieren ließ, spricht dabei gegen reaktiv expandierte Klone mit hoher Repräsentation in der jeweiligen Probe. Die kritische Beurteilung der Möglichkeit von PCR-Bias, also der Überamplifikation von IGH-Rearrangements der Hintergrundpopulation, in der IGH- oder Barcode-PCR [129-132], beispielsweise bei ungenügender DNA-Qualität oder –Menge, und folglich falscher Interpretation als klonal, bedürfte der genaueren Untersuchung reaktiver Lymphozytenproben aus FFPE-Material mittels NGS und einer Überprüfung der Reproduzierbarkeit prominenter Rearrangements in diagnostischen Proben. So ist auch bei der GeneScan-basierten Klonalitätstestung etabliert, zur Vermeidung falsch positiver Befunde, jede Probe in zwei separaten Ansätzen zu analysieren [36].

5.1.3 Vergleich der Klonalitätstestungen mittels NGS und GeneScan

Die in dieser Studie untersuchten Proben mit hohem Lymphomzellgehalt wurden zusätzlich einer GeneScan-Analyse zur Klonalitätsuntersuchung unterzogen und die Ergebnisse mit denen des NGS-Verfahrens verglichen. In sieben der 15 Fälle stimmten die Befunde der beiden Verfahren miteinander überein, sodass keine hohe Konkordanz zwischen den beiden Verfahren bestand. Im Gegensatz hierzu wiesen alle von Boyd et al. 2009 mittels NGS als klonal interpretierten Befunde auch in der GeneScan-Analyse eine klonales Bandenmuster auf, während die Proben gesunder Probanden sich jeweils in beiden Verfahren polyklonal

präsentierten [105]. Auch Schumacher et al. erzielten in 19 von 19 mittels GeneScan polyklonalen und 24 von 27 klonalen Fällen übereinstimmende Ergebnisse mittels NGS, während zwei der diskordanten Fälle ein Problem der kapillarelektrophoretischen Klonalitätsanalyse veranschaulichten [106]: So zeigten die NGS-Datensätze jeweils zwei, respektive drei unterschiedliche Rearrangements, die zwar für sich betrachtet anhand ihrer Sequenzstärke nicht die Klonalitätskriterien mittels NGS erfüllten, die jedoch zufälligerweise dieselbe Nukleotidlänge aufwiesen, sodass sie im Kapillarelektrophorese-basierten GeneScan-Verfahren in einem falsch positiven „klonalen“ Peak resultierten. Eine höhere Sensitivität ihres NGS-Verfahrens stellten die Autoren nicht fest. Eine Studie von Sufficool et al. 2015 jedoch konnte an einem Kollektiv von Mycosis-fungoides-Fällen eine höhere Sensitivität NGS-basierter TCR-Klonalitätstestung gegenüber solcher mittels Kapillarelektrophorese zeigen [107].

In einer multizentrischen technischen Machbarkeitsstudie der EuroClonality-NGS-Arbeitsgruppe verglichen Scheijnen et al. 2019 die etablierten EuroClonality/BIOMED-2 Assays mit einem durch die Arbeitsgruppe entwickelten NGS-Verfahren zur Klonalitätsdiagnostik anhand von 14 FFPE-Proben von B-NHL sowie Tonsillen gesunder Probanden [112]. Die hier verwendeten NGS-Primer basierten auf den BIOMED-2-Primern für IGH und IGK. Das NGS-Verfahren zeigte eine zuverlässige und in einigen Fällen gar verbesserte Detektion klonaler Rearrangements im untersuchten Kollektiv.

Des Weiteren bietet die NGS-basierte Klonalitätstestung Vorteile gegenüber der GeneScan-Analyse was die Befundinterpretation angeht. Die Interpretation von GeneScan-Befunden erfordert in vielen Fällen große Erfahrung seitens des Untersuchers [92]. Mittels NGS gewonnene Daten sind quantitativ und damit potenziell leichter interpretierbar, auch können sie zu einer Verringerung der interpretativen Subjektivität beitragen, wie sie der GeneScan-Analyse inhärent ist. Ein weiterer entscheidender Vorteil der NGS-basierten Klonalitätstestung ist die Möglichkeit, die gewonnene Charakterisierung eines klonalen Rearrangements im Verlauf der Lymphomkrankung zum Monitoring von Minimal Residual Disease zu nutzen (siehe 5.3).

Die Überlegenheit des NGS-basierten Verfahrens in Bezug auf bloßen Informationsgehalt wird u.a. dadurch deutlich, dass sich aus diagnostischen NGS-Daten ein „Elektropherogramm“ rekonstruieren lässt, wie an mehreren Fällen dieser Studie erfolgt. So liefert NGS potenziell alle Informationen, die auch das GeneScan-Verfahren liefert, und darüber hinaus etliche mehr, wie beispielsweise eine detaillierte Charakterisierung der polyklonalen bzw. reaktiven Hintergrundpopulation. Diese Informationen spielen zwar für die primäre Diagnose einer malignen Lymphoproliferation keine entscheidende Rolle, eröffnen aber die Möglichkeit, zu diagnostischen Zwecken routinemäßig erhobene Daten für weitergehende wissenschaftliche Fragestellungen zu nutzen.

Die Parameter „Anzahl mittels NGS gewonnener Sequenzen“ und „Anzahl einem IMGT-Clonotype (AA) zugeordneter Sequenzen“, die sich aus Analyse der NGS-Datensätze mittels IMGT/HighV-Quest ergaben, liefern außerdem Rückschlüsse auf die Effektivität der IGH-Amplifikation sowie auf die Vielfalt des amplifizierten Repertoires, was beispielsweise bei der Beurteilung von Proben mit geringem Zellgehalt oder eingeschränkter DNA-Qualität von Nutzen sein kann, wie unten diskutiert.

5.1.4 Gründe für falsch negative GeneScan- und NGS-Befunde

Die im Rahmen dieser Arbeit molekularpathologisch untersuchten Patientenproben mit hohem Lymphomzellgehalt waren zuvor histologisch und immunhistochemisch eindeutig als Lymphom klassifiziert worden, sodass GeneScan- und NGS-Ergebnisse, die keinen Klonalitätsnachweis erbrachten, als falsch negativ erachtet werden müssen. Die hier verwendeten standardisierten Primersets der GeneScan-Analyse, auf denen auch die Primer für das NGS-Verfahren basierten, haben, seit ihrer Entwicklung durch das BIOMED-2 Konsortium, breite Anwendung gefunden und ausführliche Validierung erfahren. So zeigte sich als häufige Ursache falsch negativer Klonalitätsbefunde das Vorliegen von Mutationen im Bereich der Primer-Bindungsstellen. Dieses Problem betrifft v.a. Entitäten mit einer hohen Anzahl somatischer

Hypermutationen, insbesondere FL, DLBCL und multiple Myelome (MM) [36, 133]. Auch die Verwendung von FFPE-Material ist in diesem Zusammenhang problematisch, da diese Form der Fixierung zu Sequenzalterationen und Fragmentierung der DNA führt [36, 85, 134]; die Assays des BIOMED-2 Konsortiums wurden hauptsächlich an gefrorenem Gewebe und Frischmaterial validiert [93].

In mehreren Studien wurden mit den BIOMED-2 IGH-Assays zum Klonalitätsnachweis in FL bei Verwendung der drei Primersets FR1, 2 und 3 stark divergierende Sensitivitätswerte von 37 % bis 68 % an FFPE-Gewebe ermittelt [88-91, 133, 135], in einer Studie von knapp 80 % [86]. An frischem Gewebe hingegen konnte eine Sensitivität von 96 % an FL bei Verwendung der Primersets FR1, 2 und 3 erzielt werden [136]. In vielen Studien stellten sich die BIOMED-2 IGK-Assays zum Klonalitätsnachweis bei FL als sensitiver heraus, wobei die maximale Sensitivität durch Kombination von IGH- und IGK-Assays erreicht wurde und bis zu 100 % betrug [86, 88, 90, 91, 93, 135, 136].

Ähnliche Probleme sind für die GeneScan-Analyse von DLBCL beschrieben. Eine Studie zum Klonalitätsnachweis mittels BIOMED-2 IGH-Assays untersuchte 78 FFPE-DLBCL-Fälle und ermittelten eine Sensitivität des Primersets FR3 von knapp 58 % [87]. Wie im Falle des FL konnte an frischem DLBCL-Gewebe eine höhere Sensitivität erreicht werden: 74 % bis 79 % bei Verwendung der Primersets FR1, 2 und 3 in Kombination und bis zu 100 % bei Verwendung aller BIOMED-2 Primersets [93, 136]. In den NGS-Datensätzen der DLBCL-Zelllinie OCI-LY3, konnte bei Verwendung der Primersets FR1 und FR2 dennoch das zuvor mittels Sanger-Sequenzierung ermittelte mutmaßliche Lymphom-Rearrangement nicht detektiert werden. Dies könnte mit der großen Länge der mittels Sanger-Sequenzierung ermittelten Rearrangement-Sequenz zusammenhängen (451 nt mit dem Primerset FR1): Die Amplikons des Lymphom-Rearrangements wären zu lang für eine durchgängige erfolgreiche Sequenzierung auf der Illumina MiSeq-Plattform.

Es ist zu vermuten, dass die Verwendung von FFPE-Gewebe und die Limitation auf nur eines der drei IGH-Primersets (Primerset FR2) für die GeneScan-Analyse

bzw. zwei Primersets (FR1 und 2) für die NGS-basierte Klonalitätsanalyse zu falsch negativen Befunden in den hier untersuchten FL- und DLBCL-Kollektiven beitragen. So waren von fünf FL-Fällen mit hohem Lymphomzellgehalt zwei übereinstimmend polyklonal, einer übereinstimmend klonal und zwei Fälle mit jeweils einem Verfahren klonal und im anderen polyklonal. Von den fünf untersuchten DLBCL-Fällen mit hohem Lymphomzellgehalt war ein Fall übereinstimmend klonal, während vier Fälle pseudoklonal mittels GeneScan und in mindestens einem NGS-Datensatz klonal waren. Hervorzuheben ist die Tatsache, dass drei der fünf mittels NGS klonalen DLBCL-Fälle in nur einem Datensatz klonal waren: Dies hebt erneut die Bedeutung der Verwendung mehrerer Primersets zum Klonalitätsnachweis hervor.

Für die CLL sind Sensitivitäten der drei BIOMED-2 IGH-Primersets in Kombination von 96 % an frischem und 80 % an FFPE-Gewebe beschrieben [90]. Trotz der Beschränkung auf ein Primersets in der GeneScan-Analyse in dieser Studie stellten sich vier der fünf CLL-Fälle mit hohem Lymphomzellgehalt als klonal dar. Auch mittels NGS waren vier von fünf Fällen klonal. Allerdings kamen die beiden Verfahren in zwei Fällen zu unterschiedlichen Ergebnissen: Ein mittels NGS klonaler Fall war mittels GeneScan pseudoklonal, während ein mittels GeneScan klonaler Fall in beiden NGS-Datensätzen ein polyklonales Muster zeigte.

Alle im Rahmen dieser Arbeit untersuchten Proben betrachtet, konnte in mehr Proben mit dem Primersets FR1 keine Amplifikation erzielt werden als mit dem Primersets FR2. Auch konnten mit dem Primersets FR2 in mehr Fällen starke Banden erzielt werden. Dies könnte mit der größeren Amplikonlänge mit dem Primersets FR1 (310 – 360 bp) im Vergleich zum FR2-Assay (250 – 295 bp) bei gewissem Fragmentierungsgrad der DNA zusammenhängen [89, 133]. Um diesem Problem zu begegnen, wurden auch im Rahmen der beschriebenen EuroClonality-NGS Studie von 2019 ausschließlich Primer für die FR3 eingesetzt, um mit dem Ziel höherer Sensitivität kleinere Amplikons zu generieren [112].

Beim Vergleich der NGS- und GeneScan-Befunde ist das möglicherweise unterschiedliche Annealing-Verhalten der jeweils verwendeten Multiplex-PCR-Primer zu bedenken: Während für die GeneScan-Analyse die BIOMED-2 IGH-Primer verwendet wurden, kamen für die NGS-Untersuchung modifizierte Primer mit Fluidigm-Anhang zum Einsatz. Diese besitzen zwar dieselbe 22 – 25 nt lange, zur Bindung an verschiedene VH-Gensegmente optimierte Sequenz, sind allerdings zusätzlich um einen 22 nt langen Anhang zur Bindung der Fluidigm-Barcode-Primer im zweiten PCR-Schritt erweitert. Da dieser Anhang nicht an das zu amplifizierende Rearrangement bindet, könnte er das Annealing-Verhalten negativ beeinflussen und zu einer schlechteren Amplifizierbarkeit der NHL-Proben mit Fluidigm-Primern als mit den Primersets zur GeneScan-Analyse führen.

Die Zusammenschau der NGS- und GeneScan-Befunde macht deutlich, dass viele Schwachstellen des hier erprobten NGS-basierten Verfahrens wohl primär auf die bekannten Probleme der BIOMED-2 Primersets, auf deren Design auch die NGS-Fluidigm-Primer basieren, zurückzuführen sind.

5.2 Mutationsstatus und Varianten klonaler IGH-Rearrangements und Abgrenzbarkeit zu Artefakten

5.2.1 Varianten klonaler Rearrangements als Problem in der Klonalitätsdiagnostik

Die Untersuchung der Verdünnungsproben der Zelllinie Granta-519 sowie einer Vielzahl der untersuchten Patientenproben zeigte, dass mittels NGS und Analyse durch IMGT/HighV-Quest neben dem der gegebenenfalls vorhandenen Sanger-Sequenzierung entsprechenden Clonotype (AA) auf Rang 1, aufgrund kleinerer Sequenzunterschiede zahlreiche weitere IMGT-Clonotypes (AA) innerhalb der Lymphompopulation definiert wurden. Im Datensatz FR1 der Zelllinienverdünnungsprobe 1:50 beispielsweise stellte der der Sanger-Sequenzierung entsprechende Clonotype (AA) auf Rang 1 28,9 % aller IGH-Sequenzen, während Varianten dieses Clonotypes unter den top zehn

Clonotypes insgesamt 5 % aller Sequenzen zugeordnet wurden. Weitere Varianten fanden sich darüber hinaus unter den Clonotypes jenseits Rang 10, sodass in diesem Datensatz mehr als ein Siebtel der Sequenzen der Lymphompopulation nicht dem Clonotype der Sanger-Sequenz zugeordnet wurden. Innerhalb der analysierten Population zeigte sich also eine gewisse Heterogenität der IGH-Rearrangements auf Nukleotid- und teilweise auf Aminosäureebene, die anhand der Datensätze FR1 und FR2 der Verdünnung 1:5 näher veranschaulicht wurde.

Die Vielzahl an detektierten und charakterisierten Subklonen innerhalb der durch die Zelllinien-DNA simulierten Lymphompopulation stellt ein Problem in der hier angewandten Auswertung der durch IMGT/HighV-Quest generierten Datensätze im Hinblick auf Klonalität dar. Wenn IGH-Sequenzen aufgrund geringerer Sequenzunterschiede zahlreichen verschiedenen, wenn auch ähnlichen, Clonotypes (AA) zugeordnet werden, sind diese für sich betrachtet jeweils eventuell zu sequenzschwach, um unter die top zehn Clonotypes zu gelangen. Dieser Effekt kann dazu führen, dass der Anteil der Lymphompopulation an der Gesamtzahl der IGH-Sequenzen in einem Datensatz unterschätzt wird. Dies ist ebenfalls der Grund, weshalb der hier erarbeitete Ansatz zur Auswertung der IMGT/HighV-Quest-Datensätze es lediglich erlaubt, die Mindestanzahl der einer Lymphompopulation zugehörigen Sequenzen anzugeben. Zwar lieferten die hier untersuchten Verdünnungs- und Patientenproben kein Beispiel für folgendes Szenario, aber es wäre vorstellbar, dass in einem Datensatz ein relativ sequenzstarkes Rearrangement der Hintergrundpopulation auf Rang 1 liegt, während eine gleichzeitig präsente Lymphompopulation bei hoher durch IMGT/HighV-Quest detektierter interklonaler Heterogenität sich auf verschiedene sequenzschwächere Clonotypes (AA) verteilt und ein falsch negativer Befund resultiert. Eine Möglichkeit diesem Problem zu begegnen wäre, die Betrachtung von den top Zehn auszuweiten auf eine größere Anzahl Clonotypes (AA) und die Sequenzzahlen der Clonotypes der Lymphompopulation zu summieren. Angesichts der großen Anzahl verschiedener in einem Datensatz enthaltener IMGT-Clonotypes (AA) von teilweise über 3.000 ist dies manuell allerdings kaum

praktikabel und würde die Erarbeitung eines entsprechenden Algorithmus voraussetzen.

