

MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG
INSTITUT FÜR INFORMATIK



**Morphologische und transkriptomische Analyse neuronaler Zellen
basierend auf Verfahren der
Bildverarbeitung und Sequenzanalyse.**

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

vorgelegt der

**Naturwissenschaftlichen Fakultät III
Agrar. und Ernährungswissenschaften,
Geowissenschaften und Informatik**

von

Herrn Danny Misiak

geboren am 11. April 1983 in Aschersleben

Gutachter:
Stefan Posch
Stefan Hüttelmaier
Johannes Schulte

Verteidigt am 27.06.2019

Halle (Saale), 2018

Danksagung

Einen besonderer Dank gebührt an dieser Stelle meiner Frau Claudia für ihre Geduld, das Vertrauen und die moralische Unterstützung auf dem gesamten Weg bei der Erstellung dieser Arbeit und darüber hinaus.

Ich möchte mich zudem bei Prof. Dr. Stefan Hüttelmaier und Prof. Dr. Stefan Posch für die Betreuung und die Möglichkeit dieses Forschungsthema zu bearbeiten bedanken. Weiterhin danke ich beiden sehr für ihre konstruktive Kritik und die nötige Geduld die ich ihnen bis zur Fertigstellung der nun vorliegenden Arbeit abverlangt habe. Auch danke ich der gesamten AG Hüttelmaier für die besonders angenehme Arbeitsatmosphäre und die vielen Diskussionen und Gespräche, die ich auch über fachliche Belange hinaus zu schätzen gelernt habe.

Frau Dr. Birgit Möller danke ich für ihre wertvolle Unterstützung in den vergangenen Jahren, die zum Gelingen dieser Arbeit beigetragen hat und immer Zeit für Gespräche ermöglichte.

Nicht zuletzt möchte ich meinen Eltern, meiner Familie und meinen Freunden einen großen Dank für den immerwährenden Rückhalt im Verlauf meines Studiums und bei Fertigstellung dieser Arbeit zukommen lassen.

Für Jonathan

Inhaltsverzeichnis

1. Einleitung	1
1.1. Morphologie neuronaler Zellen	1
1.1.1. Neuronale Strukturen	2
1.1.2. Lokale Translation in Neuronen	3
1.1.3. Analyse neuronaler Zellmorphologie	5
1.2. Krebserkrankung Neuroblastom	7
1.2.1. Klinische Einteilung	8
1.2.2. Therapie und Prognose	11
1.3. Untersuchung des Transkriptom	14
1.3.1. Hochdurchsatz-RNA-Sequenzierung	14
1.4. Zielstellung der Arbeit	17
2. Material und Methoden	21
2.1. Fluoreszenzbasierte Analyse neuronaler Zellmorphologie	21
2.1.1. Fluoreszenzmikroskopiebilder neuronaler Zellen	22
2.1.2. Detektion neuronaler Zellen	24
2.1.3. Identifizierung neuronaler Strukturen	33
2.1.4. Extraktion von Proteinverteilungen	42
2.2. Transkriptomanalyse	45
2.2.1. RNA-Sequenzierungsdaten	45
2.2.2. RNA-Sequenzierungsanalyse	48
2.2.3. Differentielle Expression	55
2.2.4. Identifizierung von Gensignaturen	60
2.2.5. microRNA-Targetanalyse	66
3. Ergebnisse	83
3.1. Neuronale Zellmorphologie und Proteinverteilung	83
3.1.1. Automatische Detektion der Zellmorphologie	84
3.1.2. Proteinverteilung in neuronalen Zellen	95
3.2. Transkriptomanalyse des Neuroblastoms	98
3.2.1. Sequenzierung der Neuroblastomproben	98
3.2.2. Differentielle Expression im Neuroblastom	102
3.2.3. Gensignaturen des Neuroblastoms	113
3.2.4. microRNA-Targets im Neuroblastom	134
4. Diskussion	177
5. Zusammenfassung und Ausblick	191
Literatur	199

Anhang	213
A. Neuronenmorphologie	213
B. Probeninformationen zu den Neuroblastomdaten	214
C. Differentielle Genexpression	216
D. Ermittelte Gensignaturen	216
D.1. Signaturgene	216
D.2. Analyse der Signaturen	225
E. Öffentliche Daten	225
F. microRNA-Targets	225
F.1. Informationen zu den microRNA-Bindestellen	225
F.2. Tabellen der microRNA-Targetanalyse	225
F.3. Bestenliste der microRNA-Targetanalyse	226

1. Einleitung

1.1. Morphologie neuronaler Zellen

Morphologische Veränderungen in der Zelle können durch Umstrukturierungen des Zytoskeletts bedingt werden, welches aus drei Hauptkomponenten, den Mikrotubuli, Mikrofilamenten (Aktin) sowie intermediären Filamenten besteht. Während der Entwicklung der Zelle ist der Prozess der zellulären Morphogenese von großer Bedeutung. Dieser Prozess beschreibt die Fähigkeit von Zellen, ihre Form zu verändern. Die dabei ablaufenden morphologischen Variationen spielen bei der Entwicklung des Gehirns, den Nerven und des Nervensystems eine wichtige Rolle und können Ursache verschiedener neurologischer Entwicklungsstörungen sein. Neuronale Zellen weisen dabei eine der komplexesten Zellmorphogenesen im sich entwickelnden Organismus auf [66]. Die Entwicklung eines Neurons beginnt aus einer Zelle mit einfacher Sphäroidform, deren anschließende Entwicklungsphase eine vollständige Transformation der Zellform mit sich bringt und in der diverse intrazelluläre Prozesse gleichzeitig stattfinden. Die Neuritogenese, die Initiation von Neuriten, ist das wichtigste Ereignis der neuronalen Morphogenese, um verschiedene Stadien der neuronalen Entwicklung zu erreichen und funktionsfähig in das Nervensystem integriert zu werden. Aus den zunächst entstehenden Neuriten werden im Laufe der Entwicklung Axone und Dendriten der reifen Neuronen gebildet und stellen die Basis der komplizierten Verschaltung des gesamten Nervensystems dar. In den letzten Jahren wurden erhebliche Fortschritte beim Verständnis der einzelnen Entwicklungsphasen als auch des Zusammenspiels der diversen simultan ablaufenden molekularen Prozesse erzielt [66]. Störungen der neurologischen Entwicklung, Morphologie, des Wachstums oder der Differenzierung sind für verschiedene Erkrankungen verantwortlich. So können beispielsweise Mutationen in verschiedenen Tubulin-Isoformen Fehlbildungen des Kopfes oder des Gehirns (Mikrozephalie und Lissenzephalie) zur Folge haben. Mikrotubuli-bindende Proteine (MBPs) können die Translokation des Zellkerns

(Nukleokinese) bei der Migration oder das Neuritenwachstum beeinflussen, was zu einer starken Reduktion der Migration durch eine Verringerung des Neuritenwachstums und der Verzweigungsfähigkeit führt [66]. Das Verständnis der Entwicklungsprozesse und der unterschiedlichen Entwicklungsstufen kann helfen Abweichungen der normalen Entwicklung frühzeitig zu erkennen und Ansätze zu generieren, ausgelöste Erkrankungen zu therapieren und Beschädigungen des komplizierten Nervensystems zu regenerieren. Verfahren zur Erfassung morphologischer und molekularbiologischer Veränderungen, wie sie in der vorliegenden Arbeit entwickelt wurden, helfen die Variationen bei der Entwicklung genauer zu untersuchen und morphologische Merkmale unter verschiedenen Bedingungen bzw. im Falle einer Erkrankung zu vergleichen, um darauf basierend charakteristische Unterschiede für eine Diagnose und Prognose zu identifizieren und Ansätze für Therapiemöglichkeiten zu entwickeln.

1.1.1. Neuronale Strukturen

Nicht nur der Prozess der neuronalen Morphogenese weist eine enorme Komplexität auf, auch die Strukturen der sich entwickelnden Neuronen sind im Vergleich zu anderen eukaryotischen Zellen äußerst komplex. Die Erfassung neuronaler morphologischer Merkmale, wie sie im Abschnitt 1.1.3 vorgestellt wird, liefert, neben einfachen Merkmalen wie Größe des Zellkörpers oder Zellkerns, weitere Merkmale, die sich aus diesen komplexen Strukturen ergeben, wie z. B. Neuritenlängen und Anzahl Verzweigungspunkte, und gestaltet eine automatische Erfassung dieser Merkmale vielschichtige und nicht trivial.

Die zu untersuchenden Strukturen sich entwickelnder Neuronen werden aus einfachen sphäroiden Zellen gebildet, deren Entwicklung im ersten Schritt der Neuritogenese mit der Bildung von Neuriten einhergeht. Die Voraussetzungen zur Initiation von Neuriten sind ein dynamisches Aktinnetzwerk und die Bündelung von Mikrotubuli, die das Zytoskelett zur Bildung und zum Wachstum der Neuriten benötigt. Je nach Entwicklung bilden sich verschiedenen Morphologien der Neuronen aus, die zur Unterteilung in diverse neuronale Subtypen führt [66]. Primäre Neuronen des Hippocampus eines sich entwickelnden Mausgehirns, wie sie in dieser Arbeit betrachtet werden, sind durch mehrere kleine Neuriten gekennzeichnet. Einer dieser Neuriten bildet im weiteren Verlauf der Entwicklung eine große dynamische Wachstumskrone aus und kann sich zum späteren Axon des reifen Neurons entwickeln. Ein schematischer Aufbau eines primären hippocampalen Neurons ist in Abb. 1.1 dargestellt und zeigt die strukturellen Bestandteile, die für die Erfassung der morphologischen Merkmale in dieser

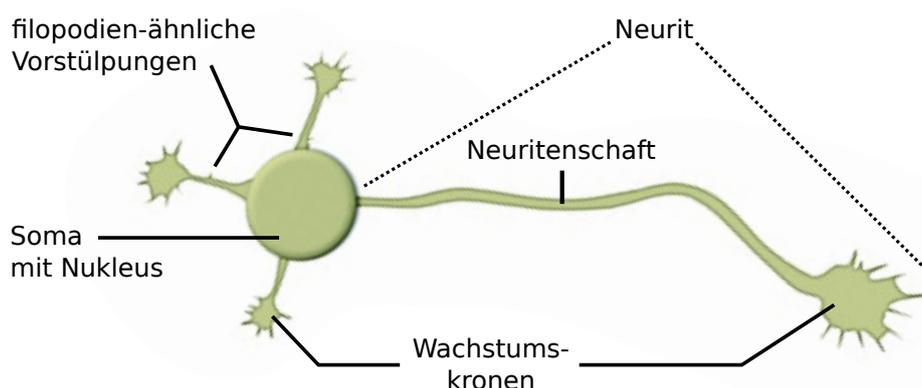


Abbildung 1.1.: Schematische Darstellung der Struktur einer sich entwickelnden neuronalen Zelle. Die strukturellen Bestandteile der Zelle gliedern sich in den Zellkörper (Soma) mit dem Zellkern (Nukleus), sowie in (sich verzweigende) Neuriten, bestehend aus Neuritenschicht, Wachstumskrone und möglichen filopodienähnliche Vorstülpungen. Erweiterte Darstellung der Abbildung nach [66].

Arbeit eine wesentliche Rolle spielen. Die anfängliche sphärische Form des Neurons mit seinem Zellkern (Nukleus) wird durch das Heraustreten der Neuriten aus dem Zellkörper (Soma) umgewandelt. Basierend auf Veränderungen der Aktin- und Mikrotubulistrukturen bilden sich die Wachstumskronen, welche aus dem Soma herauswachsen, während sich die Membran und das darunterliegende Zytoskelett beim Fortschreiten der Wachstumskronen neu organisieren, um den zylindrischen Schaft eines Neuriten zu bilden. Entlang des Neuriten können sich filopodienähnliche Vorstülpungen bilden, die der neuronalen Verzweigung und späteren synaptischen Kontaktausbildung im entwickelnden Gehirn dienen.

1.1.2. Lokale Translation in Neuronen

Das dynamische Zytoskelett der neuronalen Zelle mit seinen Mikrotubuli, Aktin- und Intermediär-Filamenten ist, wie erwähnt, für die Umstrukturierung der Zellmorphologie bei Entwicklung und Migration der Zelle verantwortlich. Die sich vom Soma herauswachsenden Wachstumskronen sowie die sich verlängernden und verzweigenden Neuritenschäfte sind im Bezug auf die dabei ablaufenden molekulare Prozesse von diversen intrazellulären Molekülen abhängig (RNAs, Proteine etc.), die diese Prozesse steuern und dem Zytoskelett die nötige Dynamik ermöglichen. Der Mechanismus der lokalen Translation von mRNAs spielt dabei eine zentrale Rolle für das Auswachsen der Wachstumskronen, der Ausrichtung der Neuriten und somit für die Entwicklung des gesamten Neurons [57, 109]. Mittels lokaler Translation wird die lokalisierte Synthese spezifischer Transkripte am Ort ihrer Funktion, wie z. B. an

Verzweigungspunkten der Neuriten oder in den Wachstumskronen, in räumlicher Distanz zur Transkription der Gene im Nukleus ermöglicht. Vorteile dieses Mechanismus sind, dass die Lokalisation der mRNA die Synthese des Proteins, an einem Ort an denen es nicht funktional oder womöglich schädlich wirken kann, verhindert. Das Protein wird somit in dem intrazellulären Kompartiment gebildet, wo es benötigt wird. Wachstumskronen bzw. Neuriten haben dadurch die Möglichkeit mit hoher zeitlicher Auflösung die lokale Proteinsynthese räumlich einzuschränken und sehr flexibel auf ihre Umgebung zu reagieren, unabhängig von Prozessen im „entfernten“ Nukleus. Zudem können auf effektive Weise die mRNA Transkripte für eine mehrfache Translation an Ort und Stelle wiederverwendet werden, anstatt die Proteine im Zellkern zu synthetisieren und bis zur Peripherie der Neuriten zu transportieren [57].