Ein weiteres Problem des hier verwendeten Auswertungsansatzes, das auf der hohen Trennschärfe der IMGT-Clonotypes (AA) beruht, ist die Entscheidungsfindung darüber, ob ein bestimmter Clonotype (AA) einer Lymphompopulation zuzuordnen ist, oder ob es sich um eine eigenständige Population des polyklonalen Hintergrundes handelt, die nur zufällig große Ähnlichkeiten mit einem in diesem Datensatz prominenten Lymphom-Rearrangement aufweist. In dieser Arbeit wurden Clonotypes (AA) anhand ihres VDJ-Rearrangements auf Gen- und Allelebene, anhand ihrer Nukleotidlänge sowie anhand der AA-Länge und –Sequenz ihrer CDR3 beurteilt. Bei gleichem VDJ-Rearrangement zweier Clonotypes auf Gen- und Allelebene, gleicher AA-Länge und bis auf eine Aminosäure identischer Sequenz der CDR3 liegt der Schluss nahe, dass es sich um verschiedene Varianten derselben Lymphom- bzw. Lymphozytenpopulation handelt, die durch somatische Hypermutation entstanden sein könnten. Auch verschiedene Allele bei gleichem V-Gen zweier Rearrangements bei ansonsten übereinstimmenden Parametern legen diesen Schluss nahe. Insbesondere die Tatsache, dass Sequenzen unter Umständen gleiche Übereinstimmungsgrade mit mehreren Allelen der IMGT-Referenzdatenbank aufweisen und daher mehrere in Frage kommende Allele dem Rearrangement zugeordnet werden können veranschaulicht, dass kleine Unterschiede auf Nukleotidebene ausreichen, um zu unterschiedlichen zugeordneten Allelen zweier Rearrangements und daher unterschiedlichen Clonotypes (AA) zu führen, obwohl beide der gleichen Lymphompopulation entstammen. Bei einer Vielzahl dem prädominanten Rearrangement ähnlicher Clonotypes (AA) ist der Übergang zu solchen Clonotypes (AA) jedoch fließend, die eher der polyklonalen Hintergrundpopulation zuzuordnen sind. In dieser Studie wurden Clonotypes (AA) nur bei starker Ähnlichkeit als einer klonalen Population zugehörig interpretiert. Doch die fehlende Quantifizierbarkeit der dieser Einordnung zugrundeliegenden Kriterien führt zu einer gewissen Subjektivität in der Beurteilung der untersuchten NGS-Datensätze. Diese war zwar, aufgrund der in den als klonal interpretierten Datensätzen stets auf Rang 1

liegenden prominenten Clonotypes, in keinem untersuchten Datensatz befundentscheidend, steht aber dem bereits diskutierten Vorteil NGS-basierter gegenüber GeneScan-basierten Methoden zur Klonalitätstestung entgegen, dass erstere eine geringere interpretative Subjektivität aufweisen sollte und weniger Erfahrung bedürfe als Letztere.

In vorausgegangenen Arbeiten zum Einsatz des NGS zu Zwecken der Klonalitäts- und MRD-Diagnostik wurde dem Problem kleinerer Sequenzalterationen innerhalb einer mutmaßlichen Lymphompopulation auf verschiedene Weise begegnet. Wie beschrieben, unterschieden Boyd et al. 2009 IGH-Rearrangements allein anhand der Kombination verwendeter V- und J-Gensegmente [105], was zu einem unscharfen Verteilungsmuster der IGH-Sequenzen auf Übergruppen mehrerer Populationen mit gleicher VJ-Segmentskombination führte. Dies lieferte zwar in Fällen mit prominenter klonaler Population ein eindeutig klonales Bild, erlaubt aber keine genaue Charakterisierung und Quantifizierung klonaler Rearrangements. Schumacher et al. untersuchten 2014 TCRG-Rearrangements und bildeten Cluster von Sequenzen mit gleichem V- und J-Gensegment, Übereinstimmung der Sequenzen der Verbindungsregion und einem Identity Score von mindestens 95 % bei Alignment [106]. Der durch Scheijen et al. 2019 im Rahmen der EuroClonality-NGS Studie verwandte bioinformatische Ansatz basierte auf dem Algorithmus „ARResT/Interrogate“, welcher, auf Grundlage von IMGT/HighV-Quest Analysen, aus Clonotypes abgeleitete sogenannte „feature types“ definiert. Je nach Fragestellung kann hier die Überrepräsentation eines „feature types“ in einer Probe quantifiziert und eine Aussage über Klonalität getroffen werden [112, 137].

In einigen Studien korrigierten die Autoren einzelne Nukleotidinsertionen oder – deletionen ansonsten gleicher Sequenzen, die sie auf den Sequenzierprozess zurückführten, durch Alignment mit prominenteren Clonotypes [104, 109, 138]. Auch die vorliegende Studie enthält Beispiele, in denen Clonotypes (AA) um wenige Nukleotidlängen voneinander abweichen, obwohl sie derselben Lymphom- oder Lymphozytenpopulation zu entstammen scheinen; so im Datensatz FR1 der Verdünnungsprobe 1:5 der Zelllinie Granta-519.

Auch quantitative Kriterien für die Zuordnung eines Clonotypes zu einer Lymphompopulation sind beschrieben, wie von Faham et al. 2012, die jeden zum Diagnosezeitpunkt charakterisierten Clonotype, dem mehr als 5 % aller Sequenzen in einem Datensatz zugeordnet wurden, als der Lymphompopulation zugehörig interpretierten, basierend auf der Annahme, dass Clonotypes der normalen B-Zellpopulation sich unterhalb dieses Schwellenwertes bewegten [139]. Um dieses Kriterium auf die Datensätze in dieser Studie anwenden zu können, wäre zuvor allerdings die Erprobung des hier praktizierten Vorgehens an eindeutig als reaktiv eingestuften Lymphknotenproben nötig, um den Maximalwert von 5 % zugeordneter Sequenzen reaktiver Lymphozytenpopulationen zu bestätigen. Teilweise wurde für die Definition eines Clonotypes vorausgesetzt, dass mindestens zwei Reads desselben vorhanden waren [128, 138, 139]. Da eine Vielzahl der in dieser Arbeit durch IMGT-HighV-Quest definierten Clonotypes (AA), v.a. der polyklonalen Hintergrundpopulation, mit nur einer zugeordneten Sequenz vertreten sind, würde die Anwendung dieses Kriteriums den Anteil der Hintergrundpopulation am Gesamt-IGH-Repertoire schmälern und könnte zu falsch positiven NGS-Befunden führen.

5.2.2 Bedeutung von Mutationen und Varianten in mittels NGS detektierten Lymphompopulationen

Neben der Beantwortung der Frage nach der Zugehörigkeit eines IMGT-Clonotypes (AA) zu einer Lymphozyten- oder Lymphompopulation, stellt eine Herausforderung auch die Beantwortung der Frage dar, ob kleinere Unterschiede zwischen IMGT-Clonotypes (AA) biologisch bedingt, also echte, durch SHM oder andere Prozesse der BCR-Generierung entstandene Varianten, oder methodisch bedingt sind, also Artefakte, die dem Einfluss der Verwendung von FFPE-Gewebe geschuldet sowie durch Fehler der PCR und des Sequenzierprozesses entstanden sind. Klare Aussagen bezüglich dieser Fragestellungen sind insbesondere wichtig bei der Beurteilung des Mutationsstatus von CLL sowie beim Versuch, klonale Evolution in Verlaufspalten durch anhaltende SHM zu detektieren.

Mutationsstatus und Varianten der Zelllinie Granta-519

Bezüglich der Frage, wie in dieser Arbeit detektierte Varianten im Sinne von echten, biologisch bedingten Subklonen einer Lymphompopulation, oder im Sinne von fixierungs-, amplifikations- oder sequenzierungsbedingten Artefakten einzuordnen sind, soll zunächst die Variabilität in den Datensätzen der Zelllinie Granta-519 diskutiert werden. In ihnen wurden in dieser Arbeit eine Vielzahl von der Sanger-Sequenz abweichender Varianten festgestellt, die Unterschiede in der CDR3, aber auch bezüglich der ihnen durch IMGT/HighV-Quest zugeordneten V- oder D-Gensegmente aufwiesen, wobei die jeweilige Segmentfamilie stets variantenübergreifend übereinstimmend war. Die Zelllinie Granta-519 ist eine B-Zellneoplasie aus dem peripheren Blut einer 58-jährigen kaukasischen Frau [140] und als Mantelzelllymphom klassifiziert [115]. Zytogenetisch betrachtet wurden in früheren Untersuchungen noch ein größerer Klon und zwei kleinere Subklone innerhalb der Population detektiert [140], in aktuelleren Untersuchungen war nur noch ein Klon nachweisbar, der sich eventuell im Laufe der kulturellen Züchtung durchsetzen konnte [141].

Die Sanger-Sequenzierung des Rearrangements der Zelllinie Granta-519 in dieser Arbeit lieferte eine Charakterisierung des V-Gens und –Alles als IGHV4-59*01 oder IGHV4-59*07 und einen Identity-Score der V-Region von 100 %, wobei die Position der Referenzsequenz, die die beiden Allele unterscheidet, außerhalb des amplifizierten Bereiches lag. Auch die sequenzstärksten Clonotypes (AA) der NGS-Datensätze wiesen eine Übereinstimmung mit einer oder beiden möglichen Gen- und Allel-Charakterisierungen, sowie einen Identity-Score von 100 % auf, wodurch das VH-Segment der Zelllinie als unmutiert einzustufen ist, ausgehend von einem Cut-Off-Wert von 98 % V-Region Identity zur Abgrenzung mutierter von unmutierten VH-Segmenten [66]. Die Charakterisierung des Rearrangements und die Einstufung als unmutiert stehen in Einklang mit den Ergebnissen einer früheren Untersuchung der Zelllinie Granta-519, die ebenfalls keine Mutationen im VH-Segment der Zelllinie feststellte und das Rearrangements IGHV4-59*01 detektierte [142]. Die durch den NGS-Prozess in dieser Arbeit detektierten, als Varianten innerhalb der Lymphompopulation eingeordneten Clonotypes wiesen einen Identity-Score von

98,9 % bis 100 % auf und gelten damit ebenfalls als unmutiert. Alle Abweichungen der Sequenzen untereinander wären damit am ehesten nicht auf SHM zurückzuführen, sondern vermutlich Artefakte des PCR- und NGS-Prozesses.

Vor diesem Hintergrund kann auch die Detektion verschiedener VH-Gensegmente und –Allele innerhalb der VH4-Familie in den Varianten des Zelllinien-Rearrangements auf die Präsenz von Artefakten zurückgeführt werden. So könnten Sequenzalterationen an entscheidenden Positionen darin resultieren, dass anstatt der Gen- und Allel-Kombination IGHV4-59*01 ein anderes Allel oder gar Gen innerhalb der Familie die größere Übereinstimmung mit der alterierten Sequenz aufweist und die Sequenz dieser abweichenden Gen- und Allel-Kombination zugeordnet wird. Das Alignment der Sanger-Sequenz der Zelllinie hatte gezeigt, dass zu den Referenzsequenzen von zwei weiteren Allelen nur an einer Position abweichende Nukleotide vorhanden waren, zu einem weiteren Allel an zwei Positionen. Da verschiedene Allele innerhalb einer VH-Segmentfamilie sich nur an einer einzigen Nukleotidposition unterscheiden können, kann ein einzelner Nukleotidaustausch an diesen Stellen zur Zuordnung eines abweichenden Allels und zur Definition eines neuen Clonotypes (AA) führen.

Als alternative Erklärung für abweichende VH-Segmente innerhalb der Population der Zelllinie Granta-519 könnte dienen, dass Clonotypes (AA) fälschlicherweise durch den Untersucher als derselben Population angehörig interpretiert wurden, obwohl es sich lediglich um ähnliche Rearrangements verschiedener Populationen handelt. Dies ist zwar in Einzelfällen nicht auszuschließen, jedoch für den Großteil der untersuchten Clonotypes (AA), sowohl der Zelllinienverdünnungsproben als auch der untersuchten Patientenproben, unwahrscheinlich, da sehr große Übereinstimmungen der Clonotypes (AA) für die Zuordnung zu einer Population vorausgesetzt wurden. So zeigten Clonotypes (AA) der Population der Zelllinie Granta-519 beispielsweise übereinstimmende DJ-Rearrangements auf Gen- und Allelebene, sowie gleiche Sequenzlängen und identische AA-Sequenzen der hypervariablen CDR3 bei lediglich abweichendem V-Gen innerhalb der VH4-Familie. Die

Wahrscheinlichkeit ist höher, dass es sich um Rearrangements derselben Population mit aus unklaren Gründen abweichenden V-Genen und –Allelen handelt, als dass Rearrangements verschiedener Populationen und aufgrund dessen mit unterschiedlichen V-Genen und –Allelen zufällig übereinstimmende Sequenzlängen, DJ-Rearrangements und CDR3 aufweisen.

Als dritte Möglichkeit könnten verschiedene VH-Segmente innerhalb einer Lymphompopulation durch VH-Replacement bedingt sein, einen Vorgang, der es B-Zellen erlaubt, nach erfolgter VDJ-Rekombination separat VH-Segmente auszutauschen und der an Maus-Lymphomzelllinien [143, 144] und menschlichen B-Zellen [145, 146] beobachtet wurde. Wie in einer Studie im Falle zweier MALT-Lymphome berichtet wurde, können so Lymphompopulationen entstehen, innerhalb derer IGH-Rearrangements mit identischen CDR3/JH-Regionen aber verschiedenen VH-Genen nachweisbar sind [147]. Interessanterweise wies einer dieser Fälle ein ähnliches IGH-Repertoire wie im hier beschriebenen Fall der Zelllinie Granta-519 auf: Neben dem in den klonalen Sequenzen hauptsächlich nachweisbaren Segment IGHV4-59, wiesen Teile der Rearrangements die Segmente IGHV4-04 und IGHV4-61 auf. Dies entspricht in weiten Teilen dem mit dem Primerset FR2 analysierten Repertoire der Zelllinie Granta-519. Ein kausaler Zusammenhang, eventuell auf Basis ähnlicher Antigen-Milieus im Reifungsprozess, lässt sich jedoch aus den vorhandenen Daten nicht ableiten. Ein 2012 publizierter Algorithmus zum MRD-Nachweis in akuten lymphatischen Leukämien (ALL) berücksichtigte die Möglichkeit durch VH-Replacement entstandener Subklone in diagnostischen NGS-Datensätzen; die hier nachgewiesenen Varianten der Zelllinie Granta-519 erfüllen die durch diese Arbeit zur Identifikation solcher Klone etablierten Kriterien [139]. Eine Charakterisierung der Zelllinie Granta-519 ergab jedoch zumindest kein VH-Replacement im prädominanten Klon mit IGHV4-59 [148], wobei nicht ausgeschlossen ist, dass kleinere Populationen mit VH-Replacement, wie sie in der vorliegenden Arbeit mittels NGS möglicherweise detektiert wurden, in jener Studie aus methodischen Gründen nicht detektiert werden konnten.