Das RNA-bindende Protein ZBP1 (*Zipcode Binding Protein 1*) ist am mRNA Transport und deren lokalen Translationskontrolle in neuronalen Prozessen beteiligt und weist seine maximale Expression während der Gehirnentwicklung auf. Mittels ZBP1 erfolgt während der neuronalen Entwicklung der Transport und die lokale Translation der β -Aktin mRNA. Die Bindung der mRNA an ZBP1 verhindert die Translation des β -Aktin Proteins und kann am benötigten Wirkungsort innerhalb der Zelle durch eine Phosphorylierung von ZBP1 mittels Scr-Kinase aktiviert und translatiert werden [57, 95, 180]. ZBP1 hat somit einen wesentlichen Einfluss auf die Ausrichtung der Wachstumskronen und die korrekte Verzweigung der Neuriten. Es hat demnach direkte Auswirkungen auf die Zellproliferation, -motilität und Morphologie und führt bei einer gestörten Expression zur Beeinträchtigung der embryonalen Entwicklung. ZBP1 aus Huhn und dessen humanes Ortholog IGF2BP1 (*IGF2 mRNA Binding Protein 1*) sind Proteine, die während der Embryogenese exprimiert werden. Die Proteine weisen ein onkofötales Expressionsmuster auf, was bedeutet, dass sie während der Embryogenese exprimiert werden, im adulten Gewebe aber weitgehend nicht vorhanden sind. In verschiedenen Tumoren und Tumor-abgeleiteten Zellen weisen diese Proteine eine *De-novo* Synthese oder signifikante Hochregulation auf, was mit dem Tumorwachstum und der Metastasierung korreliert [13, 180]. Der Zusammenhang des Expressionsmuster mit wesentlichen Aspekten der Zellfunktion, wie z. B. der Migration, Morphologie, Proliferation und Differenzierung, charakterisieren die Proteine als diagnostische Marker in aggressiven Krebsarten [13, 171] und deutet auf eine zentrale Rolle von ZBP1/IGF2BP1 bei der Differenzierung neuronaler Zellen und im Zusammenhang mit dem Neuroblastom hin [14, 180].

1.1.3. Analyse neuronaler Zellmorphologie

Um die Grundlagen neuronaler Entwicklung und Funktionen sowie neurologischer Erkrankungen besser zu verstehen, ist es wichtig, definierte morphologische Merkmale möglichst automatisch und objektiv zu erfassen, auf deren Basis morphologische Variationen ermittelt und miteinander verglichen werden können. Eine Betrachtung der Regulationsmechanismen intrinsischer Zellfunktionen anhand einer quantitativen Untersuchung der subzellulären Lokalisation von Proteinen bedingt zudem eine möglichst genaue Erfassung der neuronalen Strukturen und sich daraus ableitenden morphologischen Merkmale. Obwohl diverse Verfahren zur Untersuchung neuronaler Zellen auf Basis von hochauflösenden Fluoreszenzmikroskopiebildern etabliert sind, weisen diese eine oftmals rudimentäre und nicht automatische und somit auch eine nicht unvoreingenommene Auswertung der Bilder auf. Nicht zuletzt fehlt diesen Verfahren die Fähigkeit, in den detektierten Strukturen subzellulären Verteilungen von markierten Molekülen zu extrahieren. Die Segmentierung von langgestreckten röhrenartigen Strukturen, beispielsweise Blutgefäßen [111, 124], Pflanzenwurzeln [62] oder aber auch neuronalen Zellen [31, 205], ist bereits seit längerer Zeit und bis heute eine Herausforderung für die digitale Bildverarbeitung. Bemühungen um eine computergestützte Segmentierung von Neuronen und die Analyse der neuronalen Morphologie reichen dabei mehr als 50 Jahre zurück [73, 156]. Bis heute wurden zur Extraktion neuronaler Zellmorphologie diverse Ansätze publiziert, die in den meisten Fällen auf die Extraktion von sogenannten Neuriten Spuren (*Neurite Traces*), welche als Mittellinien entlang der Neuritenstruktur definiert sind, abzielen und die Morphologie der Zelle im Gesamten betrachten, nicht jedoch auf einzelne strukturelle Bestandteile der komplexen neuronalen Zellmorphologie eingehen. Viele dieser Algorithmen überwinden zudem nicht die Notwendigkeit manueller Eingriffe, was zu einem Engpass bei der Effizienz und der Objektivität der Analyse führt.

NeuronJ [157] und der Simple Neurite Tracer¹ sind zwei beliebte Werkzeuge, die eine halbautomatische Verfolgung und Quantifizierung von Neuriten in 2D- oder 3D-Bildern ermöglichen, die durch manuelle Einstellung von Startpunkten initialisiert werden. In [286] wird ein Tracing-Algorithmus mit automatischer Startpunkterkennung präsentiert, bei dem ohne Benutzerinteraktionen die Bereiche des Somas detektiert und die Spuren der Neuriten ermittelt werden. Die vollständigen Regionen der Neuriten werden dabei jedoch nicht lokalisiert.

¹http://imagej.net/Simple_Neurite_Tracer

Häufig erfasste morphologische Merkmale der Algorithmen sind beispielsweise Länge der Neuriten, die Anzahl von Verzweigungen und Ausstülpungen [52, 88, 157, 196]. Nur wenige generieren Daten zu weiteren Merkmalen von Neuronen oder Neuriten, wie die Anzahl der Verzweigungsschichten (Verzweigungen von primären Neuritenzweigen in sekundäre oder tertiäre Zweige), Ausläufern und Wachstumskronen oder Intensitäten in bestimmten Bereichen (maximale, mittlere und integrierte Intensität) [261, 277]. Die DIADEM Challenge² (2009-2010) generierte wesentliche Fortschritte für die computergestützten Analysen zur Rekonstruktion dreidimensionaler Neuronen. Die besten Algorithmen des Wettbewerbs funktionieren gut auf 3D-Bildern, haben aber Schwierigkeiten 2D-Bilder zu verarbeiten oder umgehen die Problematik durch die Erzeugung künstlicher 3D-Bilder aus den 2D-Daten. Ein vielversprechende Methode zur Verfolgung und Rekonstruktion von Neuronen ist in Vaa3D-Neuron implementiert [192]. Konzipiert wurde das Verfahren für 3D-Bilder, ist aber in der Lage 2D-Daten zu verarbeiten. Neben manuellen und halbautomatischen Tracing-Algorithmen enthält Vaa3D-Neuron eine automatische Methode, um die Struktur eines Neurons zu verfolgen [193]. Die Anwendung ermöglicht es die Neuritenspuren und die vollständige Fläche des Neurons zu erfassen. Mit dem Synapse Detector von [221] wurde ein automatisches Bildanalyseprogramm entwickelt, um die synaptische Rekrutierung von markierten Proteinen zu ermitteln. Darüber hinaus werden dendritische und synaptische Merkmale wie Dendritenlänge sowie Anzahl der Verzweigungen und Synapsen dokumentiert. Zur Messung der dendritischen Verzweigung wird hierbei die Sholl-Analyse angewendet [227]. Diese Analysemethode ist weit verbreitet, bietet jedoch eine begrenzte Empfindlichkeit, um Unterschiede zwischen Neuronen zu erkennen und ignoriert sowohl die Orientierung als auch die Topologie der Neuriten [254]. Die sich immer weiter entwickelnden Verfahren werden zunehmend in ihren Analysemethoden optimiert, wobei die vollständige Erkennung der Neuronen sowie deren Rekonstruktion aus diversen Bilddaten (2D, 3D, Durchlicht- oder Fluoreszenzmikroskopie) im Vordergrund steht. Einige Verfahren sind bereits in der Lage, sowohl sich überschneidende Neuriten in Kulturen von Zellen korrekt zu erkennen und den einzelnen Neuronen zuzuordnen [183, 189], als auch qualitativ schlechte Ausgangsbilder adäquat zu verarbeiten [110].

Trotz der Anwendungsvielfalt bewerten die meisten gängigen Ansätze die Morphologie der Zelle als Ganzes, ohne strukturelle Teilbereiche zu betrachten oder die räumliche Anordnung einzelner Biomoleküle innerhalb der Zelle oder gar Strukturen zu analysieren. Eine automatisierte

²<http://www.diademchallenge.org/>

Analyse der Zellmorphologie mit einer damit einhergehenden räumlichen und quantitativen Bewertung von molekularen Profilen sind derzeit nicht verfügbar und erfordern die Entwicklung neuer methodischer Ansätze. In der hier vorliegenden Arbeit wurden vor diesem Hintergrund Verfahren entwickelt, die eine automatische Erkennung der neuronalen Zellen ermöglichen, sowie deren einzelnen strukturellen Bestandteile erfassen und in der Lage sind, subzellulär markierte Moleküle zu lokalisieren und quantitativ zu bewerten.

1.2. Krebserkrankung Neuroblastom

Das Neuroblastom ist eine bösartige Erkrankung des sympathischen Nervensystems, die vor allem im frühen Kindesalter auftritt und vorwiegend Kleinkinder bis zum sechsten Lebensjahr betrifft. Der solide Tumore tritt überall dort auf, wo sich sympathisches Nervengewebe befindet, das als Teil des autonomen Nervensystems die unwillkürlichen Funktionen wie Herz- und Kreislauf, Darm- und Blasenätigkeit steuert [148]^{3,4}. Am häufigsten entstehen die Tumore im Nebennierenmark (ca. 50 % der Fälle) und im Bereich der Nervengeflechte auf beiden Seiten der Wirbelsäule, dem so genannten Grenzstrang. Ist der Grenzstrang betroffen, können auf jeder Höhe entlang der Wirbelsäule Neuroblastome vorkommen. Dazu zählen Bereiche im Bauch, dem Becken, der Brust und des Halses. Die Mehrzahl der Tumore (ca. 70 %) befindet sich im Bauchraum, etwa ein Fünftel sind im Brust- und Halsbereich lokalisiert. Die Neuroblastome können auf ihren Ursprungsort begrenzt sein oder in nahe gelegene Lymphknoten streuen. Bei einigen Patienten können Metastasen, als Absiedelungen der bösartigen Zellen, im Knochenmark, den Knochen, der Leber oder der Haut, selten im Gehirn oder der Lunge auftreten. Als Ursache für die Entstehung der Krebserkrankung ist eine bösartige Veränderung unreifer, undifferenzierter (embryonaler) Nervenzellen verantwortlich. Eine Differenzierung der Zellen kann durch verschiedene Therapien forciert und deren Effektivität mittels unterschiedlicher Kombination von Agenzien gesteigert werden [246]. Untersuchungen zu Therapiewirksamkeit setzen eine Betrachtung der differenzierenden Zellen hinsichtlich ihrer morphologischen Veränderungen voraus. Neben rein visuellen Betrachtungen mikroskopischer Bilder durch einen Wissenschaftler können automatische Verfahren zur Morphologieanalyse, wie sie innerhalb dieser Arbeit entwickelt wurden, eine Vielzahl von Bildern unter verschiedenen Bedingungen gleichzeitig auswerten, um somit vielversprechende Therapieansätze aufzuzeigen. Die Erfassung

³<https://www.kinderkrebsinfo.de>

⁴<https://www.krebsgesellschaft.de>

und der Vergleich morphologischer Merkmale, wie z. B. Zellgröße, -form oder Neuritenlänge und -anzahl, ermöglichen es zudem die Wirksamkeit einer Therapie für Differenzierung und Wachstum genauer abzuschätzen und verschiedene Therapiemöglichkeiten effizient zu kombinieren. Die Entartung der undifferenzierten Zellen im Neuroblastom findet womöglich bereits vor der Geburt statt und kann Folge von Chromosomenveränderungen und/oder Genmutationen sein.

Das Neuroblastom weist unter den menschlichen Krebsarten den höchsten Anteil an spontaner und vollständiger Regression auf, also der Rückbildung des Tumors [148]. Nach der akuten lymphatischen Leukämie (ALL) ist das Neuroblastom die zweithäufigste bösartige Krebserkrankung im Kindesalter und betrifft ca. 7 % aller Krebserkrankungen bei Kindern und Jugendlichen. Es stellt neben den Tumoren des Zentralnervensystems (Hirntumoren) die häufigsten soliden Tumore dar. Betroffenen Patienten sind überwiegend jünger als sechs Jahre (ca. 90 %), wovon etwa 40 % der Betroffenen Neugeborene und Säuglinge im ersten Lebensjahr sind. Die Wahrscheinlichkeit einer Erkrankung nimmt mit zunehmendem Lebensalter ab, kann jedoch auch bei älteren Kindern, Jugendlichen und in Einzelfällen bei Erwachsenen auftreten. Nach Angaben des Deutschen Kinderkrebsregisters erkranken in Deutschland jährlich im Durchschnitt 120 Kinder bzw. Jugendliche unter 15 Jahren an einem Neuroblastom (Zeitraum 2006 bis 2015), womit etwa eines von 100.000 Kindern pro Jahr von der Erkrankung betroffen ist [220]⁵. Die 5-Jahres-Überlebensraten haben sich im Zeitraum von 1975 bis 1977 von 52 % auf 74 % in den Jahren 1999 bis 2005 erhöht. Effektive Maßnahmen zur weiteren Verbesserung der Überlebensrate durch neue Therapieansätze sind jedoch in den letzten 20 Jahren ins Stocken geraten, da die Optimierung der Effizienz bei gleichzeitiger Verminderung der Schädigung des Patienten ihre Grenzen erreicht haben [70, 148]. Neue Forschungsansätze sowie internationale Zusammenarbeiten sind erforderlich, um diese Grenzen zu überschreiten und neue Wege für erfolgreiche Therapien zu eröffnen.

1.2.1. Klinische Einteilung

Die zur klinischen Einteilung notwendige Diagnose des Neuroblastoms erfolgt im Mittel meist zwischen dem 18. und 22. Lebensmonat. Säuglinge unter 18 Monaten tendieren dabei mehr zur spontanen Regression des Tumors oder zur Genesung nach chirurgischen Eingriffen, während ältere Kinder zu metastasierenden und aggressiven Tumoren neigen, die

⁵<http://www.unimedizin-mainz.de/?id=22802>

nicht ohne cytotoxische Therapie behandelt werden können [178]. Beim Verdacht auf ein Neuroblastom sind für eine optimale Behandlung und Prognose verschiedene Untersuchungen notwendig. Dazu gehören beispielsweise Laboruntersuchungen von Blut und Urin hinsichtlich bekannter Tumormarker, bildgebende Verfahren wie Sonographie, Magnetresonanz- oder Computertomographie zur Abgrenzung von anderen möglichen Tumorerkrankungen. Weiterhin dienen diese Untersuchungen zur Bestimmung der Lage und Größe des Tumors und zur Einschätzung des Befalls benachbarter Strukturen, wie Organe oder Blutgefäße. Zusätzlich dazu werden Knochenmarksuntersuchungen zur Prüfung auf Metastasenbildung angewendet. Eine eindeutige Diagnose der Erkrankung kann jedoch nur aufgrund einer genaueren Untersuchung des Tumorgewebes mittels histologischer Beurteilung durch einen Pathologen erfolgen. Für weitere Charakterisierungen muss die Beschaffenheit des entnommenen Tumorgewebes untersucht und eine molekulargenetische Analyse erfolgen, um die Aggressivität des Tumors zu beurteilen und eine Prognose sowie Behandlung für den Patienten festzulegen. Die Bösartigkeit des Tumors kann meist anhand bekannter Mutationen, also Veränderungen in der DNA der Tumorzellen, wie z. B. einer Amplifikation des *MYCN* Gens (*MYCN* Amplifikation) oder den Verlust von chromosomalen Bereichen bestimmt werden, die je nach Auftreten oder Fehlen mit einer günstigen oder ungünstigen Prognose einhergehen [178]. Die klinische Einteilung der Tumorerkrankung erfolgt anhand verschiedener Krankheitsstadien und/oder Risikofaktoren, basierend auf diversen Unterscheidungsmerkmalen aus den oben genannten Untersuchungen.