Unabhängig von der hier untersuchten Zelllinie sind in früheren Studien in 16 bis 34 % der MCL-Fälle mutierte VH-Segmente beschrieben [77-80], während die

Rate in aktuelleren Studien bei 60 bis 70 % lag [72, 149]. Eine überproportionale Verwendung der VH-Gensegmente IGHV3-21, IGHV4-34 sowie IGHV4-59 ist ebenfalls beschrieben [77-80, 149]. Bezüglich der Frage, ob Rearrangements mit Verwendung von IGHV4-59 gehäuft mutiert oder unmutiert vorliegen, existieren einander widersprechende Ergebnisse [77, 149]. Auch die Anwendbarkeit des zitierten Cut-Off-Wertes von 98 % V-Region Identity zur Unterscheidung mutierter und unmutierter VH-Segmente, der an CLL etabliert und seither von vielen Autoren auf MCL angewandt wurde [66], wurde von einer aktuelleren Studie mit großem untersuchtem MCL-Kollektiv für diese Entität infrage gestellt, da gezeigt werden konnte, dass auch innerhalb von Fällen mit 98 bis 100 % V-Region Identity eine Gruppe von komplett unmutierten MCL und eine Gruppe von MCL mit minimaler Mutationslast unterschieden werden können [72].

Auf Grundlage der vorliegenden, anhand der Zelllinie Granta-519 erhobenen Daten, ist es also nicht mit letzter Sicherheit möglich, Varianten innerhalb der detektierten Populationen als biologisch oder methodisch bedingt einzustufen, wobei wahrscheinlich ist, dass zumindest ein Teil der Variabilität auf Artefakte zurückzuführen ist. Insbesondere bezüglich der MCL-Diagnostik wäre es erstrebenswert ein NGS-Verfahren mit genauer Aussagekraft zu erarbeiten, da der Mutationsstatus und gewisse IGH-Rearrangement-Konstellationen bei MCL prognostische Relevanz haben können [77-80, 149].

Beurteilbarkeit des Mutationsstatus von B-CLL

Für B-CLL-Patienten ist der Mutationsstatus der VH-Region ein wichtiger prognostischer Marker, da das Vorhandensein von SHM in der VH-Region mit einem günstigeren Krankheitsverlauf assoziiert ist, als er bei Patienten zu erwarten ist, deren B-CLL eine unmutierte VH-Region aufweist [65, 150, 151]. Als Cut-Off zur Identifizierung unmutierter VH-Segmente ist der Wert von mehr als 98 % V-Region Identity verglichen mit dem nächstliegenden VH-Segment der Keimbahn-DNA etabliert [66]. Von den fünf in dieser Arbeit untersuchten CLL-Proben mit hohem Lymphomzellgehalt konnte in vier Fällen mittels Sanger-Sequenzierung und IMGT/V-Quest die VH-Region analysiert werden. In allen

Fällen ergab die Beurteilung des Mutationsstatus bezüglich des 98 % Cut-Offs dabei übereinstimmende Ergebnisse mit dem NGS-Verfahren, wobei die V-Region Identity des Clonotypes (AA) auf Rang 1 in den NGS-Datensätzen durchgehend höher ausfiel als in der Analyse der Sanger-Sequenz. Da diese Differenz bis zu 2,2 Prozentpunkte betrug, ist von der Möglichkeit auszugehen, dass im Rahmen des hier angewandten Vorgehens ein VH-Segment mit V-Region Identity nahe des Cut-Offs mittels Sanger-Sequenzierung als mutiert, mittels NGS allerdings als unmutiert eingestuft werden könnte. Die Untersuchung dieses möglichen Problems erfordert die Analyse eines größeren Kollektivs. Abweichungen zwischen den Analysen sind des Weiteren auf die Verwendung zweier verschiedener Primersets zurückzuführen. So erfolgte die Sanger-Sequenzierung nur mit VH-Primern für die FR2-Region, während das NGS-Verfahren auch mit dem Primerset FR1 durchgeführt wurde. Da mittels des Primersets FR1 ein größerer Abschnitt der VH-Region amplifiziert wird, ergeben auf diesem Primerset basierende Analysen zuverlässigere Ergebnisse bezüglich der V-Region Identity [152]. Wie auch an weiteren im Rahmen dieser Arbeit untersuchten Fällen beobachtet, lieferten die beiden Primersets zwar meist vereinbare Ergebnisse bezüglich zugeordneter V-Gene und –Allele, die auf dem Primerset FR2 basierenden Analysen lieferten jedoch häufig mehr mögliche Gen- und Allelkombinationen. Auch dies ist auf die geringere Größe der erzeugten Amplifikate mit dem Primerset FR2 zurückzuführen, deren Alignment mit der Referenzsequenz keine Differenzierung zwischen Allelen ermöglicht, die sich an Positionen außerhalb der Primerbindungsstellen unterscheiden.

Ein Vergleich NGS-basierter Untersuchung des VH-Mutationsstatus und derer mittels Sanger-Sequenzierung wurde auch von McClure et al. 2015 beschrieben [113]. In dieser Studie an 50 Patientenproben wurden ebenfalls in allen Fällen übereinstimmende Ergebnisse der beiden Methoden bezüglich Mutationsstatus und detektiertem V-Gen und –Allel erzielt. Unter den hervorgehobenen Vorteilen der NGS-basierten Untersuchung wurde auch auf die Notwendigkeit der manuellen Überprüfung und eventuell fehlerhaften Interpretation durch den Untersucher mittels Sanger-Sequenzierung gewonnener Sequenzen verwiesen, die in NGS-Verfahren entfällt. 2017 demonstrierten Stamatopoulos et al. Vorteile

der NGS-basierten Mutationsanalyse in B-CLL insbesondere im Hinblick auf die Detektion multipler Subklone, deren Präsenz auch prognostische Bedeutung beigemessen wird [153].

Wie in dieser Arbeit gezeigt, bietet NGS die Möglichkeit, den Mutationsstatus der VH-Region in B-CLL bereits im Rahmen der Klonalitätsdiagnostik zu ermitteln. Aufgrund von Überlagerungen durch eine mitamplifizierte polyklonale Hintergrundpopulation ist die Sanger-Sequenzierung klonaler IGH-Rearrangements häufig problematisch. Wie in dieser Arbeit beispielsweise an der FL-Probe S21 gezeigt, ermöglicht NGS die Sequenzierung eines klonalen Rearrangements auch bei prominenter Hintergrundpopulation und könnte in problematischen CLL-Proben mit prominenter Hintergrundpopulation zum Einsatz kommen. Im CLL-Fall S17 jedoch, misslang sowohl der Versuch der Sanger-Sequenzierung als auch der Versuch, durch NGS einen prominenten Clonotype (AA) zu sequenzieren, sodass dieser potenzielle Anwendungsbereich des NGS an weiteren Fällen untersucht werden sollte. Perspektivisch wird der NGS-basierten Mutationsdiagnostik von B-CLL sowohl hinsichtlich verbesserter Sensitivität als auch der individuellen Risikostratifizierung zukünftig ein großes Potential beigemessen [154, 155].

Verfolgbarkeit von Lymphompopulationen und Varianten in Patientenproben

Im Rahmen dieser Arbeit wurden von sechs Patienten jeweils drei Proben von unterschiedlichen Zeitpunkten untersucht. In vier der sechs Fälle gelang der Nachweis einer Lymphozytenpopulation, die zu allen drei Zeitpunkten in mindestens einem der Datensätze repräsentiert war. Besonders die Datensätze der B-CLL von Patient 18 zeigten eindeutig die Detektierbarkeit eines klonalen IGH-Rearrangements im zeitlichen Verlauf mittels NGS. Eine gewisse Heterogenität innerhalb der Population im Bereich der CDR3 auf AA- und Nukleotidebene ließ sich zwar darstellen. Die Beobachtung, dass der Clonotype (AA) auf Rang 1 in allen Datensätzen dieselbe AA-Sequenz der CDR3 aufwies, zeigt dennoch eine gewisse Reproduzierbarkeit nicht nur des Klonalitätsbefundes, sondern auch von Details des prominenten IGH-

Rearrangements. Die V-Region Identity der Clonotypes (AA) der Population von 99 bis 100 % kategorisierte die CLL als unmutiert und lässt daher vermuten, dass zumindest ein Teil der Varianten innerhalb der Population Artefakte darstellten. Auch die Datensätze von Patient 19 erlaubten die Detektion einer Population im zeitlichen Verlauf. In den beiden Proben vor Stammzelltransplantation zeigte sich die Population klonal überrepräsentiert. Nach allogener Stammzelltransplantation fanden sich Clonotypes (AA) der Population nicht unter den top zehn Clonotypes (AA), waren aber dennoch im Datensatz FR1 mit mindestens sieben zugeordneten Sequenzen vertreten. In Zusammenschau mit jenen der hochverdünnten Zelllinienverdünnungsproben verdeutlichen die Datensätze von Patient 19 die Anwendbarkeit des hier verwendeten NGS-basierten Verfahrens zum MRD-Monitoring, wie unten diskutiert. Einschränkend ist festzustellen, dass der bloße Nachweis der Sequenz des zuvor identifizierten Lymphom-Rearrangements nicht als beweisend für das Vorhandensein von MRD zu werten ist; Grund hierfür ist, dass keine Unterscheidung von möglichen prä-malignen Vorläuferzellen mit demselben IGH-Rearrangement möglich ist, die nach Eradikation der malignen Zellen noch in Proben des Patienten nachweisbar sein könnten.

Die Datensätze des DLBCL-Patienten 16 sowie des FL-Patienten 20 erlaubten zwar die Detektion einer Population zu allen drei Zeitpunkten, allerdings waren pro Patient in mehreren Datensätzen Clonotypes (AA) nachweisbar, die der Population nicht zugehörig waren und denen ein größerer Teil der Gesamtsequenzen zugeordnet wurde. Im Fall von Patient 16 betraf dieses Phänomen nur die Datensätze, zu den Zeitpunkten 1 und 2, die auf dem Primerset FR1 basierten, während in den Datensätzen FR2 stets die mutmaßliche Lymphompopulation den Clonotype (AA) auf Rang 1 stellte. Es ist demnach möglich, dass sich das Lymphom-Rearrangement aufgrund von Mutationen im Bereich der Primerbindungsstelle mit dem Primerset FR1 nur eingeschränkt amplifizieren ließ. In beiden Datensätzen zu Zeitpunkt 3, stellte die mutmaßliche Lymphompopulation den Clonotype (AA) auf Rang 1 und es waren keine prominenten Nebenpopulationen nachweisbar; dennoch wurde der Lymphompopulation im Datensatz FR1 ein deutlich geringerer Anteil der

Gesamtsequenzen zugeordnet als im Datensatz FR2, was zusätzlich für eine weniger effektive Amplifizierung mit dem Primerset FR1 spricht. In den Datensätzen von Patient 20 fanden sich in den drei vorhandenen Datensätzen der Zeitpunkte 1 und 2 Clonotypes (AA) auf Rang 1, die nicht der mutmaßlichen Lymphompopulation zuzuordnen waren. Im Datensatz FR2 zum Zeitpunkt 2 wurde einem solchen Clonotype (AA) der Großteil der Sequenzen zugeordnet. Die Bedeutung dieser Beobachtung bezüglich der Möglichkeit falsch positiver NGS-Klonalitätsbefunde und des Nutzens von Doppelansätzen zur Überprüfung ihrer Reproduzierbarkeit wurden bereits diskutiert.

In den Datensätzen der Patienten 17 und 21 gelang nicht der Nachweis einer probenübergreifenden Population. In den Datensätzen beider Fälle zeigte sich entweder ein polyklonales Verteilungsmuster der Clonotypes (AA), eine geringe Anzahl gewonnener Sequenzen oder ein geringer Anteil an Sequenzen, die einem IGH-Rearrangement zugeordnet werden konnten. Die DNA-Qualität war dabei in allen bis auf eine Probe ausreichend mit amplifizierbaren Fragmenten von mindestens 300 bp. Eventuell waren SHM im Bereich der Primerbindungsstellen ursächlich. Wie bereits diskutiert wäre in derartigen Fällen ohne Nachweis eines klonalen IGH-Rearrangements die Anwendung weiterer Primersets, beispielsweise zur IGK-Amplifikation und –Sequenzierung eine Möglichkeit des Klonalitätsnachweises und der Verfolgung der Lymphompopulation.

Zusammenfassend konnte anhand der Verlaufsfälle gezeigt werden, dass NGS und Analyse mittels IMGT/HighV-Quest prinzipiell die Möglichkeit der Detektion klonaler Populationen und deren Nachverfolgung bieten, auch bei geringer Repräsentation des entsprechenden Clonotypes (AA) in einem Datensatz. Probleme in der Detektion einer probenübergreifenden Population bei zwei Patienten, sowie die Präsenz sequenzstarker Clonotypes (AA) der mutmaßlichen Hintergrundpopulation in mehreren Datensätzen machen allerdings weitere Erprobungen dieses Vorgehens, eventuell in Kombination und im Vergleich mit weiteren molekularpathologischen Verfahren wie der GeneScan-Analyse, sowie mit Fokus auf die Möglichkeit methodisch bedingt überamplifizierter IGH-Rearrangements nötig.

Ansätze zur Beurteilung klonaler Evolution

Zum Zweck der Nachverfolgung datensatzübergreifender Lymphompräsenz, war hier das Vorgehen praktikabler, prominente Clonotypes (AA) und deren Varianten zu Populationen zusammenzufassen und vielmehr die Präsenz der Population zugehöriger Clonotypes (AA) zu überprüfen, als jede Variante für sich betrachtet nachzuverfolgen. Dieser Ansatz jedoch erlaubt nicht die Verfolgung klonaler Evolution im zeitlichen Verlauf, für die mit höherer Trennschärfe Veränderungen auf Nukleotidebene zwischen zwei Zeitpunkten detektiert werden müssten. In drei exemplarischen Fällen, der Zelllinie Granta-519, den CLL-Verlaufspuren von Patient 18, sowie den FL-Verlaufspuren von Patient 19 wurden Varianten innerhalb der jeweiligen Population zu einem Zeitpunkt, im Falle der Patientenproben auch im zeitlichen Verlauf dargestellt. Die Anzahl verschiedener Varianten innerhalb einer Population blieb dabei in Abhängigkeit der Anzahl zugeordneter Sequenzen stets ähnlich und zeigte keine Zunahme im zeitlichen Verlauf. Anhand des Beispiels der Zelllinie Granta-519 zeigte sich außerdem eine Abhängigkeit der Anzahl verschiedener Varianten vom verwendeten Primerset. Die CLL-Proben der Datensätze von Patient 18, die aufgrund ihrer V-Region Identity von 99 bis 100 % als unmutiert gilt, wies mehr Varianten innerhalb ihrer CDR3 auf AA-Ebene auf als die FL-Proben von Patient 19 mit einer V-Region Identity von 90 bis 91 %. Die Betrachtung der Clonotypes (nt) pro Clonotype (AA), also die Anzahl verschiedener Nukleotidsequenzen der CDR3 bei gleicher AA-Sequenz der jeweiligen Clonotypes (AA) auf Rang 1, ergab ebenfalls mehr Varianten innerhalb der Clonotypes (AA) der CLL-Probe sowie der Zelllinie Granta-519 als innerhalb des Clonotypes (AA) auf Rang 1 der FL-Probe. Dies steht im Gegensatz zu einer früheren Untersuchung von Logan et al., die in einem Kollektiv unmutierter CLL-Fälle nur sehr selten und in geringen Zahlen Varianten der Lymphompopulation detektierten [138].

Wenn im Rahmen dieses Verfahrens detektierte Varianten innerhalb einer prominenten Population zu großen Teilen auf methodisch bedingte Sequenzalterationen zurückzuführen wären, so wäre ein Zusammenhang zwischen der Anzahl Clonotypes (nt) pro Clonotype (AA) mit der Länge der untersuchten CDR3 und der Anzahl dem Clonotype (AA) zugeordneter

Sequenzen zu erwarten. Eine detailliertere Analyse dieses möglichen Zusammenhanges anhand einer größeren Probenzahl könnte dazu beitragen, methodisch bedingte Artefakte von biologisch bedingten Mutationen zu unterscheiden. Dies wäre Voraussetzung, um zuverlässige Aussagen bezüglich intraklonaler Evolution innerhalb der Untersuchten Populationen treffen zu können.

Ein zusätzlich zu verfolgender Ansatz, der Aufschluss über den jeweiligen Anteil biologisch und methodisch bedingter Sequenzvarianten liefern könnte, stellt die Betrachtung der Positionen von Basenaustauschen innerhalb des IGH-Rearrangements dar. Wie einleitend erläutert, stellen die drei CDRs des IGH-Rearrangements die Regionen größter Variabilität dar, indem diese, im Gegensatz zu den FRs, bevorzugt Ziel somatischer Hypermutation sind [17, 25, 31, 32]. Bei der Analyse des Basenaustauschprofils innerhalb eines IGH-Rearrangements, wäre daher auch mit einer höheren Dichte varianter Basen innerhalb der CDRs, als im Bereich der FRs zu rechnen. Ein hoher Anteil an Basenaustauschen innerhalb der FRs, sowie eine gleichmäßige Verteilung varianter Basen über alle Regionen des IGH-Rearrangements, ließe auf einen methodischen Einfluss schließen.

5.2.3 Methodische Schwächen in der Beurteilung von Varianten und klonaler Evolution

Ursachen von Artefakten in NGS-Datensätzen

In verschiedenen Schritten dieser Arbeit bestand die Möglichkeit der Entstehung von Artefakten, die sich im späteren NGS-Datensatz als Sequenzalteration und damit falsch positive Varianten manifestieren konnten.

Die Eigenschaft von FFPE-Gewebe, nicht reproduzierbare Sequenzartefakte, d.h. fixierungs- und nicht biologisch bedingte Sequenzvarianten, in PCR-Verfahren zu generieren und Schwierigkeiten, eine ausreichende Amplifikation von DNA aus FFPE-Material zu erreichen, sind bekannt [119, 134, 156-158]. Auch in mehreren Studien zum Einsatz von NGS-Verfahren an FFPE-Material

zeigte sich eine erhöhte Rate von Basenaustauschen, die schwer von tatsächlichen Varianten innerhalb einer malignen Zellpopulation zu unterscheiden sind, in erster Linie in Form der Reduktionen von Cytosin zu Thymin, sowie von Guanin zu Adenin [159-163]. In den unter 4.5.3 dargestellten Sequenzvarianten innerhalb der CDR3 von Patient 18 machten Austausche von Cytosin mit Thymin tatsächlich 34,5 % und Austausch von Thymin mit Cytosin 30,9 % aller Basenaustausche aus, während der Austausch von Guanin mit Adenin deutlich seltener auftrat. Bei Patient 19 war der Austausch von Thymin mit Cytosin mit 40 % mit Abstand am häufigsten. Beide Proben waren aus FFPE-Material isoliert worden, sodass wahrscheinlich zumindest ein Teil der Sequenzvarianten als Artefakte zu deuten ist. Allerdings wiesen auch die Varianten der Zelllinie Granta-519, deren Proben nicht aus FFPE-Material isoliert worden waren mit 27,3 % am häufigsten den Austausch von Cytosin mit Thymin und an zweithäufigster Stelle mit 18,2 % den Austausch von Guanin mit Adenin auf. Eine vergleichende Analyse des Basenaustauschprofils einer größeren Fallzahl von CDRs, sowohl aus FFPE- als auch aus Frischmaterial, könnte die Einordnung von Sequenzalterationen als biologische Variante oder als Artefakt erleichtern.

PCR-Artefakte durch die Inkorporation fehlerhafter Basen durch Tag-Polymerasen wurden in verschiedenen Arbeiten demonstriert [130, 164-167], und müssen daher auch als Ursache von Sequenzalterationen in den hier erhobenen NGS-Datensätzen in Betracht gezogen werden. Insbesondere die vergleichsweise hohe Anzahl von PCR-Zyklen im IGH-Amplifikationsprozess in dieser Arbeit könnte dabei das Auftreten derartiger Artefakte begünstigt haben.

Die häufigste Fehlerform, die direkt auf die Illumina-Sequenzierung zurückzuführen ist, sind Basensubstitutionen [168, 169]. Aktuelle Untersuchungen von Fehlerprofilen und –ursachen der Illumina MiSeq-Plattform ergaben jedoch, dass insbesondere das experimentelle Design bezüglich Library-Erstellung und verwendeter Forward- und Reverse-Primer einen Einfluss auf Fehlerraten in generierten Sequenzen hat, und dass daher für individuelle experimentelle Vorgehensweisen eigene Fehlerevaluationen vorgenommen werden sollten, um beispielsweise eine falsch hohe Einschätzung der Diversität

in Proben zu vermeiden [129, 169]. Die EuroClonaloty-NGS Arbeitsgruppe beispielsweise etablierte hier 2019 zwei Ansätze zur Qualitätskontrolle in NGS-basierten Assays zur Klonalitätsdiagnostik [170]. Eine standardisierte Mischung lymphozytärer DNA, die das volle Repertoire möglicher Immunglobulingene bzw. TCR enthält, dient dazu, in Probedurchläufen PCR-Bias oder ungewöhnliche Amplifikationsverhältnisse einzelner Primer zu detektieren. Zweitens wird den diagnostischen Proben in definierter Menge DNA einer gut charakterisierten B- bzw. T-Zelllinie beigefügt, die in den NGS-Datensätzen als interne Qualitäts- und auch Quantitätsreferenz dient.

Bewertung des bioinformatischen Auswertungsansatzes

Die oben beschriebenen Fehlerquellen in der Library-Erstellung und im NGS-Prozess stellen eine Herausforderung bei der Untersuchung von Lymphozyten- und Lymphompopulationen mittels NGS dar, in denen zwischen methodisch bedingten Artefakten und biologisch bedingten Varianten unterschieden werden muss. Dabei können Sequenzartefakte, die beispielsweise in einem frühen PCR-Zyklus entstanden sind, in mehr Sequenzen präsent sein als seltene, durch SHM entstandene Subklone einer Lymphompopulation. Eine Untersuchung von TCR-Repertoires mittels Illumina-Sequencing ergab, dass 1 bis 6 % der mittels NGS generierten Sequenzen in der CDR3 methodisch bedingte Fehler enthielten, die größtenteils einzelne Nukleotidsubstitutionen darstellten [171]. Demnach könnten fehlerhafte Sequenzen am ehesten eliminiert werden, wenn einfache Nukleotidsubstitutionen einzelner Sequenzen verglichen mit einer häufigeren Indexsequenz ignoriert würden. Dabei könnte dieses Vorgehen jedoch dazu führen, dass tatsächliche, biologisch bedingt variante Sequenzen fälschlicherweise von der Beurteilung des untersuchten Repertoires ausgeschlossen werden.

Campbell et al. untersuchten IGH-Rearrangements in B-CLL zur Detektion phylogenetischer Verhältnisse zwischen Subklonen [172]. Die Häufigkeit von Sequenzartefakten wurde ausführlich anhand eines untersuchten IGH-Locus der Keimbahn evaluiert. Charakteristika PCR- und NGS-basierter Artefakte wurden

ermittelt und ein Algorithmus zur Erkennung der Artefakte erarbeitet. Zusätzlich wurde für die Einordnung eines Klon als biologisch bedingt gefordert, dass er innerhalb eines Datensatzes mindestens drei Mal beobachtet wurde. Einige weitere Arbeiten demonstrierten bereits den Nutzen von NGS des IGH-Repertoires zur Detektion verschiedener Klone innerhalb einer Lymphompopulation und zur Verfolgung klonaler Evolution im zeitlichen Verlauf, und griffen auf ähnliche bioinformatische Auswertungsansätze zurück [127, 173-175], deren Fehlen in dieser Arbeit einen Vergleich der erzielten Ergebnisse nicht zulässt.

Zur validen Beantwortung der Teilfragestellung dieser Arbeit nach intraklonaler Diversität sowie klonaler Evolution, müssten, in Zusammenschau mit vorliegenden Arbeiten zu diesem Thema, also folgende Eckpunkte des Vorgehens miteinbezogen werden: (i) Die Ermittlung der Übereinstimmungsgrade analysierter IGH-Sequenzen nicht nur mit der Keimbahn-DNA, sondern auch untereinander bzw. mit einer prominenten Indexsequenz, zur quantitativen Beurteilung der Wahrscheinlichkeit klonaler Verwandtschaft; (ii) ein Mechanismus zur Aussortierung bzw. Korrektur seltener Sequenzen mit geringer Abweichung von einer Indexsequenz, um dem Problem artifizieller Varianten zu begegnen; Ein- und Ausschlussparameter dieses Mechanismus müssten durch (iii) systematische Untersuchung des Einflusses methodischer Faktoren (insbesondere der Verwendung von FFPE-Material) auf die Häufigkeit von Sequenzartefakten ermittelt werden, um Sensitivität und Spezifität der Detektion echter Varianten zu optimieren.

5.3 Möglichkeiten und Grenzen der NGS-Diagnostik zum Nachweis kleiner Zahlen klonaler Zellen

5.3.1 Nachweisgrenze klonaler Populationen mittels NGS und Eignung zum MRD-Nachweis

Um die Nachweisgrenze klonaler Populationen mittels NGS zu ermitteln, wurden im Rahmen dieser Arbeit Verdünnungsproben von B-Zell-Lymphom-Zelllinien-

DNA in PBMC-DNA hergestellt und untersucht. Auch in Verdünnungen von 1:1.000 konnte das IGH-Rearrangement der Lymphomzelllinie in den NGS-Datensätzen detektiert werden. Da B-Zellen allerdings nur ca. 5 – 20 % der PBMCs ausmachen [176], war die tatsächlich erzielte Verdünnung der klonalen B-Zell-DNA in polyklonaler B-Zell-DNA im Sinne der Simulation einer Lymphomprobe mit geringem Lymphomzellgehalt eine geringere als 1:1.000. In der Berechnung ausgehend von 15 % B-Zellen in den PBMCs wurde so eine Verdünnung von etwa 1:150 in der höchsten hier untersuchten Verdünnung erzielt. Diese lieferten einen Anteil des Lymphom-Rearrangements an allen als Rearrangement erkannten Sequenzen von 2,3 % (Primerset FR1) bzw. 1,1 % (Primerset FR2), sodass hier mit keiner Verdünnung das untere Detektionslimit unterschritten wurde, welches somit folglich nicht klar festgelegt werden konnte.

Mehrere Arbeiten der letzten Jahre haben sich bereits als Teil- oder Hauptfragestellung mit dem Einsatz von NGS zum Monitoring von MRD auseinandergesetzt, die Sensitivität verschiedener Vorgehensweisen ermittelt und Vorteile gegenüber anderen Methoden der MRD-Untersuchung herausgearbeitet. Bisherige Methoden zum MRD-Nachweis und –Monitoring sind insbesondere die Allelspezifische Oligonukleotid-PCR (ASO-PCR) und die Durchflusszytometrie („flow cytometry“, FC) [111, 177].

Ein dem Vorgehen dieser Arbeit ähnliches Setup zur Ermittlung der Sensitivität des NGS zur Detektion klonaler Lymphozytenpopulationen wurde von Boyd et al. 2009 verwendet [105]. Die Blutprobe eines CLL-Patienten wurde gewonnen, diese in Blut eines gesunden Spenders in verschiedenen Verhältnissen verdünnt und die Verdünnungsproben einer IGH-Amplifikation und NGS unterzogen. Die Detektion der klonalen Zellen gelang in dieser Studie bis zur Verdünnung von 1:10.000..Auch Schumacher et al. ermittelten 2014 die Sensitivität der Detektion von MRD in ähnlicher Weise in ihrer Studie zur Klonalitätstestung von T-Zell-Neoplasien mittels NGS [106] und detektierten klonale Zellen bis zu einem Gehalt von 0,1 %, in Einzelfällen bis 0,004 %.

Eine Studie von Logan et al. aus dem Jahr 2011 beschäftigte sich exklusiv mit der Anwendbarkeit von Hochdurchsatzsequenzierverfahren zum MRD-Nachweis

und –Monitoring [138], ermittelte eine Sensitivität der NGS-Methode von 1:100.000 (10^{-5}) und benannte als größten Vorteil ihres Ansatzes zum MRD-Nachweis und –Monitoring die „breite Anwendbarkeit von Konsensus-Primern, ohne die Notwendigkeit von Anpassung an einzelne Patienten“ im Gegensatz zur ASO-PCR. Dies würde zu einer „erhöhten Verfügbarkeit von MRD-Quantifizierung bei CLL-Patienten“ führen. Diese Schlussfolgerungen konnten auch von Faham et al. bestätigt werden [139]. Ihr MRD-Assay bot bei einer Sensitivität von 10^{-6} zusätzlich die Möglichkeit, die Absolutzahl der leukämischen Moleküle in der Ausgangsreaktion zu ermitteln.

Ein weiterer Vergleich von NGS-basierter gegenüber FC-basierter MRD-Testung von Wu et al. 2012 demonstrierte ebenfalls die Überlegenheit von NGS bezüglich Sensitivität anhand einer Kohorte von T-ALL-Fällen [108]. Die Möglichkeit, MRD mittels NGS zu quantifizieren, nutzen Logan et al. 2013 und 2014 um anhand von MRD-Last zu bestimmten Zeitpunkten sowie anhand der MRD-Verdopplungszeit nach allogener Stammzelltransplantation Aussagen bezüglich des individuellen Rückfallrisikos von CLL- [110] bzw. ALL-Patienten [109] zu treffen. Eine Studie von Ladetto et al. von 2014 an einem Kollektiv aus ALL-, MCL- und MM-Fällen attestierte NGS und ASO-PCR eine „vergleichbare“ Sensitivität und betonte den Nutzen von Multi-Target-Ansätzen [128]. Brüggemann et al. nutzten 2019 die EuroClonality-NGS-Primersets in einer multizentrischen Validierungsstudie zur Identifikation von MRD-Markern in ALL und konnten Vorteile gegenüber etablierten Niedrigdurchsatzverfahren zur Markeridentifikation aufzeigen [178].

2019 demonstrierten Studien von Yao et al. sowie Rustad et al. den Nutzen der kommerziell verfügbaren multiplex-PCR-basierten LymphoTrack-Assays sowohl zur Klonalitäts- als auch MRD-Diagnostik mittels NGS an MM-Patienten [179, 180].

Insbesondere die Datensätze des im Rahmen dieser Studie untersuchten FL-Patienten 19 zeigen die Möglichkeit auf, wie mittels NGS eine Evaluation von MRD erfolgen kann: Nach Detektion und Charakterisierung des klonalen Rearrangements in zwei Proben, konnte nach erfolgter allogener Stammzelltransplantation mit dem Primerset FR1 eine geringe Zahl der

Lymphompopulation zugehöriger Sequenzen detektiert werden. Die valide Reproduzierbarkeit klonaler Rearrangements in Diagnose- und Verlaufs- bzw. Rezidivproben von NHL mittels NGS und damit der potenzielle Nutzen zur MRD-Diagnostik wurden auch von Scheijen et al. 2019 bestätigt [112].

5.3.2 Aussagekraft von NGS- und GeneScan-basierter Klonalitätsuntersuchung bei Proben mit geringem Lymphomzellgehalt

In dieser Arbeit wurden drei Glaskörperproben von IOL untersucht, um Vorteile der NGS-basierten Klonalitätsstestung gegenüber der GeneScan-Analyse bei Proben mit geringem Lymphomzellgehalt zu ermitteln. Intraokuläre Lymphome sind sehr selten und entstehen entweder isoliert (Primär intraokuläre Lymphome, PIOL) oder gemeinsam mit einem Lymphom des Zentralnervensystems („primary central nervous system lymphoma“, PCNSL) [181]. Sowohl PIOL als auch PCNSL entstammen meist der B-Zell-Reihe und sind als DLBCL zu klassifizieren [10].

Die Aufwändigkeit des diagnostischen Vorgehens bei Verdacht auf ein IOL [182, 183] veranschaulicht die diagnostische Herausforderung, die dieses präsentiert. Die verwendeten Vitrektomiepräparate und Glaskörperaspirate sind oft arm an Zellmaterial und können damit sowohl zu falsch negativen als auch zu falsch positiven Befunden führen. Nicht zuletzt PCR-basierte Methoden sind in diesem Zusammenhang problematisch, da bei geringer Lymphozytenzahl und dadurch eingeschränktem Rezeptorrepertoire pseudoklonale Bandenmuster entstehen und zu falsch positiven Befunden führen können [184, 185]. Die Anzahl an enthaltenen Zellen kann noch zusätzlich verringert sein durch Lyse der Zellen nach Probeentnahme oder durch vorangegangene empirische Therapie mit Glukokortikoiden, als Resultat des Phänomens, dass IOL sich häufig klinisch wie entzündlichen Erkrankungen des Auges präsentieren [186-189]. Dem Problem falsch positiver PCR-Befunde bei geringer Zellzahl kann begegnet werden, indem die Ergebnisse in mehreren Ansätzen auf Reproduzierbarkeit hin getestet werden [95].