Krankheitsstadien des Neuroblastoms

Die Ausbreitung des Neuroblastoms innerhalb des Körpers spielt eine wesentliche Rolle bei der Einteilung in die Krankheitsstadien und für die Strategie zur Behandlung des Tumors im Patienten. Als Standard für die Einteilung der Tumore in verschiedene Krankheitsstadien dient das im Jahre 1986 entwickelte *International Neuroblastoma Staging System* (INSS). Diese Einteilung ist mit einem operativen Eingriff verbunden und berücksichtigt zur Unterscheidung verschiedene Aspekte wie die Lage des Tumors, dessen Größe, den Befall von Lymphknoten, das Vorhandensein von Metastasen sowie die Möglichkeit den Tumor mittels chirurgischen Eingriffs vollständig oder unvollständig zu entfernen [169, 178]. Das Neuroblastom wird anhand des INSS in fünf Krankheitsstadien eingeteilt, deren Unterscheidungsmerkmale in Tab. 1.1 zusammengefasst sind.

Eine Besonderheit, die beim Neuroblastom auftritt, ist die Möglichkeit zur Differenzierung der

Krankheitsstadium	Merkmal
1	Tumor begrenzt auf Ursprungsort, kein Befall von Lymphknoten, vollständig entfernbar
2	Tumor vollständig oder teilweise entfernbar, Befall nur einer Körperseite der Wirbelsäule ohne (Stadium 2a) oder mit (Stadium 2b) benachbarten Lymphknoten
3	Tumore mit Wirbelsäulenüberschreitung, nicht vollständig entfernbar, Befall von Lymphknoten der gegenüberliegenden Körperseite
4	Fernmetastasen des Tumors, z. B. in Knochenmark, Leber, Haut, oder entfernte Lymphknoten
4S	betrifft nur Säuglinge unter 12 Monaten, begrenzte Tumore mit Metastasen in Haut, Leber und/oder minimalem Befall des Knochenmarks

Tabelle 1.1.: Krankheitsstadien des Neuroblastoms, eingeteilt nach dem *International Neuroblastoma Staging System (INSS)*.

Zellen und spontanen Zurückbildung (Regression) des Tumors. Dies geschieht häufig im frühen Säuglingsalter (vor dem 12. Lebensmonat) bei Tumoren, die dem Stadium 4S zugeordnet werden können und somit nur vereinzelt Metastasen in Leber, Haut oder Knochenmark aufweisen. Diese Tumore können sich durch eine geringfügige Chemotherapiebehandlung oder, bei lokal begrenzten Tumoren, auch ohne Therapie zurückbilden. Die Tumorzellen sterben dabei durch Apoptose ab, ein Vorgang des selbstausgelösten Zelltods innerhalb der Tumorzellen. Eine mögliche Differenzierung der Zellen, z. B. durch die Hemmung von Signalwegen oder der Zugabe von Retinoiden, kann zudem zu einer Verringerung des unkontrollierten Wachstums und der Verbreitung des Tumors führen [25, 148]. Aktuelle Untersuchungen zeigen jedoch, dass die Regression oder Differenzierung des Neuroblastoms nicht nur auf das Stadium 4S begrenzt ist und auch in anderen Stadien beobachtet werden kann [25].

Die Einteilung des Neuroblastoms nach dem INSS in seine fünf Krankheitsstadien ist ein Standard, der in vielen Ländern Anwendung findet, jedoch mit den Jahren auch Defizite für die Diagnose und Therapie der Tumorerkrankung aufgezeigt hat [169]. Zum einen wird für die Einteilung nach diesem System eine Biopsie, Operationen und/oder die Untersuchung von Knochenmark benötigt, was zusätzliche Risiken für den Patienten bedeuten kann. Zum anderen weisen viele Patienten oftmals eine Erkrankung mit geringem Risiko auf, so dass keine Operation und Entnahme von Tumorgewebe stattfindet, wodurch eine Einteilung nach dem INSS nicht erfolgen kann und eine nachfolgende Charakterisierungen des Tumors sowie eine Prognose und Therapie erschwert wird. Weiterhin ist nachteilig, dass ein Tumor dem Stadium 1 oder 3

zugeordnet werden kann, abhängig vom Ausmaß des durchgeführten chirurgischen Eingriffs und der ärztlichen Einschätzung des Lymphknotenbefalls. Auch eine bereits mögliche Regression bei lokalisierten Tumoren erschwert die genaue Einschätzung mithilfe des INSS [169]. Die Einteilung nach dem INSS weist somit nachteilige Aspekte auf und ist ungeeignet für eine Vorbehandlung und Prognose der Tumorerkrankung ohne operative Eingriffe.

Aus diesen Gründen wurde 2009 das *International Neuroblastoma Risk Group* (INRG) Klassifizierungssystem für das Neuroblastom entwickelt [44]. Das damit einhergehende System zur Einteilung in verschiedene Krankheitsstadien wird als *International Neuroblastoma Risk Group Staging System* (INRGSS) bezeichnet und soll nicht als Ersatz für das bestehende INSS dienen, sondern dieses vielmehr erweitern [169]. In den meisten Ländern werden beide Systeme simultan genutzt, um die Erkrankung zu charakterisieren und eine Behandlungstherapie zu erarbeiten. Die Einteilung nach dem INRGSS basiert auf klinischen Kriterien und bildgebenden Verfahren der Tumore, ohne dass ein chirurgischer Eingriff zur Anwendung kommt. Das System soll eine zusätzliche, robuste und unabhängige Einteilung der Krebserkrankung während der Diagnose und vor jedweder Therapie ermöglichen. Wesentliche Unterschiede zum INSS sind, dass das System auf präoperativen Bildern basiert, aus denen sogenannte bildbasierte Risikofaktoren (IDRFs, *Image-Defined Risk Factors*) ermittelt werden, keine Einbeziehung der Körperseite oder Lymphknoten stattfindet und das maximale Alter von 12 Monaten im Stadium 4S auf 18 Monate erhöht wurde [169, 178]. Neben der Einteilung nach dem INRGSS, dem Grad der Tumordifferenzierung, dem Status des *MYCN* Onkogens und weiterer klinisch relevanter Faktoren, beinhaltet das INRG Klassifizierungssystem eine auf vorherigen Überlebenszeitanalysen basierende Einschätzung der Risikogruppe für die vorliegende Tumorerkrankung, welche sich in die Gruppen „sehr geringes“ bis „geringes Risiko“ (*very low risk, low risk*), „mittleres Risiko“ (*intermediate risk*) und „hochrisikoreiches Neuroblastom“ (*high risk neuroblastoma*) gliedern. Da eine Einteilung der in dieser Arbeit analysierten Tumorproben nur nach dem INSS erfolgte, werden die diversen Kriterien des INRG Klassifizierungssystem sowie die Einteilung nach dem INRGSS hier nur erwähnt, ohne diese in der hier vorliegenden Arbeit detaillierter zu betrachten.

1.2.2. Therapie und Prognose

Beim Neuroblastom handelt es sich um eine sehr heterogene Krebserkrankung, die durch diverse Faktoren charakterisiert ist. Hierzu zählen beispielsweise das Diagnosealter, Stadium der

Erkrankung und verschiedene biologische Faktoren, die für die Überlebenswahrscheinlichkeit relevant sind. Die Integration genetischer Veränderungen und prognostischer biologischer Faktoren durch neue Forschungserkenntnisse und Techniken führen zu gezielteren Behandlungen und zu einer Verbesserung der Überlebenswahrscheinlichkeit. Die Therapie des Neuroblastoms erfolgt aufgrund seiner Heterogenität auf vielfältige Art und Weise und erfordert Kombinationen multidisziplinärer Ansätze, wie Chirurgie, Chemotherapie, Bestrahlung, Stammzellenrettung oder Immuntherapie, um möglichst erfolgreiche Behandlungsergebnisse und eine günstige Prognose zu erzielen [148,178]. Dabei sollen vor allem chirurgische Komplikationen gemindert und eine Chemotherapie-Toxizität reduziert werden, indem für Tumore mit günstiger Prognose die Intensität therapeutischer Maßnahmen reduziert wird, während bei einer ungünstigen Prognose intensivere Maßnahmen ergriffen werden.

Die Therapie des Neuroblastoms richtet sich nach dem Risiko der Erkrankung für den Patienten. Bei einem geringen bis mittleren Risiko werden die Tumore operativ entfernt und zusätzlich wird der Patient mittels Chemotherapie geringer Intensität behandelt. Das Ziel der Kombination aus Operation und Chemotherapie ist eine Verringerung des Tumorumfanges um mindestens 50 % [178] sowie die Gewinnung von ausreichend Material für weitere Untersuchungen, bei gleichzeitiger Vermeidung von Schäden am umliegenden Gewebe [148]. Begrenzte Tumore können zur spontanen Regression neigen, ohne operative Eingriffe vornehmen zu müssen. Tumore des Stadiums 4S werden meist mit geringen Dosen an Chemotherapeutika behandelt, um das Risiko durch Folgeschäden zu minimieren. Die Behandlung des hochrisikoreichen Neuroblastoms teilt sich hingegen in drei Phasen. Erstens die Verringerung der Krankheitssymptome (Remission), die weitere Verstärkung der Remission und letztlich die Erhaltungsphase mit einer Zerstörung der restlichen Tumorzellen [148]. Die Therapie beinhaltet eine intensive Chemotherapie mit verschiedenen Arten von Zytostatika, wie z. B. Doxorubicin, Cisplatin und Vincristin [148,178]. Zudem kommen immer mehr biologische Therapien zum Einsatz, wie beispielsweise Retinoide (Isotretinoin, auch bekannt als 13-*cis*-Retinsäure), das zur neuronale Differenzierung *in vitro* führt, oder monoklonale Antikörper gegen das Gangliosid GD2, welches auf der Zelloberfläche der Neuroblastomzellen vorkommt und sich somit als Ziel (*Target*) für eine Krebsimmuntherapie eignet. Des Weiteren bieten Weiterentwicklungen in den Sequenzierungstechnologien die Möglichkeiten molekulare Veränderungen in die klinischen Befunde zu integrieren, um neue Therapieansätze zu entwickeln und die Behandlungsergebnisse zu verbessern [178].

Dennoch erschweren in vielen Fällen die verschiedenen Faktoren, wie das Ausmaß der Erkrankung,

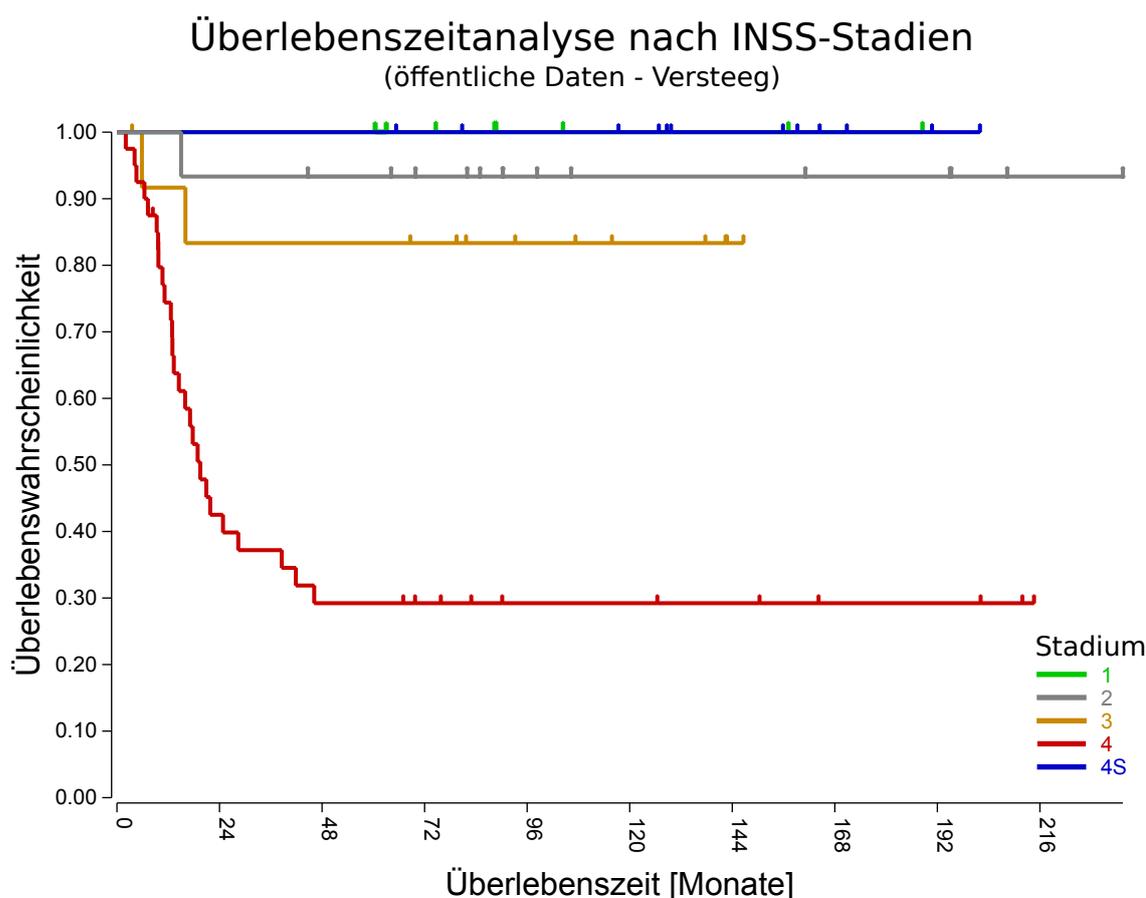


Abbildung 1.2.: Überlebenswahrscheinlichkeit am Neuroblastom erkrankter Patienten unter Betrachtung der verschiedenen Krankheitsstadien. Die Überlebenswahrscheinlichkeit in Abhängigkeit vom Stadium und der Überlebenszeit sind als Kaplan-Meier Kurven dargestellt. Abbildung erstellt mittels R2 Genomics Analysis and Visualization Platform (<http://r2.amc.nl>) und nachträglich adaptiert. Grundlage bilden die öffentlichen Daten aus [167].

die Aggressivität und das Alter zum Zeitpunkt der Diagnose, die Erstellung einer Prognose bei dieser Erkrankung. Kinder unter sechs Monaten und Kinder mit kleinen, lokalisierten Tumoren (Stadium 1 und 2) haben in der Regel eine gute Prognose. Gleiches gilt für Kinder mit Tumoren des Stadiums 4S, welche durch eine vollständige Regression der Erkrankung geheilt werden können (Überlebensrate > 90 %). Für ältere Kinder, die an metastasierenden Tumoren oder an einem hochrisikoreichen Neuroblastom erkrankt sind, liegt - trotz intensiver Therapiemöglichkeiten - immer noch eine (sehr) ungünstige Prognose vor (Überlebensraten von unter 50 %) [148, 178]. Anhand der Ergebnisse der Überlebenszeitanalyse in Abb. 1.2 werden die unterschiedlichen Überlebenswahrscheinlichkeiten und die damit einhergehenden guten bzw. schlechten Prognosen bei Betrachtung der verschiedenen Krankheitsstadien des

Neuroblastoms veranschaulicht, basierend auf den Daten von [167]. Aktuelle Forschungen und Analysen lassen hoffen, dass ein besseres Verständnis der zellulären und molekularbiologischen Prozesse im Neuroblastom verbesserte Behandlungsstrategien ermöglicht als auch neue und effektive therapeutische Targets identifiziert, die zu einer effizienten Therapie und wesentlichen Erhöhung der Überlebenschancen der erkrankten Kinder führen wird.

1.3. Untersuchung des Transkriptoms

Untersuchungen der Zellmorphologie und Proteinlokalisierung, wie sie im Abschnitt 1.1.3 vorgestellt wurden, ermöglichen die Betrachtung phänotypischer Variation im Zusammenhang mit einfachen molekularbiologischen Begebenheiten. Zusätzliche Betrachtungen transkriptomischer Veränderungen zwischen unterschiedlichen biologischen Bedingungen, Zelltypen oder Zellstadien in Kombination mit diesen Erkenntnissen ermöglichen ein besseres Verständnis von Zusammenhängen und Prozessabläufen in diversen menschlichen Erkrankungen. Das Transkriptom einer Zelle beschreibt hierbei die vollständige Menge an Transkripten, die zu einem spezifischen Zeitpunkt einer Entwicklungsphase oder einer physiologischen Bedingung in der Zelle präsent sind. Zur Interpretation gegenwärtiger Funktionen des Genoms und zum Verständnis der Entwicklung und Veränderung der Zellen sowie damit einhergehender Krankheiten, ist das Verständnis des Transkriptoms essentiell und bedarf einer Quantifizierung des sich verändernden Niveaus exprimierter Transkripte.

1.3.1. Hochdurchsatz-RNA-Sequenzierung

Durch die Entwicklung und Verbesserung von *Next Generation Sequencing* Verfahren können heutzutage mittels paralleler Sequenzierungen von kurzen DNA-Molekülen eukaryotische Genome in geringer Zeit vollständig und kostengünstig sequenziert werden. So betrug die durchschnittliche Sequenzierung eines menschlichen Genoms im Jahre 2017⁶ rund 1000 €. Die Sequenzierung von RNA-Transkripten ermöglicht tiefere Einblicke in das Transkriptom und damit auf die exprimierten Transkripte innerhalb einer untersuchten Probe. Unter Anwendung der Hochdurchsatz-RNA-Sequenzierung (RNA-Seq) können bekannte als auch noch unbekannte Transkripte identifiziert und quantifiziert werden. Veränderungen in der Genexpression unter verschiedenen Bedingungen, Entwicklungsstadien oder zwischen gesundem

⁶www.genome.gov/sequencingcosts/

und erkranktem Gewebe können somit quantitativ erfasst und gegenübergestellt werden. Verglichen mit anderen Techniken, wie z. B. Sanger-Sequenzierungen oder Microarrays, weist RNA-Seq diverse Vorteile auf, wie beispielsweise die erwähnten geringen Kosten, Benötigung geringer RNA-Mengen, eine hohe Auflösung im Bereich von Einzelbasen, geringe Hintergrundsignale, einer höheren Dynamik, die Möglichkeit Isoformen zu detektieren und alle vorhandenen Transkripte zu quantifizieren [263]. Hybridisierungsbasierte Methoden, wie Microarrays, beinhalten fluoreszenzmarkierte Moleküle und benötigen Kenntnisse über das untersuchte Genom, dessen Transkripte bekannt sein müssen. Diese Verfahren sind vergleichsweise teuer und weisen hohe Hintergrundsignale oder Sättigungen bedingt durch Kreuzhybridisierungen auf. Zudem ist ein Vergleich von einzelnen Proben nicht trivial, da die Effizienz der Hybridisierung nicht vollumfänglich geprüft werden kann, und eine geeignete, oftmals komplizierte, Normalisierung der Daten benötigt wird [263].

Sequenzierung

Um mittels RNA-Seq das Transkriptom einer Probe zu bestimmen, wird zunächst die RNA einer Probe isoliert, welche aus der total oder fraktionierten RNA bestehen kann. Die isolierte RNA wird mittels reverser Transkription in eine Bibliothek von cDNA-Fragmenten (*Libraries*) umgewandelt. Da die Zuverlässigkeit der katalysierenden Polymerase mit zunehmender Sequenzlänge abnimmt und um mit den heutigen Sequenzierungstechnologien kompatibel zu sein, werden lange RNAs auf Längen von ca. 200 bis 500 Basen fragmentiert, wodurch überlappende Teilsequenzen langer Transkripte erzeugt werden. Je nach Art der Fragmentierung hat dies Auswirkungen auf die erzeugten Fragmente, wie z. B. eine schlechtere Qualität der Transkriptenden bei vorheriger RNA-Fragmentierung, während eine cDNA-Fragmentierung bessere Qualität der 5'-Enden der Sequenzen generiert, innerhalb eines Transkriptes jedoch schlechtere Qualitäten erzeugt [263]. Zur Sequenzierung kurzer RNA-Moleküle, wie beispielsweise microRNAs (kurze einzelsträngige RNAs einer Länge von ca. 22 Nukleotiden), findet keine Fragmentierung, sondern eine Größenselektion der Transkripte, z. B. mittels Gelelektrophorese, statt. Für microRNAs erfolgt die Selektion für Moleküle, die eine Größe von ca. 30 Nukleotiden aufweisen, wobei neben microRNAs auch weitere kurze RNA-Moleküle (rRNAs, tRNAs etc.) in den erzeugten Libraries enthalten sind. Die umgewandelten cDNA-Fragmente der Libraries enthalten Sequenzierungsadapter zur Immobilisierung der Fragmente, welche an einem oder beiden Enden angehängt sind. Diese

Oligonukleotidadapter sind für die nachfolgenden Sequenzierungsschritte notwendig und können zur Amplifikation der Transkripte mittels Polymerasekettenreaktion (PCR) genutzt werden. Mittels Hochdurchsatz werden anschließend die cDNA-Fragmente von nur einem Fragmentende (*Single-End*) oder beiden Enden (*Paired-End* Verfahren) sequenziert, um kurze Sequenzen mit ca. 30 bis 400 Basen der Transkripte zu erhalten. Spezifische Protokolle ermöglichen es, verschiedene Adapter an das 5'- bzw. 3'-Ende der Fragmente zu ligieren, um eine strangspezifische Sequenzierung der Transkripte von einer definierten Seite des DNA-Stranges zu ermöglichen. Dadurch ist die Orientierung der Transkripte bekannt, was von Vorteil für die spätere Annotation der Transkripte ist. Somit können überlappenden Transkripte vom gegenüberliegenden Strang getrennt und korrekt dem Genom zugeordnet werden. Nachteilig können sich hierbei jedoch direkte RNA-RNA-Ligationen auswirken oder die Möglichkeit, dass vermeintliche Antisense-Transkript als Artefakte der reversen Transkription entstehen [263]. Das Prinzip der Hochdurchsatzsequenzierung folgt trotz diverse Sequenzierungstechniken (Illumina, Roche, Applied Biosystems, Oxford Nanopore etc.) ähnlichen Ansätzen. Die cDNA-Fragmente werden nach ihrer Einzelstrang-Denaturierung immobilisiert und amplifiziert. Damit werden die Sequenzen auf der Durchflusszelle (*Flow Cell*) im Sequenzierer räumlich angeordnet und es kommt zur Verstärkung der Signalstärke pro Sequenz. Die amplifizierten Sequenzen dienen als Template für den Einbau komplementärer Nukleotide mit einer nukleotidspezifischen Fluoreszenz, welche vom Gerät optisch erfasst wird. Millionen von Sequenzen können somit nukleotidweise registriert und kurze Nukleotidsequenzen, sogenannte Reads, von den Templates erzeugt werden. Abhängig von der eingesetzten Technik bzw. Sequenzierungsmaschine variiert die mögliche Länge der erzeugten Reads als auch deren Anzahl. Neue Geräte zur Analyse vollständiger Transkripte, können Sequenzlängen von bis zu 30000 Basen verarbeiten, wobei die Anzahl der Reads (ca. 500000) oftmals geringer ausfällt (Pacific Biosciences). Geräte, die auf die Erzeugung kurzer Sequenzen spezialisiert sind (50 bis 500 Basen) erreichen dabei die enorme Anzahl von bis zu 130 Millionen Reads (Illumina).

Verarbeitung

Die Verarbeitung der erzeugten Reads der Sequenzierung erfolgt in mehreren Schritten. Zunächst werden in den Reads enthaltenen Nukleotide mit schlechter Sequenzierungsqualität oder gegebenenfalls enthaltene Adaptersequenzen entfernt. Anschließend folgt als aufwändigster Schritt die Zuordnung, dem sogenannten *Mapping* oder *Alignment*, der einzelnen Reads zu

einem Referenzgenom. Für diesen Schritt existieren heutzutage zahlreiche Algorithmen, die auf unterschiedlichste Fragestellungen, Sequenzlängen oder Präzision optimiert sind und Millionen von Reads dank aktueller Hardware in wenigen Stunden verarbeiten können. Außerdem ist es möglich, auch ohne vorhandenes Referenzgenom, das Transkriptom, z. B. unbekannter oder noch nicht sequenzierter Organismen, mithilfe überlappender Sequenzen der Reads zusammensetzen. Probleme beim Alignment können generell bei Reads auftreten, die mehreren Positionen im Genom zugeordnet werden können. Dies ist beispielsweise bei microRNAs sowie bei repetitiven Sequenzen möglich, welche ähnliche oder identische Sequenzen aufweisen. Eine nicht korrekte Zuordnung kann dann in folgenden Schritten zu einer fehlerhaften Quantifizierung der Gene führen. Maßnahmen wie die Verwendung längerer Sequenzen oder die Anwendung von *Paired-End* Sequenzierungen können diese sogenannten *Multiplen Alignments* reduzieren [263]. Im Anschluss an das Alignment erfolgt unter Verwendung bekannter Genannotationen der zugeordneten Reads eine Abschätzung der relativen Abundanz der sequenzierten Transkripte. Dabei wird die Anzahl der Reads bestimmt, die einen entsprechenden Bereich des Genoms überlagern. Für nachfolgende biologisch relevante Analysen der Genexpression ist zu beachten, dass neben der Anzahl der zugeordneten Reads pro Transkript auch die Länge der untersuchten Transkripte eine Rolle spielt, da längere Transkripte durch eine höhere Anzahl an Reads repräsentiert werden könnten. Zudem ist es wichtig, besonders beim Vergleich zwischen den sequenzierten Proben, die Zusammensetzung der sequenzierten RNA-Population zu betrachten als auch die unterschiedliche Gesamtanzahl an Reads in den zu vergleichenen Proben. Für die Betrachtung einer veränderten Genexpression zwischen zwei oder mehreren Proben sollte dies berücksichtigt und mittels geeigneter Normalisierungsmethoden reguliert werden [55, 211]. Die Herausforderung in der Verarbeitung der sequenzierten Reads besteht nicht zuletzt jedoch auch im Umgang mit der enormen Datenmenge, für die entsprechende Hardware und effiziente Techniken zum Speichern, Austauschen und Weiterverarbeiten nötig sind.

1.4. Zielstellung der Arbeit

Um morphologische Veränderungen in Zellen, z. B. unter verschiedenen Bedingungen wie Stress, Wachstums, Differenzierung etc., untersuchen zu können, ist eine Erfassung der morphologischer Merkmale notwendig, die je nach Zelltyp sehr unterschiedlich gekennzeichnet sind. Besondere Herausforderungen stellen dabei neuronale Zellen dar, deren Ausläufer (Neuriten) die Erfassung

der Zellgrenzen erschweren, da sie den großen Bereich eines typischen (meist rundlichen) Zellkörpers um lange, sich oftmals verzweigenden, dünne Neuriten erweitern. Die Lokalisation und Verteilung enthaltener Proteinen in der Zelle kann zusätzlich zur Betrachtung der phänotypischen Variation der Morphologie Hinweise auf ablaufende und sich verändernde molekularbiologischer Prozesse in der gesamten Zelle und deren einzelnen Strukturen geben. Detailliertere molekularbiologische Analysen durch neueste Sequenzierungsverfahren geben hierzu einen noch weitaus tieferen Einblick in die genomischen Veränderungen zwischen unterschiedlichen biologischen Bedingungen, Zelltypen oder z. B. im Vergleich von Zellen gesunder und erkrankter Gewebe. Untersuchungen der Zellmorphologie, Proteinlokalisierung und Veränderungen in der Genexpression bilden nur den Bruchteil zum Verständnis der äußerst komplexen Abläufe innerhalb diverser menschlicher Krankheiten. Zusammen betrachtet können sie jedoch wichtige Hinweise auf biologische Vorgänge und sich wandelnde Prozesse innerhalb der Zelle oder des Organismus offen legen und neue Zusammenhänge und Möglichkeiten zur Behandlung aufzeigen. Insbesondere im Bereich der Krebsforschung bilden diese Punkte relevante Untersuchungsansätze.