Die Anwendbarkeit und Überlegenheit gegenüber FC und Zytologie verschiedener PCR-basierter Methoden, teils in Kombination mit Mikrodissektion und Bestimmung des intraokulären Zytokinprofils konnten gezeigt werden [190-194]. Ein weiterer Ansatz in Erprobung macht sich die hohe Frequenz von MYD88-Mutationen in PCNSL bzw. PIOL zunutze [195], indem der Nachweis der Mutation in Glaskörperproben zur Detektion maligner Zellen genutzt wird [196].

Für den Nachweis maligner Zellen in Liquor („cerebrospinal fluid“, CSF) sind ähnliche Probleme bekannt, wie für die IOL-Diagnostik [197-199]. Die Überlegenheit gegenüber FC und Zytologie zweier semi-nested PCR-Ansätze konnte auch an CSF gezeigt werden, [200, 201]. Jedoch kann die hierfür nötige DNA-Menge in ausreichender Qualität häufig nicht gewonnen werden [201].

Bisher existieren wenige Daten zur Performance der BIOMED-2 Primersets für die Fragestellung nach Klonalität in Glaskörperproben und CSF und die beschriebenen Ansätze sind aufgrund verschiedener Primersets und PCR-Modalitäten schwer mit dem Vorgehen in dieser Studie vergleichbar. Arbeiten zum Einsatz von NGS-basierten Methoden zum Nachweis von Lymphomzellen in Glaskörperaspiraten und CSF liegen bisher nicht vor. Eine Studie von Liu et al. 2015 beschäftigte sich jedoch mit der Detektion maligner B-Lymphozyten mittels BIOMED-2 Primersets in CSF und Glaskörperaspiraten [202]. Die Gruppe analysierte 19 CSF- und eine Glaskörperprobe mittels BIOMED-2 Primersets FR1, FR2 und FR3 und verglichen zusätzlich standardmäßige Säulen-basierte DNA-Isolationsmethoden mit direkter Lyse mittels Detergenzien und Proteinase K. Letzteres Vorgehen erwies sich dabei als besser geeignet bei Proben mit geringer Zellzahl als das auch in dieser Arbeit verwendete Säulen-basierte Verfahren. Mittels direkter Lyse konnte in Verdünnungsansätzen mit einer Mindestzahl von 20 Lymphomzellen ohne Hintergrundpopulation Klonalität nachgewiesen werden. Dennoch erwiesen elf der 20 untersuchten Fälle sich mittels BIOMED-2 PCR als falsch negativ, was auch mit der geringen Anzahl verwendeter Primersets zusammenhängen könnte, wie bereits diskutiert. Aufgrund der geringen Zahl durch Liu et al. untersuchter Fälle und dem Fehlen von Follow-Up-Daten der untersuchten Patienten ließ sich schließlich keine Aussage zu Sensitivität und Bedeutung des erprobten BIOMED-2 PCR-

Protokolls an CSF- und Glaskörperproben treffen. Auch eine Studie von Bonzheim et al. wies auf hohe Raten falsch negativer Klonalitätsbefunde der BIOMED-2 Protokolle an Glaskörperaspiraten hin [196].

Das in dieser Arbeit erprobte Vorgehen der Amplifikation mittels BIOMED-2 FR1- und FR2-Primersets mit Fluidigm-Anhang und anschließender massiv-paralleler Sequenzierung lieferte in den drei untersuchten Proben eine Charakterisierung des IGH-Repertoires und bestätigte den klonalen GeneScan-Befund der Probe S46. Im Falle der Probe S47 mit pseudoklonalem und der Probe S48 mit polyklonalem GeneScan-Befund, blieb auch das NGS-Verfahren falsch negativ. Aufgrund der jeweils zu geringen Zahl gewonnener und als IGH-Rearrangement erkannter Sequenzen ließ sich keine Aussage über das Vorliegen von Klonalität treffen. Bei Vorliegen pseudoklonaler GeneScan-Analysen könnte NGS jedoch einen Beitrag dazu leisten, falsch positive Befunde zu identifizieren, indem über die Anzahl verschiedener nachweisbarer Clonotypes das für falsch positive Befunde verantwortliche eingeschränkte IGH-Repertoire dargestellt wird. Dies bleibt mittels systematischer Untersuchung größerer Probenkollektive zu ermitteln.

5.4 Schlussfolgerungen und Ausblick

Die vorliegende Arbeit veranschaulicht die vielfältigen Anwendungsmöglichkeiten einer NGS-basierten Vorgehensweise zur Klonalitätsdiagnostik in malignen Lymphoproliferationen, die über die Beantwortung der Frage nach Klonalität hinausgehen und zusätzlich weitergehende molekularpathologisch relevante Informationen über die untersuchte Probe liefern können. Die Arbeit verdeutlicht gleichzeitig die Herausforderungen und Probleme, die sich bei der Interpretation der NGS-Daten verglichen mit etablierten Methoden der Klonalitätsdiagnostik ergeben.

Anhand der Zelllinienverdünnungsproben, konnten zunächst Charakteristika klonaler Populationen und Merkmale zur Unterscheidung von Rearrangements der Hintergrundpopulation ermittelt werden. Die erhobenen Daten bestätigten

dabei Ergebnisse vorausgegangener Arbeiten zu diesem Thema, insbesondere bezüglich der zuverlässigen Unterscheidbarkeit zwischen klonalen Rearrangements und Rearrangements der reaktiven Hintergrundpopulation. Die Charakterisierung überrepräsentierter klonaler IGH-Rearrangements in Proben mit hohem Lymphomzellgehalt erlaubte sowohl in Zelllinienverdünnungsproben als auch in den FFPE-Verlaufsproben von Patienten die Detektion mutmaßlicher Lymphompopulationen bei geringerem Lymphomzellgehalt und demonstrierte damit die Anwendbarkeit der Vorgehensweise zum Nachweis und Monitoring von MRD. Dabei könnten in weitergehenden Versuchsreihen höhergradige Verdünnungen von Zelllinienproben untersucht werden, um die untere Detektionsgrenze zu unterschreiten und diese definitiv festzulegen.

In der vorliegenden Arbeit wurden weder FFPE- noch Proben aus Frischmaterial von rein reaktiven Lymphozytenpopulationen mittels NGS untersucht. Die Erprobung des hier praktizierten Vorgehens an derartigen Kollektiven wäre ein nützlicher nächster Schritt, um weitere Unterscheidungsmerkmale zwischen poly- und monoklonalen Rearrangements zu erarbeiten und damit die Auswertungsansätze der NGS-Daten in Richtung höherer Sensitivität und Spezifität zu modifizieren.

Die Anwendung der erarbeiteten Auswertungsansätze zum Klonalitätsnachweis in Patientenproben aus FFPE-Gewebe demonstrierte die prinzipielle Eignung der Vorgehensweise, lieferte jedoch teils falsch negative Befunde. Zusammenfassend lässt sich dies insbesondere auf bekannte Schwächen der BIOMED-2 Primersets zurückführen, die in modifizierter Form zum Einsatz kamen. Die Erweiterung der NGS-basierten Analyse um zusätzliche Primersets des BIOMED-2 Repertoires, bspw. zur IGK-Amplifikation könnte die Sensitivität steigern. Gleichzeitig müsste dem Problem falsch positiver NGS-Befunde begegnet werden; den ersten Schritt hierzu kann die vergleichende Analyse von Doppelansätzen, beispielsweise mit unterschiedlicher DNA-Ausgangsmenge wie im GeneScan-Verfahren üblich, darstellen. Dies könnte ebenfalls der besseren Aussagekraft bei Proben mit geringem Lymphomzellgehalt, wie den in der vorliegenden Arbeit untersuchten Glaskörperaspiraten, dienen. Eine deutlich verbesserte Aussagekraft verglichen mit der GeneScan-Analyse konnte anhand

der drei untersuchten Fälle nicht erzielt werden, weswegen zusätzlich größere Fallzahlen analysiert werden sollten.

Ebenfalls konnte reproduziert werden, wie im Zuge der Klonalitätsdiagnostik mittels NGS eine Aussage über den Mutationsstatus detektierter IGH-Rearrangements getroffen werden kann, indem deren Sequenzen mit Indexsequenzen der Keimbahn-DNA verglichen und der Grad der Abweichung quantifiziert wird. Dies kann im Kontext mancher Entitäten wie der CLL von prognostischer Bedeutung sein, weshalb die Untersuchung größerer derartiger Probenkollektive und der Vergleich der Ergebnisse mit jenen der Sanger-Sequenzierung anzustreben ist.

Ein in der vorliegenden Arbeit nicht abschließend gelöstes Problem stellt die Unterscheidung biologisch bedingter Sequenzvarianten von methodisch bedingten Sequenzartefakten dar. Vergleichbare vorausgegangene Arbeiten implementierten verschiedene eigens entwickelte Auswertungsalgorithmen, um diesem Problem zu begegnen. Die Verwendung von IMGT/HighV-Quest erwies sich dabei als nur bedingt geeignet, da es keine Filterfunktion für seltene oder biologisch nicht plausible Rearrangement-Varianten bietet. Die Amplifikation und Untersuchung eines einzelnen, gut charakterisierten IGH-Rearrangements, könnte dazu beitragen, die Rate methodisch bedingter Sequenzalterationen zu ermitteln und Ansätze zu entwickeln, diese von biologisch bedingten Sequenzvarianten zu unterscheiden.

Massiv parallele Sequenzierverfahren halten großflächig Einzug in die molekularpathologische Routinediagnostik. Folglich ist zu vermuten, dass auch NGS-basierte Verfahren zum Klonalitätsnachweis maligner Lymphoproliferationen zunehmend an Bedeutung gewinnen werden. Zusätzlich verspricht die Technologie, durch weitere Anpassung und Optimierung der Amplifikations- und Sequenziermethodik, sowie wachsende Erfahrung in der komplexen Interpretation der generierten Daten, Fortschritte in der Evaluation von MRD, des Mutationsstatus maligner Lymphome, sowie intraklonaler Evolution.

6. Verzeichnisse

6.1 Abkürzungsverzeichnis

°C	Grad Celsius
µl	Mikroliter
µm	Mikrometer
µM	mikromolar
AA	Aminosäure, "amino acid"
ABC	"activated b-cell like"
ALL	akute lymphatische Leukämie
Anm.	Anmerkung
ASO-PCR	Allelspezifische Oligonukleotid-PCR
BCL-1	"B-cell lymphoma 1"
Bcl-2	"B-cell lymphoma 2"
BCR	B-Zell-Rezeptor, "B-cell receptor"
B-NHL	B-Non-Hodgkin-Lymphom
bp	Basenpaare
C	konstante Region des Immunglobulins, "constant region"
ca.	circa
CD	"cluster of differentiation"
CDR	"complementary determining region"
CLL	chronisch lymphatische Leukämie
CSF	„cerebrospinal fluid“
CSR	Klassenwechsel, "class switch recombination"
D	"diversity"-Gensegment des IGH-Rearrangements
ddNTP	Didesoxynukleosidtriphosphat
DLBCL	diffus großzelliges B-Zell-Lymphom, "diffuse large B-cell lymphoma"
DNA	Desoxyribonukleinsäure, "deoxyribonucleic acid"
dNTP	Desoxyribonukleosidtriphosphat
FC	Durchflusszytometrie, "flow cytometrie"

Verzeichnisse

FDC	follikuläre dendritische Zelle, "follicular dendritic cell"
FFPE	Formalin-fixiert Paraffin-eingebettet
FL	follikuläres Lymphom
FR	"framework region"
g	Gramm
GC	Keimzentrum, "germinal center"
GK	Glaskörper
h	Stunde
HIV	Humanes Immundefizienz-Virus
HTS	Hochdurchsatzsequenzierung, "high-throuput sequencing"
ID	Identifikator
Ig	Immunglobulin
IgH	schwere Kette des Immunglobulins, "immunoglobulin heavy chain"
IGH	rearrangierter DNA-Abschnitt, der für ein IgH-Molekül kodiert
IGK	rearrangierter DNA-Abschnitt, der für den κ -Isotyp der leichten Kette kodiert
IgL	leichte Kette des Immunglobulins, "immunoglobulin light chain"
IGL	rearrangierter DNA-Abschnitt, der für den λ -Isotyp der leichten Kette kodiert
IMGT	"international ImMunoGeneTics information system"
IOL	Intraokuläres Lymphom
J(H)	"joining"-Gensegment eines Antigenrezeptor-Gen-Rearrangements (H, bezogen auf die schwere Kette)
k.a.	kaum amplifiziert
KM	Knochenmark
LK	Lymphknoten
MCL	Mantelzelllymphom, "mantle cell lymphoma"
mg	Milligramm
min	Minute
ml	Milliliter
mM	millimolar
MM	Multiples Myelom

Verzeichnisse

MRD	"minimal residual disease"
ng	Nanogramm
NGS	"Next Generation Sequencing"
NHL	Non-Hodgkin-Lymphom
nmol	Nanomol
nt	Nukleotid
o.a.	oder andere
PBMC	"peripheral blood mononuclear cells"
PCNSL	"primary central nervous system lymphoma"
PCR	Polymerase-Kettenreaktion, "polymerase chain reaction"
PE	Probeexzision
pH	pH-Wert
PIOL	primär intraokuläres Lymphom
PTLD	"post transplant lymphoproliferative disease"
RAG	Rekombination-aktivierendes Gen
RFU	„relative fluorescence units“
RNA	Ribonukleinsäure, "ribonucleic acid"
rpm	rounds per minute
RSS	"recombination signal sequence"
s	Sekunde
SHM	somatische Hypermutation
sog.	sogenannt
SPRI	solide-phase reverse immobilization
TCR	T-Zell-Rezeptor, "T-cell receptor"
TdT	Deoxynukleotidyl-Transferase
U	unit
UV	ultraviolett
V	variable Region des Immunglobulins, "variable region"
V	Volt
V(H)	"variable"-Gensegment eines Antigenrezeptor-Gen- Rearrangements <u>oder</u> Abschnitt eines Antigenrezeptor-Gen-Rearrangements, der

für die V-Region des Antigenrezeptors kodiert (H: bezogen auf die schwere Kette)

WHO World Health Organisation

6.2 Literaturverzeichnis

1. Evans, L.S. and B.W. Hancock, *Non-Hodgkin lymphoma*. Lancet, 2003. **362**(9378): p. 139-46.
2. Shankland, K.R., J.O. Armitage, and B.W. Hancock, *Non-Hodgkin lymphoma*. Lancet, 2012. **380**(9844): p. 848-57.
3. *Krebs in Deutschland 2011/2012. 10. Ausgabe. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg). Berlin, 2015.*
4. Office for National Statistics. *Cancer statistics registrations: registrations of cancer diagnosed in 2007, England. Series MB1 no. 38. Surrey, UK: Office of Public Sector Information, 2010.*
5. Anderson, J.R., J.O. Armitage, and D.D. Weisenburger, *Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. Ann Oncol, 1998. 9(7): p. 717-20.*
6. Naresh, K.N., V. Srinivas, and C.S. Soman, *Distribution of various subtypes of non-Hodgkin's lymphoma in India: a study of 2773 lymphomas using R.E.A.L. and WHO Classifications. Ann Oncol, 2000. 11 Suppl 1: p. 63-7.*
7. Seifert, M., R. Scholtysik, and R. Kuppers, *Origin and pathogenesis of B cell lymphomas. Methods Mol Biol, 2013. 971: p. 1-25.*
8. Kuppers, R. and R. Dalla-Favera, *Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene, 2001. 20(40): p. 5580-94.*
9. Kuppers, R., et al., *Cellular origin of human B-cell lymphomas. N Engl J Med, 1999. 341(20): p. 1520-9.*
10. Swerdlow, S., et al., *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 2008: IARC Press.*
11. Swerdlow, S.H., et al., *The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood, 2016. 127(20): p. 2375-90.*
12. Campo, E., et al., *The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. Blood, 2011. 117(19): p. 5019-5032.*
13. Cheson, B.D., et al., *Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol, 2014. 32(27): p. 3059-68.*