Die in der vorliegenden Arbeit untersuchten Gebiete Morphologie und Genexpression bilden die Grundlage zur Untersuchung wesentlicher Faktoren von Tumorerkrankungen, wie beispielsweise Migration der Zellen, Invasivität, Differenzierung, Wachstum, und sind kennzeichnend für die Diagnose, Prognose und Therapie solcher Erkrankungen. Im Rahmen der vorliegenden Arbeit sollten Verfahren entwickelt und implementiert werden, die basierend auf Methoden der Bild- und Genexpressionsanalyse zur automatischen Erfassung morphologischer Merkmale und Veränderungen in der Genexpression dienen. Grundlage dafür bilden zweidimensionalen Mikroskopbilder fluoreszenzmarkierter neuronaler Zellen sowie RNA-Sequenzierungsdaten zur transkriptomischen Untersuchung primärer Tumorproben verschiedener Krankheitsstadien am Neuroblastom erkrankter Kleinkinder. Ziel ist es zum einen eine frei verfügbare Anwendung zu generieren, mit deren Hilfe Fluoreszenzbilder neuronaler Zellen hinsichtlich ihrer morphologischen Merkmale untersucht und zum anderen intrazellulär markierten Proteine der Zellen lokalisiert werden können. Dabei soll im Speziellen die Lokalisation von ZBP1 innerhalb primärer hippocampaler Neuronen aus Mausembryonen erfasst und erstmals ein vollständiges Verteilungsprofil des Proteins entlang der Neuriten vom Soma bis zur Wachstumskrone dargestellt werden. Weiterhin soll ein geeigneter Arbeitsablaufs zur Analyse von RNA-Sequenzierungsdaten von Tumorproben mit nachfolgender Untersuchung der Veränderungen

in der Genexpression der Krankheitsstadien des Neuroblastoms entwickelt werden. Dabei sollen insbesondere signifikant veränderte mRNAs Protein-codierender Gene sowie microRNAs und deren Targets zwischen den Krankheitsstadien betrachtet und charakteristische Gene in Form von Gensignaturen ermittelt werden, um mögliche Ansätze für therapeutische Targets oder Signalwege aufzuzeigen. Die in dieser Dissertation entwickelten Methoden und daraus gewonnen Erkenntnisse aus Morphologie-, Transkriptom- und microRNA-Target-Analyse sollen genutzt werden, um verschiedenste Krebserkrankungen hinsichtlich phänotypischer als auch genotypischer Veränderungen zukünftig umfassender untersuchen zu können und ein besseres Verständnis über die Abläufe im Neuroblastom und anderen Erkrankung zu erlangen.

2. Material und Methoden

2.1. Fluoreszenzbasierte Analyse neuronaler Zellmorphologie

Die Erfassung morphologischer sowie molekularbiologischer Veränderungen innerhalb neuronaler Zellen dienen dem besseren Verständnis der Entwicklungsprozesse und der unterschiedlichen Entwicklungsstufen der neuronalen Morphogenese und helfen die Variationen der Entwicklung, bei Krankheiten oder unter Umwelteinflüssen, zu untersuchen und untereinander zu vergleichen. Die erfassten Veränderungen in Form definierter Merkmale können genutzt werden, um charakteristische Unterschiede der Entwicklung festzuhalten oder im Falle einer Erkrankung, um eine Diagnose, Prognose und mögliche Therapieansätze zu erarbeiten. Die in dieser Arbeit entwickelten Methoden basieren auf zweidimensionalen Fluoreszenzmikroskopiebildern und dienen der automatischen Erfassung definierter morphologischer Merkmale der neuronalen Zellen. Neben üblichen Merkmalen, wie Zellfläche und Neuritenlängen, werden weitere Merkmale anhand struktureller Unterteilungen der Zellen vorgenommen. Zu den morphologischen Merkmalen werden des Weiteren intrazellulär markierte Moleküle entlang der Neuriten vom Beginn des Neuritenschaftes bis hin zur Peripherie der Wachstumskronen quantifiziert und deren Verteilung anhand eines Proteinprofils graphisch dargestellt. Dies ermöglicht neben der morphologischen Betrachtung die Einbeziehung molekularbiologischer Prozesse innerhalb der sich verändernden Neuronen und lässt Analysen zur Proteinlokalisierung und Verteilung sowie daran gekoppelter Prozesse, wie beispielsweise der lokalen Translation von Proteinen, innerhalb der Zellen zu. Die entwickelte fluoreszenzbasierte Analyse neuronaler Zellen in der vorliegenden Arbeit gliedert sich in drei Teile, bestehend aus der Detektion der Zellen in den Mikroskopbildern, der Identifizierung der strukturellen Bestandteile der einzelnen Neuronen und der Extraktion der markierten Proteine innerhalb der Neuriten der detektierten Zellen.

2.1.1. Fluoreszenzmikroskopiebilder neuronaler Zellen

Grundlage der fluoreszenzbasierten Analyse der neuronalen Zellen bilden zweidimensionale Bilder fluoreszenzmarkierter Neuronen. Die abgebildeten Zellen stammen aus primäre Neuronen des Hippocampus von E17 Mausembryonen (Stadium der embryonalen Mausentwicklung nach 17 Tagen). Die verwendeten Neuronen wurden für zwei Tage im Zellkulturmedium kultiviert und nach diesem Zeitraum mithilfe von Paraformaldehyd (PFA) fixiert und immunhistochemisch gefärbt. Die Kerne der Neuronen wurden mittels 4',6-diamidino-2-phenylindole (DAPI) [104] angefärbt. Über eine spezifische Antikörperfärbung wurden die zwei Proteine α -Tubulin, welches Bestandteil der Mikrotubuli ist, und das mRNA bindende Protein ZBP1 markiert (vgl. Abschnitt 1.1). Zur Markierung der fixierten Aktinfilamente wurde mittels Phalloidinfarbstoff (Toxin des Grünen Knollenblätterpilzes) das Protein F-Aktin markiert, was vornehmlich die Wachstumskronen und filopodienähnliche Ausläufer der Neuriten des Neurons färbt (vgl. 1.1). Die Akquise der Bilder erfolgte an einem Epifluoreszenzmikroskop (Nikon TE-2000) unter Verwendung eines Öbobjektivs mit einer 60-fachen Vergrößerung und einer numerische Apertur (Auflösungsvermögen) von 1,4. Die einzelnen Bilder weisen eine Größe von 1000×1000 Pixel mit einer physikalischen Größe von $191 \mu\text{m} \times 191 \mu\text{m}$, bei einer Auflösung von $0,191 \mu\text{m} \times 0,191 \mu\text{m}$ pro Pixel auf. Diese relativ hohe Auflösung der Bilder wurde verwendet, um die markierten Strukturen und Proteine weitestgehend vollständig sichtbar zu machen und eine exakte Lokalisierung zu ermöglichen. Die für die Analyse aufgenommenen Bilder bestehen aus 4 Einzelbildern (Kanälen) mit einer effektiven Bildtiefe von 12 Bit pro Pixel und Kanal. Jeder Kanal zeigt einen der fluoreszenzmarkierten Bestandteile des Neurons. Für die Analyse stehen 60 Bilder zur Verfügung, welche ein Neuron als auch Fragmente von toten oder nicht vollständig abgebildeten Zellen enthalten. Die einzelnen Neuronenbilder sind (nicht lückenlos aufeinanderfolgend) nummeriert, um im Ergebnisteil detaillierter auf die jeweiligen Neuronen eingehen zu können (vgl. hierzu 3.1.1). Abbildung. 2.1 zeigt die vier Kanäle des Bildes einer Mikroskopaufnahme, die die fluoreszenzmarkierten Bestandteile eines der untersuchten Neuronen abbilden.

Die entwickelte Methodik zur fluoreszenzbasierten Analyse der neuronalen Zellen besteht aus vier einzelnen Phasen, die unabhängig auf den zugrunde liegenden Fluoreszenzbildern ausgeführt werden. In Abbildung 2.2 sind diese Phasen dargestellt, die als Ergebnis das vollständig segmentierte Neuron, sowie die morphologischen Merkmale der Zelle und die extrahierten

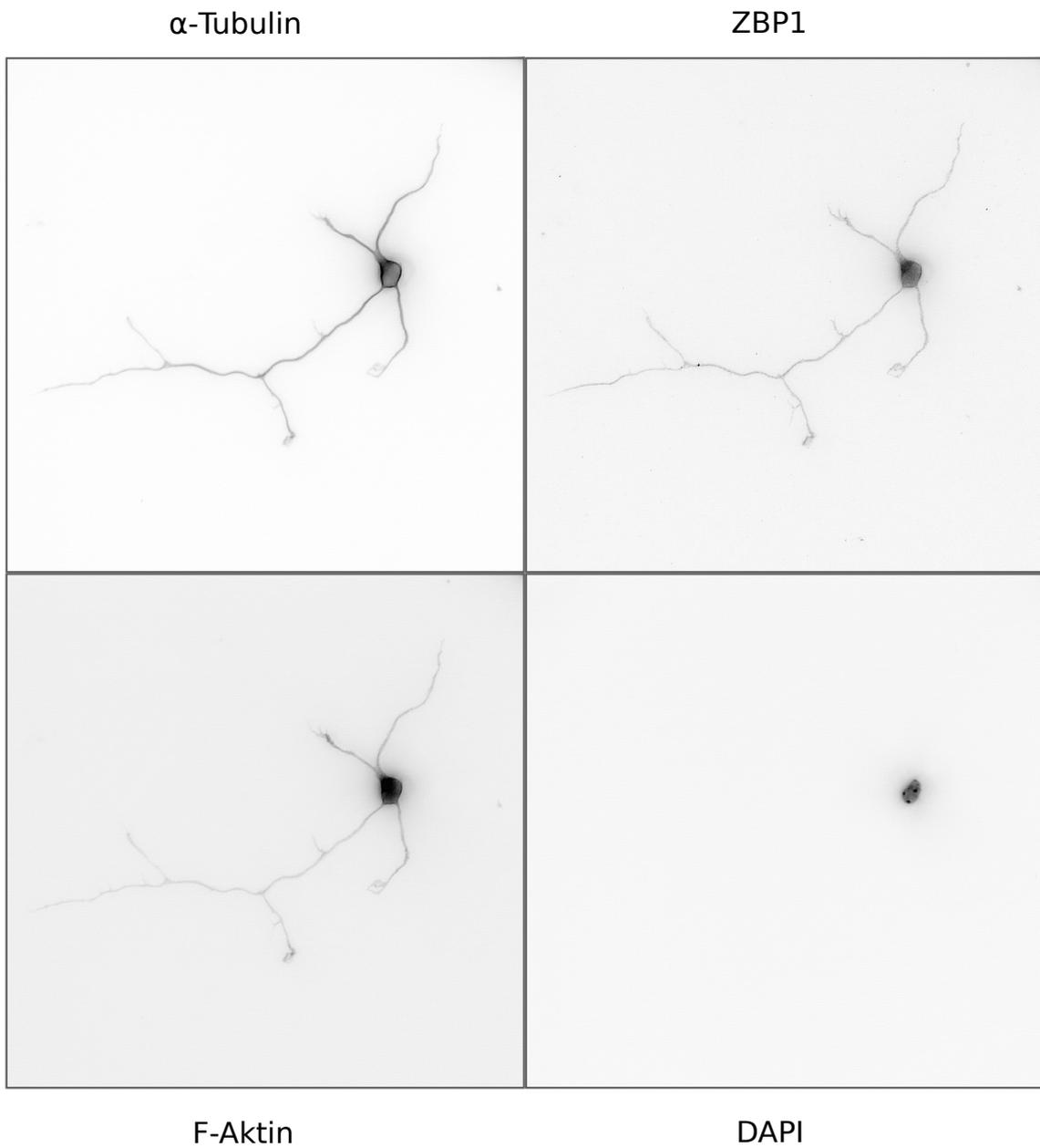


Abbildung 2.1.: Mikroskopaufnahme der 4 Kanäle eines fluoreszenzmarkierten primären hippocampalen Neurons. Die ersten drei Kanäle des Bildes zeigen die markierten Proteine α -Tubulin, ZBP1 und F-Aktin. Der vierte Kanal zeigt den mittels DAPI gefärbten Zellkern des Neurons. Zu einer verbesserten Darstellung wurde das Bild invertiert und die Helligkeit sowie der Kontrast angepasst.

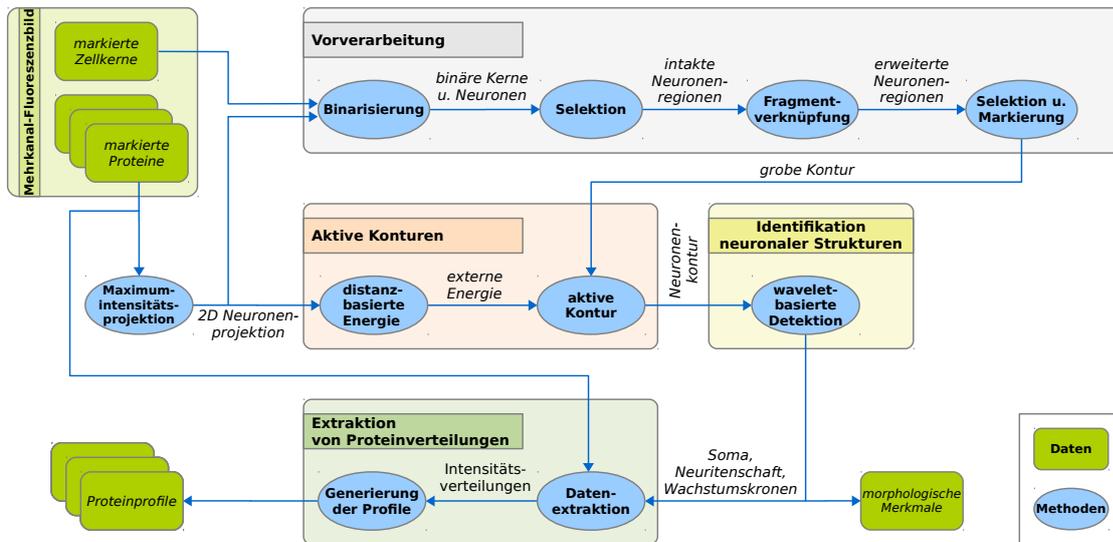


Abbildung 2.2.: Datenflussdiagramm der automatischen Neuronendetektion und Profilextraktion des entwickelten NeuronAnalyzer2D-Algorithmus. Die vier Phasen werden unabhängig auf einzelnen Kanälen der Mehrkanal-Fluoreszenzbildern ausgeführt und ergeben eine vollständige Segmentierung des Neurons, dessen morphologischen Merkmale und die extrahierten Proteinprofile.