14. Kho, M.E., et al., *Reliability of staging, prognosis, and comorbidity data collection in the National Comprehensive Cancer Network (NCCN) non-Hodgkin lymphoma (NHL) multicenter outcomes database*. *Cancer*, 2008. **113**(11): p. 3209-12.
15. Carbone, P.P., et al., *Report of the Committee on Hodgkin's Disease Staging Classification*. *Cancer Res*, 1971. **31**(11): p. 1860-1.
16. Rosenberg, S.A., *Validity of the Ann Arbor staging classification for the non-Hodgkin's lymphomas*. *Cancer Treat Rep*, 1977. **61**(6): p. 1023-7.
17. Tonegawa, S., *Somatic generation of antibody diversity*. *Nature*, 1983. **302**(5909): p. 575-81.
18. Chen, S., Z. Zuo, and D. Jones, *Molecular Diagnostics and Cytogenetic Testing*, in *Neoplastic Hematopathology: Experimental and Clinical Approaches*, D. Jones, Editor. 2010, Humana Press: Totowa, NJ. p. 61-95.
19. Urso, P. and T. Makinodan, *The Roles of Cellular Division and Maturation in the Formation of Precipitating Antibody*. *J Immunol*, 1963. **90**: p. 897-907.
20. Dreyer, W.J. and J.C. Bennett, *The molecular basis of antibody formation: a paradox*. *Proc Natl Acad Sci U S A*, 1965. **54**(3): p. 864-9.
21. Lewis, S.M., *The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses*. *Adv Immunol*, 1994. **56**: p. 27-150.
22. Edelman, G.M. and B. Benacerraf, *On structural and functional relations between antibodies and proteins of the gamma-system*. *Proc Natl Acad Sci U S A*, 1962. **48**: p. 1035-42.
23. Edelman, G.M. and J.A. Gally, *The nature of Bence-Jones proteins. Chemical similarities to polypeptide chains of myeloma globulins and normal gamma-globulins*. *J Exp Med*, 1962. **116**: p. 207-27.
24. Bennett, J.C., et al., *Evidence for Amino Acid Sequence Differences among Proteins Resembling the L-Chain Subunits of Immunoglobulins*. *J Mol Biol*, 1965. **12**: p. 81-7.
25. Schatz, D.G., *Antigen receptor genes and the evolution of a recombinase*. *Semin Immunol*, 2004. **16**(4): p. 245-56.
26. Chua, K.F., F.W. Alt, and J.P. Manis, *The function of AID in somatic mutation and class switch recombination: upstream or downstream of DNA breaks*. *J Exp Med*, 2002. **195**(9): p. F37-41.
27. Manis, J.P., M. Tian, and F.W. Alt, *Mechanism and control of class-switch recombination*. *Trends Immunol*, 2002. **23**(1): p. 31-9.
28. Jung, D., et al., *Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus*. *Annu Rev Immunol*, 2006. **24**: p. 541-70.
29. Flajnik, M.F., *Comparative analyses of immunoglobulin genes: surprises and portents*. *Nat Rev Immunol*, 2002. **2**(9): p. 688-98.
30. Storb, U., *The molecular basis of somatic hypermutation of immunoglobulin genes*. *Curr Opin Immunol*, 1996. **8**(2): p. 206-14.
31. Davis, M.M., *The evolutionary and structural 'logic' of antigen receptor diversity*. *Semin Immunol*, 2004. **16**(4): p. 239-43.

32. Ohno, S., N. Mori, and T. Matsunaga, *Antigen-binding specificities of antibodies are primarily determined by seven residues of VH*. Proc Natl Acad Sci U S A, 1985. **82**(9): p. 2945-9.
33. Arnaout, R., et al., *High-resolution description of antibody heavy-chain repertoires in humans*. PLoS One, 2011. **6**(8): p. e22365.
34. Fanning, L.J., A.M. Connor, and G.E. Wu, *Development of the immunoglobulin repertoire*. Clin Immunol Immunopathol, 1996. **79**(1): p. 1-14.
35. Lefranc, M.P., et al., *IMGT, the international ImMunoGeneTics information system*. Nucleic Acids Res, 2009. **37**(Database issue): p. D1006-12.
36. van Dongen, J.J., et al., *Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936*. Leukemia, 2003. **17**(12): p. 2257-317.
37. van Gent, D.C., et al., *Initiation of V(D)J recombinations in a cell-free system by RAG1 and RAG2 proteins*. Curr Top Microbiol Immunol, 1996. **217**: p. 1-10.
38. Oettinger, M.A., et al., *RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination*. Science, 1990. **248**(4962): p. 1517-23.
39. Akira, S., K. Okazaki, and H. Sakano, *Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining*. Science, 1987. **238**(4830): p. 1134-8.
40. van Gent, D.C., D.A. Ramsden, and M. Gellert, *The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination*. Cell, 1996. **85**(1): p. 107-13.
41. Kiyoi, H. and T. Naoe, *Immunoglobulin variable region structure and B-cell malignancies*. Int J Hematol, 2001. **73**(1): p. 47-53.
42. Gilfillan, S., et al., *Mice lacking TdT: mature animals with an immature lymphocyte repertoire*. Science, 1993. **261**(5125): p. 1175-8.
43. Komori, T. and H. Sugiyama, *N sequences, P nucleotides and short sequence homologies at junctional sites in VH to VHDJH and VHDJH to JH joining*. Mol Immunol, 1993. **30**(16): p. 1393-8.
44. Jacobs, H. and L. Bross, *Towards an understanding of somatic hypermutation*. Curr Opin Immunol, 2001. **13**(2): p. 208-18.
45. Hwang, J.K., F.W. Alt, and L.S. Yeap, *Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination*. Microbiol Spectr, 2015. **3**(1): p. MDNA3-0037-2014.
46. Eisen, H.N. and G.W. Siskind, *Variations in Affinities of Antibodies during the Immune Response*. Biochemistry, 1964. **3**: p. 996-1008.
47. Stevenson, F., et al., *Insight into the origin and clonal history of B-cell tumors as revealed by analysis of immunoglobulin variable region genes*. Immunol Rev, 1998. **162**: p. 247-59.
48. Bende, R.J., L.A. Smit, and C.J. van Noesel, *Molecular pathways in follicular lymphoma*. Leukemia, 2007. **21**(1): p. 18-29.
49. Pieper, K., B. Grimbacher, and H. Eibel, *B-cell biology and development*. Journal of Allergy and Clinical Immunology, 2013. **131**(4): p. 959-971.

50. Medina, K.L. and H. Singh, *Genetic networks that regulate B lymphopoiesis*. *Curr Opin Hematol*, 2005. **12**(3): p. 203-9.
51. Rajewsky, K., *Clonal selection and learning in the antibody system*. *Nature*, 1996. **381**(6585): p. 751-8.
52. Gorman, J.R. and F.W. Alt, *Regulation of immunoglobulin light chain isotype expression*. *Adv Immunol*, 1998. **69**: p. 113-81.
53. Schatz, D.G. and Y. Ji, *Recombination centres and the orchestration of V(D)J recombination*. *Nat Rev Immunol*, 2011. **11**(4): p. 251-63.
54. Tiegs, S.L., D.M. Russell, and D. Nemazee, *Receptor editing in self-reactive bone marrow B cells*. *J Exp Med*, 1993. **177**(4): p. 1009-20.
55. MacLennan, I.C., *Germinal centers*. *Annu Rev Immunol*, 1994. **12**: p. 117-39.
56. Jacob, J., et al., *Intraclonal generation of antibody mutants in germinal centres*. *Nature*, 1991. **354**(6352): p. 389-92.
57. Smith, A., et al., *Lymphoma incidence, survival and prevalence 2004-2014: sub-type analyses from the UK's Haematological Malignancy Research Network*. *Br J Cancer*, 2015. **112**(9): p. 1575-84.
58. Cleary, M.L. and J. Sklar, *Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome 18*. *Proc Natl Acad Sci U S A*, 1985. **82**(21): p. 7439-43.
59. Tsujimoto, Y., et al., *Involvement of the bcl-2 gene in human follicular lymphoma*. *Science*, 1985. **228**(4706): p. 1440-3.
60. Tsujimoto, Y., et al., *The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining*. *Science*, 1985. **229**(4720): p. 1390-3.
61. Kelly, R.S., et al., *Determinants of the t(14;18) translocation and their role in t(14;18)-positive follicular lymphoma*. *Cancer Causes Control*, 2015. **26**(12): p. 1845-55.
62. Zhu, D., et al., *Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma*. *Blood*, 2002. **99**(7): p. 2562-8.
63. Aarts, W.M., et al., *Variable heavy chain gene analysis of follicular lymphomas: correlation between heavy chain isotype expression and somatic mutation load*. *Blood*, 2000. **95**(9): p. 2922-9.
64. Zenz, T., et al., *From pathogenesis to treatment of chronic lymphocytic leukaemia*. *Nat Rev Cancer*, 2010. **10**(1): p. 37-50.
65. Damle, R.N., et al., *Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia*. *Blood*, 1999. **94**(6): p. 1840-7.
66. Hamblin, T.J., et al., *Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia*. *Blood*, 1999. **94**(6): p. 1848-54.
67. Seifert, M., et al., *Cellular origin and pathophysiology of chronic lymphocytic leukemia*. *J Exp Med*, 2012. **209**(12): p. 2183-98.
68. Rosenwald, A., et al., *The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma*. *N Engl J Med*, 2002. **346**(25): p. 1937-47.

69. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling*. Nature, 2000. **403**(6769): p. 503-11.
70. Lossos, I.S., et al., *Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas*. Proc Natl Acad Sci U S A, 2000. **97**(18): p. 10209-13.
71. Skarbnik, A.P. and A.H. Goy, *Mantle cell lymphoma: state of the art*. Clin Adv Hematol Oncol, 2015. **13**(1): p. 44-55.
72. Hadzidimitriou, A., et al., *Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases*. Blood, 2011. **118**(11): p. 3088-95.
73. Rosenberg, C.L., et al., *PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma*. Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9638-42.
74. Williams, M.E., et al., *Chromosome 11 translocation breakpoints at the PRAD1/cyclin D1 gene locus in centrocytic lymphoma*. Leukemia, 1993. **7**(2): p. 241-5.
75. Lovec, H., et al., *Cyclin D1/bcl-1 cooperates with myc genes in the generation of B-cell lymphoma in transgenic mice*. EMBO J, 1994. **13**(15): p. 3487-95.
76. Frater, J.L. and E.D. Hsi, *Properties of the mantle cell and mantle cell lymphoma*. Curr Opin Hematol, 2002. **9**(1): p. 56-62.
77. Kienle, D., et al., *VH mutation status and VDJ rearrangement structure in mantle cell lymphoma: correlation with genomic aberrations, clinical characteristics, and outcome*. Blood, 2003. **102**(8): p. 3003-9.
78. Orchard, J., et al., *A subset of t(11;14) lymphoma with mantle cell features displays mutated IgVH genes and includes patients with good prognosis, nonnodal disease*. Blood, 2003. **101**(12): p. 4975-81.
79. Camacho, F.I., et al., *Molecular heterogeneity in MCL defined by the use of specific VH genes and the frequency of somatic mutations*. Blood, 2003. **101**(10): p. 4042-6.
80. Walsh, S.H., et al., *Mutated VH genes and preferential VH3-21 use define new subsets of mantle cell lymphoma*. Blood, 2003. **101**(10): p. 4047-54.
81. Nowell, P.C., *The clonal evolution of tumor cell populations*. Science, 1976. **194**(4260): p. 23-8.
82. Sandberg, Y., et al., *BIOMED-2 multiplex immunoglobulin/T-cell receptor polymerase chain reaction protocols can reliably replace Southern blot analysis in routine clonality diagnostics*. J Mol Diagn, 2005. **7**(4): p. 495-503.
83. van Dongen, J.J. and I.L. Wolvers-Tettero, *Analysis of immunoglobulin and T cell receptor genes. Part I: Basic and technical aspects*. Clin Chim Acta, 1991. **198**(1-2): p. 1-91.
84. Boone, E., B. Verhaaf, and A.W. Langerak, *PCR-based analysis of rearranged immunoglobulin or T-cell receptor genes by GeneScan analysis or heteroduplex analysis for clonality assessment in lymphoma diagnostics*. Methods Mol Biol, 2013. **971**: p. 65-91.
85. Amara, K., et al., *PCR-based clonality analysis of B-cell lymphomas in paraffin-embedded tissues: diagnostic value of immunoglobulin kappa and*

- lambda light chain gene rearrangement investigation*. *Pathol Res Pract*, 2006. **202**(6): p. 425-31.
86. Berget, E., et al., *Detection of clonality in follicular lymphoma using formalin-fixed, paraffin-embedded tissue samples and BIOMED-2 immunoglobulin primers*. *J Clin Pathol*, 2011. **64**(1): p. 37-41.
 87. Kosari, F., et al., *PCR-based clonality analysis in diffuse large B-cell lymphoma using BIOMED-2 primers of IgH (FR3) on formalin-fixed paraffin-embedded tissue*. *Arch Iran Med*, 2013. **16**(9): p. 526-9.
 88. Liu, H., et al., *A practical strategy for the routine use of BIOMED-2 PCR assays for detection of B- and T-cell clonality in diagnostic haematopathology*. *Br J Haematol*, 2007. **138**(1): p. 31-43.
 89. Lukowsky, A., et al., *Evaluation of B-cell clonality in archival skin biopsy samples of cutaneous B-cell lymphoma by immunoglobulin heavy chain gene polymerase chain reaction*. *Leuk Lymphoma*, 2006. **47**(3): p. 487-93.
 90. McClure, R.F., et al., *Validation of immunoglobulin gene rearrangement detection by PCR using commercially available BIOMED-2 primers*. *Leukemia*, 2006. **20**(1): p. 176-9.
 91. Morales, A.V., et al., *Evaluation of B-cell clonality using the BIOMED-2 PCR method effectively distinguishes cutaneous B-cell lymphoma from benign lymphoid infiltrates*. *Am J Dermatopathol*, 2008. **30**(5): p. 425-30.
 92. van Krieken, J.H., et al., *Improved reliability of lymphoma diagnostics via PCR-based clonality testing: report of the BIOMED-2 Concerted Action BHM4-CT98-3936*. *Leukemia*, 2007. **21**(2): p. 201-6.
 93. Evans, P.A., et al., *Significantly improved PCR-based clonality testing in B-cell malignancies by use of multiple immunoglobulin gene targets. Report of the BIOMED-2 Concerted Action BHM4-CT98-3936*. *Leukemia*, 2007. **21**(2): p. 207-14.
 94. Langerak, A.W., et al., *Polymerase chain reaction-based clonality testing in tissue samples with reactive lymphoproliferations: usefulness and pitfalls. A report of the BIOMED-2 Concerted Action BMH4-CT98-3936*. *Leukemia*, 2007. **21**(2): p. 222-9.
 95. Langerak, A.W., et al., *EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations*. *Leukemia*, 2012. **26**(10): p. 2159-71.
 96. Hsu, F.J. and R. Levy, *Preferential use of the VH4 Ig gene family by diffuse large-cell lymphoma*. *Blood*, 1995. **86**(8): p. 3072-82.
 97. Lossos, I.S. and R. Levy, *Higher-grade transformation of follicle center lymphoma is associated with somatic mutation of the 5' noncoding regulatory region of the BCL-6 gene*. *Blood*, 2000. **96**(2): p. 635-9.
 98. Lossos, I.S., et al., *Molecular analysis of immunoglobulin genes in diffuse large B-cell lymphomas*. *Blood*, 2000. **95**(5): p. 1797-803.
 99. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. *Proc Natl Acad Sci U S A*, 1977. **74**(12): p. 5463-7.
 100. Mardis, E.R., *Next-generation DNA sequencing methods*. *Annu Rev Genomics Hum Genet*, 2008. **9**: p. 387-402.
 101. Zhou, X., et al., *The next-generation sequencing technology and application*. *Protein Cell*, 2010. **1**(6): p. 520-36.

102. Metzker, M.L., *Sequencing technologies - the next generation*. Nat Rev Genet, 2010. **11**(1): p. 31-46.
103. Merker, J.D., A. Valouev, and J. Gotlib, *Next-generation sequencing in hematologic malignancies: what will be the dividends?* Ther Adv Hematol, 2012. **3**(6): p. 333-9.
104. Amato, T., et al., *Clonality Analysis of Immunoglobulin Gene Rearrangement by Next-Generation Sequencing in Endemic Burkitt Lymphoma Suggests Antigen Drive Activation of BCR as Opposed to Sporadic Burkitt Lymphoma*. Am J Clin Pathol, 2016. **145**(1): p. 116-27.
105. Boyd, S.D., et al., *Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing*. Sci Transl Med, 2009. **1**(12): p. 12ra23.
106. Schumacher, J.A., et al., *A comparison of deep sequencing of TCRG rearrangements vs traditional capillary electrophoresis for assessment of clonality in T-Cell lymphoproliferative disorders*. Am J Clin Pathol, 2014. **141**(3): p. 348-59.
107. Sufficool, K.E., et al., *T-cell clonality assessment by next-generation sequencing improves detection sensitivity in mycosis fungoides*. J Am Acad Dermatol, 2015. **73**(2): p. 228-36 e2.
108. Wu, D., et al., *High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia*. Sci Transl Med, 2012. **4**(134): p. 134ra63.
109. Logan, A.C., et al., *Immunoglobulin and T cell receptor gene high-throughput sequencing quantifies minimal residual disease in acute lymphoblastic leukemia and predicts post-transplantation relapse and survival*. Biol Blood Marrow Transplant, 2014. **20**(9): p. 1307-13.
110. Logan, A.C., et al., *Minimal residual disease quantification using consensus primers and high-throughput IGH sequencing predicts post-transplant relapse in chronic lymphocytic leukemia*. Leukemia, 2013. **27**(8): p. 1659-65.
111. Sayala, H.A., A.C. Rawstron, and P. Hillmen, *Minimal residual disease assessment in chronic lymphocytic leukaemia*. Best Pract Res Clin Haematol, 2007. **20**(3): p. 499-512.
112. Scheijen, B., et al., *Next-generation sequencing of immunoglobulin gene rearrangements for clonality assessment: a technical feasibility study by EuroClonality-NGS*. Leukemia, 2019. **33**(9): p. 2227-2240.
113. McClure, R., M. Mai, and S. McClure, *High-throughput sequencing using the Ion Torrent personal genome machine for clinical evaluation of somatic hypermutation status in chronic lymphocytic leukemia*. J Mol Diagn, 2015. **17**(2): p. 145-54.
114. Alamyar, E., et al., *IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS*. Methods Mol Biol, 2012. **882**: p. 569-604.
115. Amin, H.M., et al., *Characterization of 4 mantle cell lymphoma cell lines*. Arch Pathol Lab Med, 2003. **127**(4): p. 424-31.

116. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling*. Nature, 2000. **403**(6769): p. 503-511.
117. Gilbert, M.T., et al., *The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when?* PLoS One, 2007. **2**(6): p. e537.
118. Bonin, S., et al., *PCR analysis in archival postmortem tissues*. Mol Pathol, 2003. **56**(3): p. 184-6.
119. Quach, N., M.F. Goodman, and D. Shibata, *In vitro mutation artifacts after formalin fixation and error prone translesion synthesis during PCR*. BMC Clin Pathol, 2004. **4**(1): p. 1.
120. DeAngelis, M.M., D.G. Wang, and T.L. Hawkins, *Solid-phase reversible immobilization for the isolation of PCR products*. Nucleic Acids Res, 1995. **23**(22): p. 4742-3.
121. Hawkins, T.L., et al., *DNA purification and isolation using a solid-phase*. Nucleic Acids Res, 1994. **22**(21): p. 4543-4.
122. Brochet, X., M.P. Lefranc, and V. Giudicelli, *IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis*. Nucleic Acids Res, 2008. **36**(Web Server issue): p. W503-8.
123. Liu, L., et al., *Comparison of next-generation sequencing systems*. J Biomed Biotechnol, 2012. **2012**: p. 251364.
124. Fedurco, M., et al., *BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies*. Nucleic Acids Res, 2006. **34**(3): p. e22.
125. Li, S., et al., *IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling*. Nat Commun, 2013. **4**: p. 2333.
126. Yousfi Monod, M., et al., *IMGT/JunctionAnalysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONS*. Bioinformatics, 2004. **20 Suppl 1**: p. i379-85.
127. Jiang, Y., et al., *Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas*. Genome Biol, 2014. **15**(8): p. 432.
128. Ladetto, M., et al., *Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders*. Leukemia, 2014. **28**(6): p. 1299-307.
129. Schirmer, M., et al., *Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform*. Nucleic Acids Res, 2015. **43**(6): p. e37.
130. Acinas, S.G., et al., *PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample*. Appl Environ Microbiol, 2005. **71**(12): p. 8966-9.
131. Park, J.W. and D.E. Crowley, *Nested PCR bias: a case study of Pseudomonas spp. in soil microcosms*. J Environ Monit, 2010. **12**(4): p. 985-8.
132. Kanagawa, T., *Bias and artifacts in multitemplate polymerase chain reactions (PCR)*. J Biosci Bioeng, 2003. **96**(4): p. 317-23.

133. Payne, K., et al., *BIOMED-2 PCR assays for IGK gene rearrangements are essential for B-cell clonality analysis in follicular lymphoma*. Br J Haematol, 2011. **155**(1): p. 84-92.
134. Williams, C., et al., *A high frequency of sequence alterations is due to formalin fixation of archival specimens*. Am J Pathol, 1999. **155**(5): p. 1467-71.
135. Halldorsdottir, A.M., B.A. Zehnbauer, and W.R. Burack, *Application of BIOMED-2 clonality assays to formalin-fixed paraffin embedded follicular lymphoma specimens: superior performance of the IGK assays compared to IGH for suboptimal specimens*. Leuk Lymphoma, 2007. **48**(7): p. 1338-43.
136. Catherwood, M.A., et al., *Improved clonality assessment in germinal centre/post-germinal centre non-Hodgkin's lymphomas with high rates of somatic hypermutation*. J Clin Pathol, 2007. **60**(5): p. 524-8.
137. Bystry, V., et al., *ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data*. Bioinformatics, 2017. **33**(3): p. 435-437.
138. Logan, A.C., et al., *High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment*. Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21194-9.
139. Faham, M., et al., *Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia*. Blood, 2012. **120**(26): p. 5173-80.
140. Jadayel, D.M., et al., *Potential role for concurrent abnormalities of the cyclin D1, p16CDKN2 and p15CDKN2B genes in certain B cell non-Hodgkin's lymphomas. Functional studies in a cell line (Granta 519)*. Leukemia, 1997. **11**(1): p. 64-72.
141. Rudolph, C., et al., *Molecular cytogenetic characterization of the mantle cell lymphoma cell line GRANTA-519*. Cancer Genet Cytogenet, 2004. **153**(2): p. 144-50.
142. Ek, S., E. Ortega, and C.A. Borrebaeck, *Transcriptional profiling and assessment of cell lines as in vitro models for mantle cell lymphoma*. Leuk Res, 2005. **29**(2): p. 205-13.
143. Kleinfield, R., et al., *Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1+ B-cell lymphoma*. Nature, 1986. **322**(6082): p. 843-6.
144. Reth, M., et al., *A novel VH to VHDJH joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production*. Nature, 1986. **322**(6082): p. 840-2.
145. Itoh, K., et al., *Immunoglobulin heavy chain variable region gene replacement As a mechanism for receptor revision in rheumatoid arthritis synovial tissue B lymphocytes*. J Exp Med, 2000. **192**(8): p. 1151-64.
146. Wilson, P.C., et al., *Receptor revision of immunoglobulin heavy chain variable region genes in normal human B lymphocytes*. J Exp Med, 2000. **191**(11): p. 1881-94.
147. Lenze, D., et al., *Receptor revision of immunoglobulin heavy chain genes in human MALT lymphomas*. Mol Pathol, 2003. **56**(5): p. 249-55.

148. Sprangers, M., et al., *SLP65 deficiency results in perpetual V(D)J recombinase activity in pre-B-lymphoblastic leukemia and B-cell lymphoma cells*. *Oncogene*, 2006. **25**(37): p. 5180-6.
149. Lai, R., et al., *Immunoglobulin VH somatic hypermutation in mantle cell lymphoma: mutated genotype correlates with better clinical outcome*. *Mod Pathol*, 2006. **19**(11): p. 1498-505.
150. Jelinek, D.F., et al., *Analysis of clonal B-cell CD38 and immunoglobulin variable region sequence status in relation to clinical outcome for B-chronic lymphocytic leukaemia*. *Br J Haematol*, 2001. **115**(4): p. 854-61.
151. Shanafelt, T.D., et al., *Age at diagnosis and the utility of prognostic testing in patients with chronic lymphocytic leukemia*. *Cancer*, 2010. **116**(20): p. 4777-87.
152. Ghia, P., et al., *ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia*. *Leukemia*, 2007. **21**(1): p. 1-3.
153. Stamatopoulos, B., et al., *Targeted deep sequencing reveals clinically relevant subclonal IgHV rearrangements in chronic lymphocytic leukemia*. *Leukemia*, 2017. **31**(4): p. 837-845.
154. Davi, F., et al., *Immunoglobulin gene analysis in chronic lymphocytic leukemia in the era of next generation sequencing*. *Leukemia*, 2020. **34**(10): p. 2545-2551.
155. Gupta, S.K., D.S. Viswanatha, and K.P. Patel, *Evaluation of Somatic Hypermutation Status in Chronic Lymphocytic Leukemia (CLL) in the Era of Next Generation Sequencing*. *Front Cell Dev Biol*, 2020. **8**: p. 357.
156. Hofreiter, M., et al., *DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA*. *Nucleic Acids Res*, 2001. **29**(23): p. 4793-9.
157. Rait, V.K., et al., *Conversions of formaldehyde-modified 2'-deoxyadenosine 5'-monophosphate in conditions modeling formalin-fixed tissue dehydration*. *J Histochem Cytochem*, 2006. **54**(3): p. 301-10.
158. Kapp, J.R., et al., *Variation in pre-PCR processing of FFPE samples leads to discrepancies in BRAF and EGFR mutation detection: a diagnostic RING trial*. *J Clin Pathol*, 2015. **68**(2): p. 111-8.
159. Bourgon, R., et al., *High-Throughput Detection of Clinically Relevant Mutations in Archived Tumor Samples by Multiplexed PCR and Next-Generation Sequencing*. *Clinical Cancer Research*, 2014. **20**(8): p. 2080-2091.
160. Do, H., et al., *Reducing Sequence Artifacts in Amplicon-Based Massively Parallel Sequencing of Formalin-Fixed Paraffin-Embedded DNA by Enzymatic Depletion of Uracil-Containing Templates*. *Clinical Chemistry*, 2013. **59**(9): p. 1376-1383.
161. Halbritter, J., et al., *High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing*. *J Med Genet*, 2012. **49**(12): p. 756-67.
162. Do, H. and A. Dobrovic, *Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil-DNA glycosylase*. *Oncotarget*, 2012. **3**(5): p. 546-58.

163. Wang, M., et al., *Somatic Mutation Screening Using Archival Formalin-Fixed, Paraffin-Embedded Tissues by Fluidigm Multiplex PCR and Illumina Sequencing*. J Mol Diagn, 2015. **17**(5): p. 521-32.
164. Blais, J., et al., *Risk of Misdiagnosis Due to Allele Dropout and False-Positive PCR Artifacts in Molecular Diagnostics: Analysis of 30,769 Genotypes*. J Mol Diagn, 2015. **17**(5): p. 505-14.
165. Bracho, M.A., A. Moya, and E. Barrio, *Contribution of Taq polymerase-induced errors to the estimation of RNA virus diversity*. Journal of General Virology, 1998. **79**(12): p. 2921-2928.
166. Cline, J., J.C. Braman, and H.H. Hogrefe, *PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases*. Nucleic Acids Res, 1996. **24**(18): p. 3546-51.
167. Qiu, X., et al., *Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning*. Appl Environ Microbiol, 2001. **67**(2): p. 880-7.
168. Kircher, M., U. Stenzel, and J. Kelso, *Improved base calling for the Illumina Genome Analyzer using machine learning strategies*. Genome Biol, 2009. **10**(8): p. R83.
169. Ross, M.G., et al., *Characterizing and measuring bias in sequence data*. Genome Biol, 2013. **14**(5): p. R51.
170. Knecht, H., et al., *Quality control and quantification in IG/TR next-generation sequencing marker identification: protocols and bioinformatic functionalities by EuroClonality-NGS*. Leukemia, 2019. **33**(9): p. 2254-2265.
171. Nguyen, P., et al., *Identification of errors introduced during high throughput sequencing of the T cell receptor repertoire*. BMC Genomics, 2011. **12**: p. 106.
172. Campbell, P.J., et al., *Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing*. Proc Natl Acad Sci U S A, 2008. **105**(35): p. 13081-6.
173. Gawad, C., et al., *Massive evolution of the immunoglobulin heavy chain locus in children with B precursor acute lymphoblastic leukemia*. Blood, 2012. **120**(22): p. 4407-17.
174. Carlotti, E., et al., *High Throughput Sequencing Analysis of the Immunoglobulin Heavy Chain Gene from Flow-Sorted B Cell Sub-Populations Define the Dynamics of Follicular Lymphoma Clonal Evolution*. PLoS One, 2015. **10**(9): p. e0134833.
175. Jiang, Y., et al., *VDJ-Seq: Deep Sequencing Analysis of Rearranged Immunoglobulin Heavy Chain Gene to Reveal Clonal Evolution Patterns of B Cell Lymphoma*. J Vis Exp, 2015(106): p. e53215.
176. Hwang, S.-A. and J.K. Actor, *Lymphocytes*, in eLS. 2001, John Wiley & Sons, Ltd.
177. Bottcher, S., et al., *Comparative analysis of minimal residual disease detection using four-color flow cytometry, consensus IgH-PCR, and quantitative IgH PCR in CLL after allogeneic and autologous stem cell transplantation*. Leukemia, 2004. **18**(10): p. 1637-45.
178. Bruggemann, M., et al., *Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker*

- identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study.* Leukemia, 2019. **33**(9): p. 2241-2253.
179. Yao, Q., et al., *Standardized Minimal Residual Disease Detection by Next-Generation Sequencing in Multiple Myeloma.* Front Oncol, 2019. **9**: p. 449.
 180. Rustad, E.H., et al., *Baseline identification of clonal V(D)J sequences for DNA-based minimal residual disease detection in multiple myeloma.* PLoS One, 2019. **14**(3): p. e0211600.
 181. Coupland, S.E., H. Heimann, and N.E. Bechrakis, *Primary intraocular lymphoma: a review of the clinical, histopathological and molecular biological features.* Graefes Arch Clin Exp Ophthalmol, 2004. **242**(11): p. 901-13.
 182. Sen, H.N., et al., *Primary intraocular lymphoma: diagnosis and differential diagnosis.* Ocul Immunol Inflamm, 2009. **17**(3): p. 133-41.
 183. Jahnke, K., et al., *Diagnosis and management of primary intraocular lymphoma: an update.* Clin Ophthalmol, 2007. **1**(3): p. 247-58.
 184. Elenitoba-Johnson, K.S., et al., *PCR analysis of the immunoglobulin heavy chain gene in polyclonal processes can yield pseudoclonal bands as an artifact of low B cell number.* J Mol Diagn, 2000. **2**(2): p. 92-6.
 185. Wan, J.H., et al., *Rapid method for detecting monoclonality in B cell lymphoma in lymph node aspirates using the polymerase chain reaction.* J Clin Pathol, 1992. **45**(5): p. 420-3.
 186. Grange, L.K., et al., *Neoplastic masquerade syndromes in patients with uveitis.* Am J Ophthalmol, 2014. **157**(3): p. 526-31.
 187. Lobo, A., et al., *Pseudo-hypopyon as the presenting feature in B-cell and T-cell intraocular lymphoma.* Clin Exp Ophthalmol, 2003. **31**(2): p. 155-8.
 188. O'Keefe, J.S., et al., *Anterior chamber infiltrates associated with systemic lymphoma: report of two cases and review of the literature.* Ophthalmology, 2002. **109**(2): p. 253-7.
 189. Read, R.W., E. Zamir, and N.A. Rao, *Neoplastic masquerade syndromes.* Survey of Ophthalmology, 2002. **47**(2): p. 81-124.
 190. Baehring, J.M., et al., *Analysis of clonal immunoglobulin heavy chain rearrangements in ocular lymphoma.* Cancer, 2005. **104**(3): p. 591-7.
 191. Feilmeier, M.R., et al., *Successful use of semi-nested PCR for the diagnosis of primary intraocular lymphoma.* Leuk Lymphoma, 2008. **49**(8): p. 1612-5.
 192. Sugita, S., et al., *Diagnosis of intraocular lymphoma by polymerase chain reaction analysis and cytokine profiling of the vitreous fluid.* Jpn J Ophthalmol, 2009. **53**(3): p. 209-14.
 193. Kimura, K., et al., *Clinical features and diagnostic significance of the intraocular fluid of 217 patients with intraocular lymphoma.* Jpn J Ophthalmol, 2012. **56**(4): p. 383-9.
 194. Whitcup, S.M., et al., *Association of interleukin 10 in the vitreous and cerebrospinal fluid and primary central nervous system lymphoma.* Arch Ophthalmol, 1997. **115**(9): p. 1157-60.
 195. Bruno, A., et al., *Mutational analysis of primary central nervous system lymphoma.* Oncotarget, 2014. **5**(13): p. 5065-75.