Profile der markierten Proteine entlang der Neuriten liefern. Die hier dargestellte Methodik wurde in [163] beschrieben, und der entwickelte NeuronAnalyzer2D Algorithmus⁷ wird über die *Microscope Image Analysis Toolbox* MiToBo [168] zur Verfügung gestellt. Darüber hinaus kommen weitere in MiToBo implementierte Methoden der Bildverarbeitung zur Anwendung, wie beispielsweise die Maximalprojektion oder ein Algorithmus zur Regionenmarkierung.

2.1.2. Detektion neuronaler Zellen

Als Grundlage für die automatische Detektion der neuronalen Zellen dient die Maximalprojektion des Mehrkanalbildes. Diese Projektion wird aus den maximalen Intensitätswerten der einzelnen Kanäle der drei Proteinfärbungen generiert, um eine 2D Projektion der Neuronen zu erhalten. Diese Neuronenprojektion enthält mindestens ein vollständig abgebildetes Neuron, es können jedoch auch unvollständige Zellen oder Zellfragmente abgebildet sein. Anhand der Maximalprojektion und des Kanals der Zellkerne findet in den ersten zwei Phasen der Vorverarbeitung und der aktiven Konturen die vollständig automatische Detektion der neuronalen Zellen statt.

⁷<http://mitobo.informatik.uni-halle.de/index.php/Applications/NeuronAnalyzer2D>

Vorverarbeitung

In der ersten Phase der Vorverarbeitung werden die Zellkerne und Neuronen lokalisiert, um eine grober Kontur der strukturell intakten Neuronen im Bild zu bestimmen. Dabei werden mögliche durch die Fixierung vorhandene tote Zellen erkannt und nicht mit in die Analyse eingeschlossen. Der Erkennung von nicht intakten Zellen liegt die Annahme zugrunde, dass diese keinen strukturell intakten Zellkern mehr aufweisen. Kerne nicht intakter Zellen werden dabei im Kanal der Zellkerne durch zahlreiche nicht zusammenhängende kleine Regionen innerhalb der Zelle sichtbar. Als biologisch relevanter Parameter wird für die hier abgebildeten Neuronen eine minimale Fläche der Zellkerngröße mit $19,1 \mu\text{m}^2$ (100 Pixel) als Standardwert festgelegt. Die Lokalisierung der Zellkerne erfolgt durch das globale Schwellwertverfahren von Otsu [187]. Dadurch werden die Kerne im Zellkernkanal binarisiert und es erfolgt eine Segmentierung der Pixel in die einzelnen Zellkernregionen mithilfe eines sequentiellen Algorithmus zum *Component Labeling*, welcher in MiToBo enthalten ist. Für die Lokalisierung der Neuronen wird das Niblack-Schwellwertverfahren [179] auf der Grundlage des Bildes der Maximalprojektion verwendet. Dieses lokale Schwellwertverfahren arbeitet auf Teilbereichen des Bildes und schätzt den Mittelwert und die Standardabweichung innerhalb der Umgebung eines definierten Punktes im aktuellen Bildbereich. Durch vorausgegangene Studien hat sich gezeigt, dass die Betrachtung des gesamten Bildes als Nachbarschaft optimale Ergebnisse auf den Neuronenbildern liefert. Das Niblack-Verfahren wird somit als globale Schwellwertoperation zur Binarisierung des gesamten Bildes verwendet.

Anschließend findet eine Selektion der intakten Zellen statt, nicht intakte oder unvollständig abgebildeten Neuronen im Bild werden entfernt. Ein Neuron wird als intakt definiert, wenn es innerhalb der detektierten Zellregion mindestens 90% der Fläche genau eines Zellkerns enthält. Mögliche, durch das Niblackverfahren entstandenen, Fragmente in Form einzelner Regionen im Bild werden durch einen MiToBo basierten *Linking*-Algorithmus mit der intakten Neuronenregion verknüpft. Dabei werden die Fragmente, bestehend aus länglichen Regionen mit einem Exzentrizitätsmaß $\geq 0,5$ und einer Fläche von mindestens $2 \mu\text{m}^2$ (ca. 10 Pixel), die nicht mit der Kernregion überlappen und sich innerhalb einer definierten Distanz von ca. $10 \mu\text{m}$ (50 Pixel) befinden, als Kandidaten für die Verknüpfung verwendet (vgl. Abb. 2.3). Das hier angewendete Maß zur Exzentrizität-Bestimmung einer Region beruht auf der Verwendung zentraler Momente der Ordnung p, q , bei denen der Ursprung des Koordinatensystems in

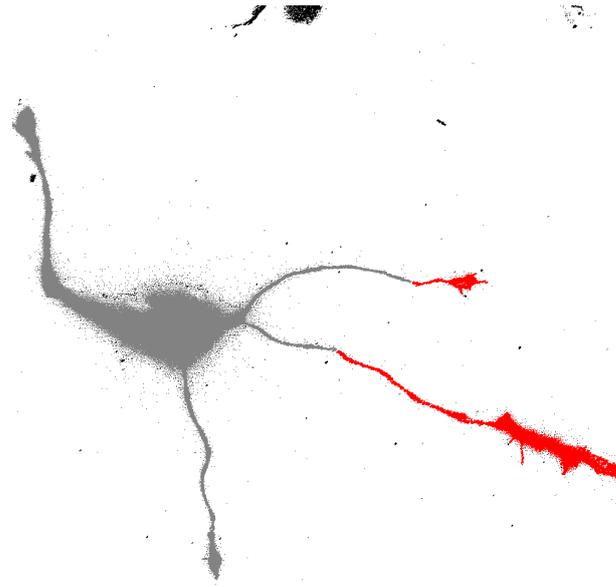


Abbildung 2.3.: Binarisiertes Bild einer neuronalen Zelle. Die Komponenten die nicht zur Region des Neurons gehören sind schwarz markiert, durch die Binarisierung entstandene Regionen von Zellfragmenten sind rot dargestellt. Identifizierte Kandidatenregionen (rot) werden basierend auf Distanzkriterien mit der Neuronenregion (grau) verknüpft.

den Schwerpunkt $\bar{s} = (\bar{x}, \bar{y})$ der betrachteten Fragmentregion verschoben wird. Das zentralen Moment μ der Ordnung p, q einer Region R des Bildes $I(u, v)$ ist definiert als

$$\mu_{p,q} = \sum_{(u,v) \in R} I(u, v) \cdot (u - \bar{x})^p \cdot (v - \bar{y})^q. \quad (2.1)$$

Das verwendete Exzentrizitätsmaß *Exzentrizität* (R) zur Bestimmung der länglichen Fragmentregionen hat die Form:

$$\text{Exzentrizität}(R) = \frac{(\mu_{20}(R) - \mu_{02}(R))^2 + 4 \cdot (\mu_{11}(R))^2}{(\mu_{20}(R) + \mu_{02}(R))^2}. \quad (2.2)$$

Die Werte des Exzentrizitätsmaßes liegen im Bereich $[0, 1]$, wobei eine runde Region einen Wert von 0, eine langgezogene Region einen Wert von 1 annimmt. Die Verknüpfung der einzelnen Kandidaten mit der Neuronenregion erfolgt mittels einer direkten linearen Verbindung zwischen den Pixeln der Fragmente und Neuronenregionen mit kleinster Distanz zueinander. Die Neuronenregion wird dadurch um die Regionen der Kandidaten erweitert und bildet die vollständigen Neuronenregion des betrachteten Bildes. Nachfolgend werden die Konturen dieser Neuronen, die als intakt und vollständig lokalisiert wurden, durch eine Konturverfolgung

extrahiert, um eine grobe Lokalisierung der vollständigen Neuronenregion zu erhalten. Diese Approximation ist jedoch nicht ausreichend, um genaue morphologische Merkmale zu ermitteln oder eine exakte Quantifizierung der markierten Proteine vorzunehmen. Dennoch eignet sie sich als Initialisierung der aktiven Konturen, die in der anschließenden zweiten Phase der automatischen Detektion zu einer Verbesserung der Kontur und zur genauen Lokalisierung der Neuronenregion verwendet werden.

Aktive Konturen (Snakes)

Die aktiven Konturen stellen verformbare Modelle dar, die durch Bildeinflüsse und externe Beschränkungen beeinflusst werden und mit deren Form und Lage ein energetischer Zustand verbunden ist. Das Ziel dieser Modelle ist die Detektion von Kanten dargestellter Bildobjekte mithilfe einer iterativen Minimierung eines gegebenen Energiefunktional. Anwendungen finden diese Modelle beispielsweise in der Segmentierung, Kantendetektion und Objektverfolgung [131, 225, 253]. Aktive Konturen können als implizites oder explizites Modell repräsentiert sein. Implizite Modelle verwendet sogenannte Level-Set-Methoden [35, 184], die keine Parametrisierung der Kontur benötigen oder Informationen über deren Topologie voraussetzen. In der hier vorgestellten Methode werden die aktiven Konturen durch ein explizites Modell repräsentiert [106]. Die Kontur liegt dabei als geschlossene parametrisierte Kurve $v(s)$ vor [217], die auf eine einheitliche Länge normalisiert ist:

$$v(s) = (x(s), y(s)), \text{ mit } s \in [0, 1]. \quad (2.3)$$

Durch die schrittweise Minimierung eines Energiefunktional erfolgt eine dynamische Bewegung der aktiven Kontur in Richtung der gesuchten Objektkontur. Wegen der Art wie die Konturen während der Minimierung „gleiten“, werden die aktiven Konturen auch als *Snakes* bezeichnet. Das zu minimierende Energiefunktional wird gesamtheitlich als Snake-Energie $E_{snake}(v(s))$ bezeichnet und umfasst interne und externe Energieterme, die die Bewegung der aktiven Kontur im Bild steuern. Die Snake-Energie ist definiert als die Summe der beiden unabhängigen Energien,

$$E_{snake}(v(s)) = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s)) ds. \quad (2.4)$$

Unter Verwendung des Verfahrens der *Variationsrechnung* wird mittels Gradientenabstieg eine Kontur $\tilde{v}(s)$ gesucht, die das Energiefunktional der hier verwendeten Snake-Energie (2.4) minimiert. Das Ergebnis der Variationsrechnung von (2.4) liefert, für die Betrachtung der x- und y-Richtung der aktiven Kontur, zwei unabhängige Euler-Lagrange-Gleichungen. Zur numerischen Lösung der Gleichungen werden diese diskretisiert, indem die aktive Kontur in äquidistante Stützstellen mit einem Abstand $h > 0$ eingeteilt wird, so dass eine Menge von Kontrollpunkten v_i mit $i = \{1, 2, \dots, n\}$ entsteht, die definiert sind als:

$$v_i = (x_i, y_i) = (x(i \cdot h), y(i \cdot h)) \quad (2.5)$$

und für die bei einer geschlossenen Kurve $v_1 = v_n$ gilt. Die durch das Minimierungsproblem gesuchte Snake-Energie sei im Folgenden als E_{snake}^* bezeichnet. Mithilfe der Diskretisierung lässt sich E_{snake}^* als Summe über den Energien an den Kontrollpunkte der Kurve darstellen. Daraus ergibt sich für die zu minimierende Snake-Energie die Gleichung:

$$E_{snake}^* = \sum_{i=1}^n E_{int}(v_i) + E_{ext}(v_i) . \quad (2.6)$$

Unter der Voraussetzung, dass die Kontrollpunkte äquidistant über die Kurve verteilt sind, lassen sich die Ableitungen an einer Stelle der Kurve mittels Differenzenquotienten approximieren. Unter Verwendung dieser Diskretisierung der Gleichungen und der Approximation der Ableitungen ergeben sich korrespondierende Euler-Lagrange-Gleichung an einem Kontrollpunkt v_i der aktiven Kontur. Diese Gleichungen lassen sich in Matrixform darstellen und bilden ein System aus zwei linearen Matrixgleichungen [106], mit der sich die genannten unabhängige Euler-Lagrange-Gleichungen für die x- und y-Richtung der aktiven Kontur lösen lassen:

$$Ax - f_x(x, y) = 0 , \quad (2.7)$$

$$Ay - f_y(x, y) = 0 . \quad (2.8)$$

Dabei bezeichnet A eine Pentadiagonalmatrix, welche die Koeffizienten der internen Energie der Kontrollpunkte beinhaltet, während die Vektoren f_x und f_y die externen Energien an den Kontrollpunkte der Snake enthalten. Um dieses System direkt lösen zu können, wird eine iterative Lösung zur Entwicklung einer aktiven Kontur mit minimaler Energie

angewendet (vgl. [106]). Die iterative Lösung erfolgt über eine Kombination aus explizitem und implizitem Euler-Verfahren, einem einfachen Verfahren zur Lösung von Anfangswertproblemen mittels Schrittweitenapproximation zur Annäherung an die exakte Lösung. Dabei wird des explizite Euler-Verfahren, zur Bestimmung der Lösung anhand der Werte zum vorherigen Zeitpunkt, auf die externe Energie angewendet, da diese innerhalb eines Zeitschrittes konstant ist. Das implizite Euler-Verfahren berechnet in jedem Schritt der Iteration die Lösung und wird bei der internen Energie angewendet, da die internen Energien der Kurve zum Zeitpunkt t bereits innerhalb der Koeffizientenmatrix A enthalten sind. Die Lösung des Optimierungsproblems resultiert somit in zwei Gleichungen, mit denen direkt die Koordinaten der Kontrollpunkte der aktiven Kontur zum Zeitpunkt $t + 1$ berechnet werden können. Die Koordinaten und Energien des vorherigen Zeitschrittes t dienen dabei als Berechnungsgrundlage. Dadurch ist die Snake neben dem Parameter s der parametrisierten Kurve auch von der Zeitvariablen t abhängig. Die Matrixgleichungen (2.7) und (2.8) können somit als ein Paar linearer Differentialgleichungen, mit konstanten Koeffizienten zum Zeitpunkt t , aufgefasst werden (vgl. [106]).

$$x_{t+1} = (A + \gamma I)^{-1} (\gamma x_t + f_x(x_t, y_t)) \quad (2.9)$$

$$y_{t+1} = (A + \gamma I)^{-1} (\gamma y_t + f_y(x_t, y_t)) \quad (2.10)$$

Die Vektoren x_t und y_t enthalten die x- und y-Koordinaten der Kontrollpunkte zum Zeitschritt t . A stellt die bekannte pentadiagonale Koeffizientenmatrix dar und die Vektoren f_x und f_y die externen Energien für jeden der Kontrollpunkte. I bezeichnet die Einheitsmatrix und der Parameter γ die Schrittweite der einzelnen Kontrollpunkte zwischen den Zeitpunkten t und $t + 1$. Die Schrittweite kontrolliert die Geschwindigkeit der Snake-Bewegung. Anhand der resultierenden linearen Gleichungen (2.9) und (2.10) können die Kontrollpunkte der gesuchten Kurve $\tilde{v}(s)$, die das Energiefunktional (vgl. (2.4)) minimiert, direkt berechnet werden.

Die in der vorliegenden Arbeit verwendete interne Energie $E_{int}(v(s))$ (vgl. (2.4)) leitet sich von der parametrisierten Kurve $(v(s))$ ab und umfasst die Summen der mit $\alpha(s)$ und $\beta(s)$ gewichteten ersten und zweiten Ableitungen der Kurve $(v'(s)$ und $v''(s))$ (vgl. [106]). Die interne Energie hat die Form:

$$E_{int}(v(s)) = \frac{1}{2} (\alpha(s) |v'(s)|^2 + \beta(s) |v''(s)|^2) . \quad (2.11)$$

Dabei beschreibt die erste Ableitung die Elastizität der Kontur, während die zweite Ableitung die Steifheit der Kontur beschreibt, wodurch die hier genutzte interne Energie einen Einfluss auf Krümmung der Kontur hat und deren Glattheit beeinflusst. Durch die Gewichtungen findet eine Balance zwischen Elastizität und Steifheit sowie zwischen der internen und externen Energie statt. Die externe Energie $E_{ext}(v(s))$ (vgl. (2.4)) wird anhand des Bildes definiert, um die aktive Kontur zu den gesuchten Bereichen im Bild zu führen, indem sich die Kontur in ein lokales Energieminimum bewegen kann. Häufig basieren verwendete externe Energien auf den Gradienten von Grauwertbildern, wie beispielsweise der negative Gradientenbetrag eines (gaußgefilterten) Bildes. Der Vorteil eines gaußgefilterten Bildes gegenüber dem ungefilterten Gradientenbild besteht darin, dass die Objektkonturen geglättet werden und der Gradient eine erhöhte Reichweite auf die aktive Kontur als bei ungefilterten Bildern aufweist. Dies wiederum erlaubt einen größeren Abstand der initialen aktiven Kontur zum gesuchten Objekt. Eine hauptsächliche Limitierung gradientenbasierter Energien bleibt jedoch die dennoch geringe Reichweite und die mangelhafte Anpassung der Kontur an die Objektgrenzen, z. B. bei auftretenden Konkavitäten an den Grenzen [274]. Sogenannte *Gradientenvektorfelder* (GVF) stellen eine alternative externe Energie auf Basis der Diffusion von Gradientenvektoren dar. Sie ermöglichen die Bewegung der aktiven Kontur in Konkavitäten hinein, wodurch eine verbesserte Angleichung an gekrümmte Objektgrenzen erzielt wird, und erlauben eine ungenauere Initialisierung, z. B. einer größeren Entfernung der initialen Kontur [162, 274]. Für die zu detektierenden Neuriten, als auch für andere tubuläre Objekte [9, 29, 36], kann das GVF zur Kontraktion der gegenüberliegenden Grenzen führen, indem ein Sprung der aktiven Kontur in die tubulären Strukturen erfolgt. Dies geschieht bei röhrenförmigen Objekt dadurch, dass die Gradientenvektoren, basierend auf der Verteilung der Gradienteninformation in Richtung der Mitte des Objektes, direkt auf die Mittellinie der tubuläre Objekte [9] und somit auch der Neuriten zeigen. Folglich sind diese alternativen Energien, basierend auf der Diffusion von Gradientenvektoren, nicht optimale im Sinne einer korrekten Lokalisierung der Kontur im Bereich der Neuriten und eignen sich nicht für die Analyse der vorliegenden Neuronenbilder, da sie falsche Ergebnisse bei der Detektion entlang der Neuritenregionen liefern (vgl. blaue Kontur in Abb. 2.4).

Um dieser Problematik zu begegnen und eine möglichst genaue Detektion zu erzielen, wurde im Rahmen dieser Arbeit eine Energie entwickelt, die, basierend auf einer Distanztransformation, eine große Entfernung der initialen Kontur als auch eine korrekte Lokalisierung der

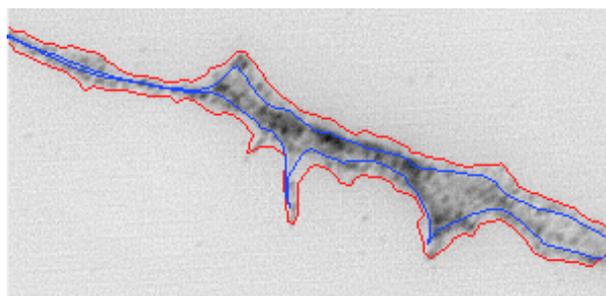


Abbildung 2.4.: Details einer detektierten Neuritenregion der Maximalprojektion unter Verwendung verschiedener externer Energien. Das Bild wurde zur besseren Veranschaulichung invertiert. In blau ist das Detektionsergebnis basierend auf dem Gradientenvektorfeld (GVF) dargestellt. Diese externe Energie bewirkt eine Bewegung der aktiven Kontur in das Innere der Neuritenregion und lässt die Kontur teilweise zu einer Linie kontrahieren. Unter Verwendung der distanzbasierten externen Energie (rot) erfolgt eine genaue Lokalisierung der Neuritengrenze.

Objektgrenzen ermöglicht [163]. Die externe Energie wird über eine Distanzkarte definiert, die sich aus der Berechnung der Distanzen der einzelnen Bildpixel zu den Pixeln der Objekte im Bild ergibt. Die aktive Kontur wird somit entlang der Distanzen der Transformation zu den Neuronen geführt. Es findet dabei keine Glättung der Objektkonturen statt, jedoch ist eine genaue (nicht notwendigerweise zusammenhängende) Lokalisierung der Objektgrenzen erforderlich, womit die Neuronenregionen der Vorverarbeitung zur Initialisierung der aktiven Konturen hier keine Anwendung finden. Um eine möglichst genaue Objektgrenze der Neuronenregion zu ermitteln, die als Grundlage für die Berechnung der Distanzkarte des Bildes dient, wird das Bild der Maximalprojektion mithilfe der nachfolgenden Schritte binarisiert:

1. Anwendung eines Medianfilters (5×5 Pixel),
2. Berechnung des Gradientenbetrags (zentrale Differenzen),
3. Anwendung eines Medianfilters (5×5 Pixel) auf den Gradientenbetrag,
4. Binarisierung des gefilterten Gradientenbetragbilds durch einen Otsu-Schwelldwert und
5. Verbesserung der binären Regionen durch morphologisches Closing (5×5 Pixel).

Hierbei wird entgegen der Vorverarbeitung das Schwellwertverfahren von Otsu für die Detektion der Neuronenregionen verwendet. Dadurch werden unter Umständen mehr Fragmente der Neuronenregionen generiert, jedoch erlaubt es eine bessere Lokalisierung der Neuronenkonturen. Entstandene Konturlücken zwischen den Fragmenten werden durch die Interpolation der Bereiche durch die aktive Kontur kompensiert. Die Verwendung des Otsu-Schwelldwertes erweist

sich in dieser Phase der Detektion als Vorteil und ermöglicht die Erzeugung einer präziseren externen Energie. Zur schnellen Berechnung der Distanzkarte wird der *Chamfer-Algorithmus* verwendet, welcher eine Approximation der Euklidischen-Distanztransformation darstellt [20]. Bei dieser Methode werden mithilfe einer Fenstermaske die Abstände von benachbarten Pixeln in zwei Bilddurchgängen, von der oberen linken Ecke zur unteren rechten Ecke und der zweite Durchgang von der unteren rechten Ecke zurück zur oberen linken Ecke, propagiert („*chamfered out*“). Im zweiten Durchlauf werden die aus dem ersten Durchlauf berechnete Distanzen ersetzt, so diese kleiner als zuvor berechnet sind. Dieser Fall tritt ein, wenn sich das Pixel näher am unteren oder rechten Rand des Objektes befindet. In dieser Arbeit wird für die angewendete Maske ein Fenster mit 3×3 Pixel und einer Euklidischen-Distanzmetrik verwendet. Die Distanzkarte wird für die Hintergrundregion des zuvor erstellten binären Bildes berechnet (siehe Schritte oben). Den Vordergrund der Distanztransformation stellt die binarisierte Neuronenregion dar, deren enthaltene Pixel eine Distanz von 0 in der Distanzkarte zugewiesen werden. Somit nehmen die Distanzen innerhalb der Neuronenregion einen Wert von Null an, außerhalb der Region nimmt die Distanz mit der Entfernung zur Neuronenregion zu, was zu einer erhöhten Reichweite des Einflussbereichs der Region auf die initiale aktive Kontur führt. Anhand des Vergleichs in Abb. 2.4 wird deutlich, dass die distanzbasierte Energie gegenüber dem Gradientenvektorfeld eine korrekte Lokalisierung der Objektgrenzen ermöglicht, was zu einem verbesserten Gesamtergebnis der Detektion führt.

Als Parameter für die analysierten Bilder dieser Arbeit wurden verschiedene Gewichte der internen Energie getestet und auf die Werte $\alpha(s) \equiv \alpha := 0,2$ und $\beta(s) \equiv \beta := 0,0$ festgelegt (vgl. (2.11)). Diese Werte erzielten gute Ergebnisse für die Konturbestimmung, indem die Kontur innerhalb von Neuritenverzweigungen bewegt werden kann und kleinere Unebenheiten entlang der Kontur überbrückt werden können. Das Festlegen von β auf den Wert 0 führt dazu, dass die interne Energie unabhängig von der zweiten Ableitung und somit der Steifheit der Kontur ist, was es ermöglicht, die aktive Kontur in Ecken bzw. Einbuchtungen der Objekte zu bewegen. Dies ist beispielsweise bei Verzweigungen der Neuriten notwendig. Die Schrittweite γ der Kontrollpunkte der aktiven Kontur aus (2.9) und (2.10) wurde initial für den Gradientenabstieg auf einen sehr kleinen Wert von 0,05 festgelegt und wird individuell für jeden einzelnen Punkt während der Minimierung adaptiert. Die zugrunde liegende Adaptierungsfunktion für einen Kontrollpunkte ist definiert als die Wurzel aus der Distanz an diesem Punkt in der Distanzkarte. Mithilfe der Adaptierungsfunktion wird jeder Kontrollpunkt durch eine energiebasierte Schrittweite bewegt,

was zu einem dynamischeren Verhalten der Kontur hingegen einer festen Schrittweite führt. Eine hoher Wert der externen Energie entspricht einer großen Distanz zur Neuronenregion, was zu einem hohen Wert der Schrittweite γ führt. Eine geringe Schrittweite dagegen stimmt mit einem geringen Energiewert überein. Anhand der hier definierten externen Energie sind die Distanzen innerhalb der Neuronenregion auf 0 gesetzt. Durch die interne Energie ist es jedoch möglich, dass die aktive Kontur sich dennoch in diese Region bewegt. Die so gewählte Adaptierungsfunktion hat zur Folge, dass die aktive Kontur nicht über die Grenzen hinweg geführt wird, da sich die Schrittweite automatisch in der Nähe der Regionengrenzen verringert. Die Terminierung des gesamten Minimierungsprozesses erfolgt, wenn weniger als 5% der Kontrollpunkte der aktiven Kontur innerhalb einer Iteration bewegt werden, oder maximal 120 Iterationsschritte ausgeführt wurden. Die Minimierung kann für mehrere aktive Konturen durchgeführt werden, wodurch die Konturen verschiedener Neuronen im Bild ermittelt werden können. Das Ergebnis des Optimierungsprozesses der aktiven Konturen liefert die finalen Konturen der automatisch detektierten Neuronen und dient als Grundlage für die Identifizierung der neuronalen Strukturen in der dritten Phase der fluoreszenzbasierten Analyse.

2.1.3. Identifizierung neuronaler Strukturen

Grundsätzlich lassen sich neuronale Zellen in zwei wesentliche strukturelle Teile gliedern, das Soma (auch Zellkörper oder Perikaryon genannt) und die einzelnen Ausläufer, die Neuriten. Ein vollständiger Neurit erstreckt sich dabei von seinem Ursprung an der Grenze des Somas über den Neuritenschaft bis hin zur Wachstumskrone (vgl. hierzu 1.1). Ziel der dritten Phase der fluoreszenzbasierten Analyse ist es, diese strukturellen Teile des Neurons, insbesondere der Neuriten, zu identifizieren.

Basierend auf der segmentierten Neuronenregion wird dessen Skelett extrahiert [285], aus welchem ein in dieser Arbeit entwickelter gerichteter azyklischer Graph aufgebaut wird (Skelettgraph). Die Verzweigungs- und Endpunkte des Skeletts repräsentieren die Knoten des Graphen, während die Kanten durch die Skelettsegmente zwischen zwei Knoten repräsentiert werden. Diese Kanten bestehen aus einer sortierten Liste von Skelettpixeln, deren Richtung beim Aufbau des Graphen von einem Start- zu einem Endpunkt vorgegeben ist. Der Aufbau des Skelettgraphen erlaubt eine geeignete Behandlung des Skeletts für spätere Schritte. Die Region des Somas weisen einen großen Bereich des Zellkörpers auf, aus dem die vergleichsweise dünnen Neuriten heraustreten. Um das Soma grob zu lokalisieren, wurde ein morphologisches Opening