196. Bonzheim, I., et al., *High frequency of MYD88 mutations in vitreoretinal B-cell lymphoma: a valuable tool to improve diagnostic yield of vitreous aspirates*. Blood, 2015. **126**(1): p. 76-9.
197. Glass, J.P., et al., *Malignant cells in cerebrospinal fluid (CSF): the meaning of a positive CSF cytology*. Neurology, 1979. **29**(10): p. 1369-75.
198. Craig, F.E., et al., *Flow cytometric immunophenotyping of cerebrospinal fluid specimens*. Am J Clin Pathol, 2011. **135**(1): p. 22-34.
199. Fischer, L., et al., *Meningeal dissemination in primary CNS lymphoma: prospective evaluation of 282 patients*. Neurology, 2008. **71**(14): p. 1102-8.
200. Baehring, J.M., et al., *Immunoglobulin gene rearrangement analysis in cerebrospinal fluid of patients with lymphoproliferative processes*. J Neurol Sci, 2006. **247**(2): p. 208-16.
201. Ekstein, D., et al., *CSF analysis of IgH gene rearrangement in CNS lymphoma: relationship to the disease course*. J Neurol Sci, 2006. **247**(1): p. 39-46.
202. Liu, L., et al., *Detection of malignant B lymphocytes by PCR clonality assay using direct lysis of cerebrospinal fluid and low volume specimens*. Int J Lab Hematol, 2015. **37**(2): p. 165-73.

6.3 Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung eines IGH-Rearrangements mit „framework regions“ (FRs) und „complementary determining regions“	7
Abbildung 2: VDJ-Rekombination des IGH-Locus.....	8
Abbildung 3: Ursprung ausgewählter B-NHL in Relation zur Keimzentrumsreaktion. ...	14
Abbildung 4: Klonalitätsdiagnostik mittels IGH-Amplifikation und Fragmentanalyse.	17
Abbildung 5: Repräsentatives Elektropherogramm einer pseudoklonalen bzw. oligoklonalen Rearrangement-Verteilung.	19
Abbildung 6: Schematische Darstellung der PCR-Schritte und Sequenzierung.....	22
Abbildung 7: Beispiel der Analyse der IGH-Sequenz eines DLBCL mittels IMGT/V-Quest, Ergebniszusammenfassung.....	41
Abbildung 8: Ausschnitt aus der Ausgabedatei „IMGT clonotypes (AA) per Nb“ der statistischen Analyse.	48
Abbildung 9: Exemplarische Darstellung der gelelektrophoretischen Auftrennung der Produkte der PCR zur Überprüfung der DNA-Qualität von Patientenproben mit hohem Lymphomzellgehalt.....	53
Abbildung 10: Prozentualer Anteil der aus FFPE-Gewebe isolierten DNA-Proben mit jeweiliger maximaler Fragmentlänge, die sich mittels Qualitäts-PCR amplifizieren ließ.	54

Abbildung 11: Gelelektrophoretische Darstellung der IGH-Amplifikation der Patientenproben mit hohem Lymphomzellgehalt (S13 bis S27) unter Verwendung der DNA-Polymerase AmpliTaq Gold, Primerset FR1.....	55
Abbildung 12: NGS-Analyse der IGH-Rearrangements der Zelllinie Granta-519 in verschiedenen Verdünnungsstufen.....	60
Abbildung 13: NGS-Analyse der IGH-Rearrangements der Zelllinie OCI-LY3 in verschiedenen Verdünnungsstufen.....	62
Abbildung 14: Verteilung der IGH-Sequenzen auf verschiedene Clonotypes (AA) in der Probe Granta-519, Verdünnung 1:5, Primerset FR1.	66
Abbildung 15: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe Granta-519, Verdünnung 1:5, Primerset FR1.....	67
Abbildung 16: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in der Probe Granta-519, Verdünnung 1:5, Primerset FR1.	68
Abbildung 17: Verteilung der IGH-Sequenzen auf verschiedene Clonotypes (AA) in der Probe Granta-519, Verdünnung 1:1.000, Primerset FR2.	71
Abbildung 18: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe Granta-519, Verdünnung 1:1.000, Primerset FR2.	72
Abbildung 19: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in der Probe Granta-519, Verdünnung 1:1000, Primerset FR2.	73
Abbildung 20: Verteilung der analysierten IGH-Sequenzen nach ermittelter Sequenzlänge in den Probe OCI-LY3, Verdünnung 1:5 sowie Verdünnung 1:1.000, Primerset FR2.....	76
Abbildung 21: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe OCI-LY3, Verdünnung 1:5, Primerset FR2.	77
Abbildung 22: Verteilung der analysierten IGH-Sequenzen nach ermitteltem V-Gen und –Allel in der Probe OCI-LY3, Verdünnung 1:1.000, Primerset FR2.	78
Abbildung 23: Probe S17, CLL; Anzahl zugeordneter Sequenzen pro Clonotype (AA).	87
Abbildung 24: Probe S18, FL, Datensatz FR2; Anzahl zugeordneter Sequenzen pro Sequenzlänge und pro Clonotype (AA).....	88
Abbildung 25: Probe S19, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA)..	88
Abbildung 26: ProbeS20, FL; Anzahl zugeordneter Sequenzen pro Sequenzlänge, Datensatz FR1, sowie Anzahl zugeordneter Sequenzen pro Clonotype, Datensatz FR2.	89
Abbildung 27: Probe S21, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) und pro Sequenzlänge.....	90
Abbildung 28: Probe S22, FL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) für die Datensätze FR1 und FR2.....	91
Abbildung 29: Probe S24, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Clonotype (AA) sowie pro Sequenzlänge, Datensatz FR2.	92
Abbildung 30: Probe S26, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Sequenzlänge sowie pro Clonotype (AA), Datensatz FR2.	92

Abbildung 31: Probe S27, ABC DLBCL; Anzahl zugeordneter Sequenzen pro Clonotype (AA). 93

Abbildung 32: Glaskörperprobe S46, Anzahl zugeordneter Sequenzen pro Clonotype (AA). 117

6.4 Tabellenverzeichnis

Tabelle 1: PCR zur Bestimmung der Amplifizierbarkeit der DNA. 31

Tabelle 2: PCR zur Amplifikation klonaler IGH-Rearrangements. 33

Tabelle 3: Standardisierte EuroClonality-Terminologie zur technischen Beschreibung der PCR-Ergebnisse. 34

Tabelle 4: PCR zur Kettenabbruchsynthese. 38

Tabelle 5: PCR zur IGH-Amplifikation in Zelllinienverdünnungs- und Patientenproben. 42

Tabelle 6: Barcode-PCR. 43

Tabelle 7: Übersicht über Verdünnungsproben, die aus B-NHL-Zelllinien-DNA durch Verdünnung in PBMC-DNA hergestellt wurden. 50

Tabelle 8: Übersicht über die in dieser Studie untersuchten Patientenproben. 52

Tabelle 9: Charakterisierung des IGH-Rearrangements der Zelllinie Granta-519 mit dem FR1-Primerset durch IMGT/V-Quest. 57

Tabelle 10: Charakterisierung des IGH-Rearrangements der Zelllinie OCI-LY3 mit dem FR1-Primerset durch IMGT/V-Quest. 58

Tabelle 11: Darstellung der zehn Clonotypes (AA) mit den meisten zugeordneten Sequenzen im Ansatz der Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1 nach IMGT/HighV-Quest-Analyse. 65

Tabelle 12: Darstellung der zehn Clonotypes (AA) mit den meisten zugeordneten Sequenzen im Ansatz der Zelllinie Granta-519, Verdünnung 1:1.000, Primerset FR2. 69

Tabelle 13: Übersicht der NGS-Ergebnisse der Verdünnungsproben Zelllinie Granta-519, Primersets FR1 und FR2 nach IMGT/HighV-Quest-Auswertung; 74

Tabelle 14: Sequenzierbarkeit nach Sanger von Patientenproben mit hohem Lymphomzellgehalt. 80

Tabelle 15: Charakterisierung durch IMGT/V-Quest der durch Sanger-Sequenzierung ermittelten klonalen IGH-Sequenzen aus Patientenproben mit hohem Lymphomzellgehalt. 81

Tabelle 16: Anzahl mittels NGS gewonnener und einem IMGT-Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT-Clonotypes (AA) in Patientenproben mit hohem Lymphomzellgehalt nach IMGT/HighV-Quest-Analyse. .. 82

Tabelle 17: Übersicht der NGS-Ergebnisse der Patientenproben mit hohem Lymphomzellgehalt bei denen das klonale Rearrangement zuvor mittels Sanger-Sequenzierung bestimmt werden konnte 84

Tabelle 18: V-Region Identity untersuchter CLL-Fälle mit hohem Lymphomzellgehalt in Sanger-Sequenzierung und Clonotypes (AA) auf Rang 1 der NGS-Datensätze.....	86
Tabelle 19: Überblick über die Ergebnisse von GeneScan-Analyse und Klonalitätsbefund in NGS-Datensätzen sowie die Sequenzierbarkeit nach Sanger in Patientenproben mit hohem Lymphomzellgehalt.	94
Tabelle 20: Anzahl mittels NGS gewonnener und einem IMGT Clonotype (AA) zugeordneter Sequenzen sowie Anzahl verschiedener IMGT Clonotypes (AA) in Verlaufsproben.	96
Tabelle 21: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 16 nach IMGT/HighV-Quest-Auswertung; Anzahl und Anteile der IGH-Sequenzen, die den sequenzstärksten Rearrangement-Populationen bzw. der Hintergrundpopulation zugeordnet wurden.	99
Tabelle 22: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 18 nach IMGT/HighV-Quest-Auswertung;	101
Tabelle 23: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 19 nach IMGT/HighV-Quest-Auswertung;	102
Tabelle 24: Übersicht der NGS-Ergebnisse der Verlaufsproben des Patienten 20 nach IMGT/HighV-Quest-Auswertung;	103
Tabelle 25: Patient 18, CLL, Primerset FR1. Verschiedene Varianten in der CDR3 des klonalen IGH-Rearrangements und Anzahlen zugeordneter Sequenzen im zeitlichen Verlauf.	106
Tabelle 26: Patient 18, CLL, Probe 1, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) der klonalen Rearrangement-Population auf Nukleotidebene. V-Region Identity des Clonotypes (AA) 100 %.....	106
Tabelle 27: Patient 19, FL. Primerset FR1. Verschiedene Varianten in der CDR3 des klonalen IGH-Rearrangements und Anzahlen zugeordneter Sequenzen im zeitlichen Verlauf.	109
Tabelle 28: Patient 19, FL, Probe 1, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) der klonalen Rearrangement-Population auf Nukleotidebene. V-Region Identity des Clonotypes (AA) ca. 90 %.....	109
Tabelle 29: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1, Sequenzstarke Clonotypes (AA) der Lymphomzelllinie.	112
Tabelle 30: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR2, Sequenzstarke Clonotypes (AA) der Lymphomzelllinie.	113
Tabelle 31: Zelllinie Granta-519, Verdünnung 1:5, Primerset FR1, CDR3 des sequenzstärksten Clonotypes (AA) des Lymphom-Rearrangements auf Nukleotidebene. V-Region-Identity des Clonotypes (AA) 100 %.....	114
Tabelle 32: Anzahl mittels NGS gewonnener und einem IMGT Clonotype (AA) zugeordneter Sequenzen, sowie Anzahl verschiedener IMGT Clonotypes (AA) in Glaskörperproben mit geringem Lymphomzellgehalt.	116

7. Erklärung zum Eigenanteil der Dissertationsschrift

Die Arbeit wurde im Department für Pathologie und Neuropathologie unter Betreuung von Herrn Prof. Dr. Falko Fend und Frau Dr. Irina Bonzheim durchgeführt.

Die Konzeption der Studie erfolgte in Zusammenarbeit mit Prof. Dr. Falko Fend und Frau Dr. Irina Bonzheim.

Sämtliche Versuche wurden (nach Einarbeitung durch die Labormitglieder Frau Sieglinde Baisch, Frau Sema Colak, Frau Julia Baturin, Frau Dr. Julia Steinhilber und Frau Dr. Janine Schmidt) von mir durchgeführt.

Der Sequenzierungsprozess nach Erstellung der DNA-Pools auf der Illumina MiSeq-Plattform wurde durch Herrn Dr. Martin Schulze und seinen beteiligten Mitarbeitern am Institut für Medizinische Genetik und angewandte Genomik durchgeführt.

Die angeschlossene bioinformatische Auswertung erfolgte eigenständig durch mich.

Ich versichere, das Manuskript selbständig nach Anleitung von Frau Dr. Irina Bonzheim verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.



Edingen-Neckarhausen, 08.01.2021

8. Danksagung

An erster Stelle gilt mein Dank Herrn Prof. Dr. Falko Fend: Zum einen danke ich ihm für die Überlassung des interessanten Themas, das zu Beginn spannend und herausfordernd klang, und sich zu jedem Zeitpunkt als derart erwies; andererseits dafür, dass ich am Institut für Pathologie exzellente Bedingungen für die Anfertigung der vorliegenden Arbeit vorfinden durfte. Prof. Fend, stand mir jeder Zeit mit klaren Vorstellungen und konkreten Vorschlägen zur Seite und stets kurzfristig für Fragen zur Verfügung.

Frau Dr. Irina Bonzheim danke ich für die herausragende Betreuung. Über die gesamte Entstehungszeit dieser Arbeit hinweg war sie stets anwesend oder schnell und unkompliziert erreichbar. Mit kompetentem Rat, mit Anregungen und Kritik trug sie maßgeblich zum Gelingen dieser Arbeit bei. In für mich vorbildhafter Weise jonglierte sie Routine- und wissenschaftliche Arbeiten neben großen privaten Projekten und vermittelte mir dennoch bei der Beantwortung meiner Fragen den Eindruck ungeteilter Aufmerksamkeit.

Frau Dr. Julia Steinhilber und Frau Dr. Janine Schmidt möchte ich für die vielen Male danken, da sie mir Lösungen für theoretische, und Hilfestellungen in praktischen Fragen gaben. Auch Frau Sieglinde Baisch, Frau Sema Colak und Frau Julia Baturin danke ich für die Hilfsbereitschaft und Geduld, die sie mir bei meinen Laborarbeiten entgegenbrachten. Ich danke auch Herrn Dr. Martin Schulze und seinen beteiligten Mitarbeitern am Institut für Medizinische Genetik und angewandte Genomik für die Sequenzierung der Proben.

Ebenfalls bin ich dankbar für die Förderung durch das IZKF-Promotionskolleg. Namentlich möchte ich Frau Prof. Marlies Knipper, Frau Dr. Inka Montero sowie Frau Pia-Sarah Henger für die Organisation der lehrreichen Seminare und die offene Betreuung danken.

Schließlich möchte ich meiner Familie für die Unterstützung danken, im Besonderen meinen Eltern und meiner Verlobten.