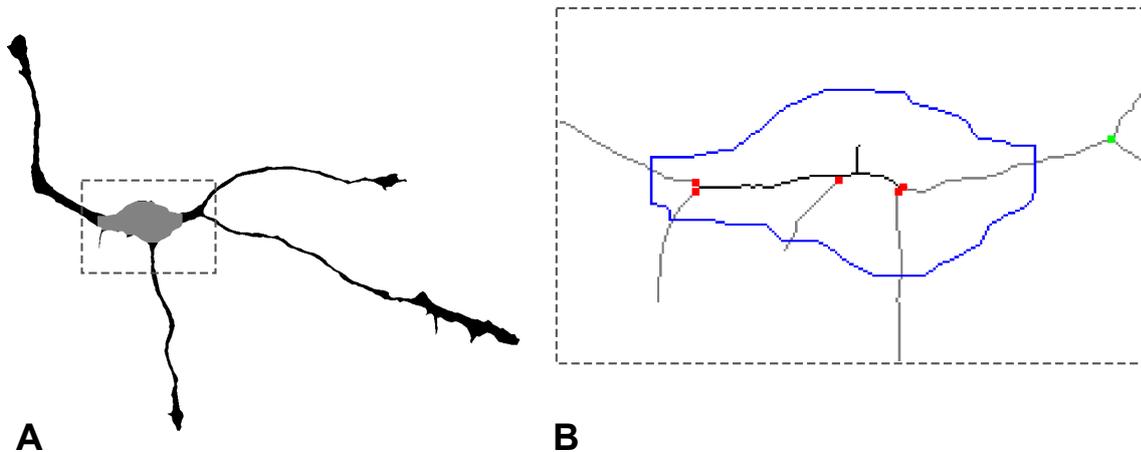


Abbildung 2.5.: Identifizierung des Somas und der Neuriten als strukturelle Bereiche des Neurons. (A) Die Somaregion (grau) des detektierten Neurons (schwarz) wird durch morphologische Opening grob lokalisiert. (B) Vergrößerung des Ausschnitts der Somaregion mit enthaltenem Skelettgraph der Neuronenregion. Der Bereich des Somas ist mit einem blauen Polygon markiert, während die Kanten des Graphen als graue oder schwarze Linien dargestellt sind. Kanten des Skelettgraphen, die ausschließlich innerhalb des Somas liegen (schwarz), werden aus dem Skelettgraphen entfernt. Kanten des Graphen, die die Grenze des Somas schneiden, werden zur Zerlegung des Graphen in die einzelnen Neuritenskelettgraphen (grau) verwendet. Die somit entstehenden Neuritenskelettgraphen repräsentieren die einzelnen Neuriten des Neurons. Rote Knoten kennzeichnen die Startknoten des gerichteten azyklischen Neuritenskelettgraphen, während grüne Knoten deren Verzweigungspunkt markieren.

mittels eines großen strukturierten Elements von $110 \times 110 \mu\text{m}$ (21×21 Pixel) angewendet. Dadurch wird der große Bereich des Somas detektiert, während die Bereiche der Neuriten von der Detektion ausgeschlossen werden (vgl. Abb. 2.5 (A)). Anhand der so ermittelten Region des Somas werden anschließend die Neuritenursprünge an der Grenze zum Soma lokalisiert. Dafür wird der erzeugte Skelettgraph in mehrere Teilgraphen zerlegt. Weist der Skelettgraph innerhalb des Somas mehrere Verzweigungspunkte auf (vgl. Abb. 2.5 (B), rote Knoten), bilden diese die Startpunkte zum Aufbau neuer gerichteter azyklischer Teilskelettgraphen der einzelnen Neuriten, welche im Folgenden als *Neuritenskelettgraphen* bezeichnet werden. Dabei werden nur Teilskelettgraphen verwendet, die eine Kante aufweisen, welche die Grenze des Somas schneidet (vgl. Abb. 2.5 (B)). Tritt der Fall ein, dass kein Verzweigungspunkt innerhalb der Somaregion vorhanden ist, wird der Graph an den Stellen geteilt an denen die Graphkante die Grenze der Somaregion schneidet. Die Schnittpunkte der Kante mit dem Soma bilden dabei die Startpunkte für den Aufbau der beiden Neuritenskelettgraphen. Dieser Fall tritt genau dann ein, wenn höchstens zwei Neuriten vorhanden sind und es somit nicht zu einer Verzweigung des Graphen innerhalb des Somas kommt. Die anhand der Zerlegung erzeugten

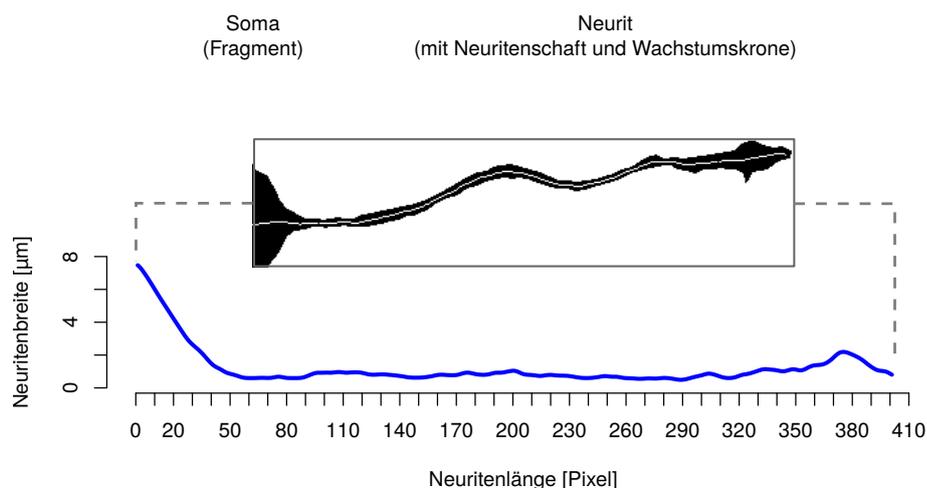


Abbildung 2.6.: Binäres Bild eines Neuriten (schwarz) mit dazugehörigem Breitenprofil (blau) entlang des erzeugten Neuritenskelettgraphen (graue Linie im Neuriten) vom Soma zur Wachstumskrone.

Neuritenskelettgraphen repräsentieren die einzelnen Neuriten und bilden die Basis dafür, die Anfangs- und Endpunkte der Neuriten zu lokalisieren. Um diese Punkte möglichst exakt zu bestimmen, ist es nötig die Grenzen zwischen dem Soma und dem Neuritenursprung als auch zwischen Neuritenschaft und Wachstumskrone zu detektieren. Die einzelnen strukturellen Teile unterscheiden sich stark in ihrer Morphologie (vgl. Abb. 2.6). So ist beispielsweise das Soma durch eine breite kompakte Region gekennzeichnet, gefolgt von einem lang gezogenen eher schmalen Neuriten, der wiederum aus einem dünnen Schaft und der endständig verdickten Wachstumskrone besteht. Die Breite der einzelnen Strukturen scheint somit ein adäquates Maß zu sein, um die neuronalen Strukturen über ihre morphologische Charakteristik zu unterscheiden. Entlang des Neuritenskelettgraphen wird die Breite der Neuronenregion als eindimensionale Funktion angesehen und im Folgenden als *Breitenprofil* bezeichnet. Die zugehörige Breite der Region entlang des Neuritenskelettgraphen wird durch eine Distanztransformation auf der segmentierten Neuronenregion ermittelt. Abbildung 2.6 zeigt das Breitenprofil zu einem binarisierten Neuriten. Das Profil weist an den Übergängen der strukturellen Teile einen steileren Anstieg auf als beispielsweise innerhalb des Neuritenschafts. Entsprechend dieser Eigenschaft wird das Breitenprofil dazu verwendet, die Grenzen zwischen den strukturellen Teile des Neurons und insbesondere die einzelnen Neuriten zu identifizieren. Um diese Identifizierung automatisch anhand das Breitenprofils zu ermöglichen, wurde ein Ansatz, basierend auf einer Wavelet-Transformation, adaptiert.

Wavelet-Transformation

Wavelet-Transformationen (WT) sind Zeit-Frequenz-Transformationen und finden Anwendung bei der Analyse von Signalstrukturen, bei der die Signale mittels sogenannten Wavelets zerlegt und skaliert werden. Zur Analyse eines Signals $f(t)$ wird eine Familie von normalisierten Wavelet-Funktionen genutzt, die dadurch entsteht, dass die Wavelet-Funktion $\psi(t)$, oft als *Mutter-Wavelet* bezeichnet, mit dem Parameter s und τ skaliert und verschoben wird:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t - \tau}{s}\right), \quad s \in \mathbb{R}^+, \tau \in \mathbb{R}. \quad (2.12)$$

Die Wavelet-Transformation des Signals $f(t)$ ist definiert als:

$$W_f(\tau, s) = \int_{-\infty}^{\infty} f(t) \psi_{\tau,s}(t) dt, \quad (2.13)$$

woraus eine Matrix mit den Wavelet-Koeffizienten W_f für diskrete Werte von s und τ resultiert. Die Koeffizienten beschreiben über die Skale s und die Verschiebung τ die Frequenzanteile von $f(t)$. Die variablen Breite des Abfragefensters wächst proportional zum Betrag der Skale, wodurch sowohl hohe als auch niedrige Frequenzanteile in der Analyse erfasst werden [50, 146, 245].

Ein häufige verwendetes Mutter-Wavelet ist das, aufgrund seiner Form, als mexikanischer Hut benannte, *Mexican-Hat-Wavelet*. Dieses Wavelet entsteht durch die normalisierte 2. Ableitung der Gaußfunktion:

$$\psi(t) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} (1 - t^2) e^{-\frac{t^2}{2}}. \quad (2.14)$$

Unter Einsetzen von (2.14) in Gleichung (2.12) wird eine Familie von skalierten und translatierten Mexican-Hat-Wavelets erzeugt. Als Beispiel zeigt Abb. 2.7 drei Wavelets für verschiedene Skalen bei einer festen Verschiebung.

Bei Wavelet-Transformationen können unterschiedliche Transformationsmethoden zur Anwendung kommen. Die als kontinuierlichen Wavelet-Transformation bezeichnete Variante (*Continuous Wavelet Transform, CWT*) ermöglicht Transformationen über allen Skalen hinweg ($s > 0$) mit einer kontinuierlichen Verschiebung. Dadurch entstehende Redundanzen stellen z. B. Informationen über die Form und Zusammensetzung von Peaks in Massenspektrometrie-Daten besser dar und machen diese leichter interpretierbar, als bei Anwendung diskrete

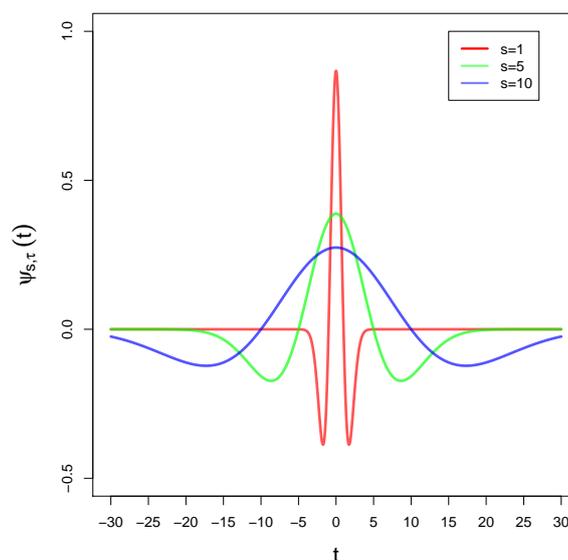


Abbildung 2.7.: Mexican-Hat-Wavelet verschiedener Skalierungen $s = (1, 5, 10)$ und bei einer Verschiebung von $\tau = 0$ im Intervall $t = [-30, 30]$

Wavelet-Transformation (*Discrete Wavelet Transform, DWT*). Bei der DWT wird der Skalierungsparameter zu ganzzahligen Potenzen von 2 diskretisiert, was eine nicht redundante und effiziente Auswertung der Signale ermöglicht, wie sie beispielsweise in der Datenkomprimierung und Merkmalsextraktion angewendet wird [58]. Die DWT erlaubt es anhand einer Teilmenge der Koeffizienten das ursprüngliche Signal zu erfassen und somit die Datenmenge zu komprimieren. Die Anzahl der verwendeten Koeffizienten entspricht der Menge der Koeffizienten des Signals. Da jedoch das Rauschen in vielen kleinen Koeffizienten resultiert, können diese verworfen werden und dennoch eine qualitativ hochwertige Signalapproximation liefern.

Identifikation der Strukturen

Der adaptierte Ansatz zur automatische Identifizierung der strukturellen Teile anhand des Breitenprofils ist vom *centWave* Algorithmus abgeleitet, einem Algorithmus zur Detektion von charakteristischen Merkmalen in Massenspektrometrie-Daten [245]. Im Speziellen wird darauf abgezielt, chromatographische Maxima, sogenannte *Peaks*, und deren Grenzen basierend auf einem 2D Signal von LC/MS Daten (Flüssigchromatographie mit Massenspektrometrie-Kopplung) zu erkennen [117, 176, 245]. Ziel von *centWave* ist es, die Peaks verschiedener Breite innerhalb des Chromatographiesignals zu detektieren. Diese erfolgt in zwei Schritten,

erstens der Identifizierung von möglichen Peakregionen aus dem Massensignal und zweitens die genaue Analyse der Signalintensitäten dieser Regionen unter Verwendung der kontinuierlichen Wavelet-Transformation (CWT). Da nicht von einer konstanten Breite der Peaks ausgegangen wird, können durch die verschiedenen Skalierungen, die die CWT ermöglicht, Peaks mit einer hohen Variabilität in der Breite als auch aneinandergrenzende Peaks detektiert werden. Das bereits erwähnte Mexican-Hat-Wavelet wird im centWave Algorithmus als Modell-Peak verwendet. Für die automatische Identifizierung der neuronalen Strukturen in dieser Arbeit wird das Eingabechromatogramm des Algorithmus durch das ermittelte Breitenprofil des Neuritenskelettgraphen der einzelnen Neuriten ersetzt. Infolgedessen wird die Intensität des Massesignals über den zeitlichen Verlauf durch die Breite des Neuriten entlang der Länge vom Soma zur Wachstumskrone ersetzt.

Betrachtet man die Veränderungen im Breitenprofil, so ist erkennbar, dass zwei Peaks das Soma und die Wachstumskrone kennzeichnen (vgl. Abb. 2.6). Um deren Grenzen zu lokalisieren und die strukturellen Teile voneinander zu trennen, wird nach Positionen an den Rändern des Breitenprofils gesucht, in denen sich der Anstieg im Profil deutlich verändert. Für die Identifizierung der Strukturen werden hierfür nur Breitenprofile verwendet, die eine Mindestlänge von ca. $19 \mu\text{m}$ (100 Pixel) aufweisen. Kürzere Profile könnten sogenannte filopodienähnliche Vorstülpungen der Zelloberfläche kennzeichnen. Da jedoch nur charakteristisch eindeutige Neuriten analysiert werden sollen, werden Profile kürzerer Strukturen nicht in die Identifizierung der Strukturen und der nachfolgenden Extraktion der Proteinprofile einbezogen. Die verwendeten Breitenprofile werden mithilfe eines eindimensionalen Gauß bei einem Sigma von 2 geglättet. Um eine Abschätzung der Basislinie der Breitenprofile im centWave Algorithmus zu verbessern, wurde die Profiliränder erweitert und das Profil in zwei Teile zerlegt. Die Erweiterung der Ränder erfolgte durch eine Spiegelung der 20 ersten und letzten Werte des Breitenprofils. Die Zerlegung erfolgte durch die Aufteilung des Profils in zwei möglichst gleich große Teilprofile. Beide Teilprofile werden anschließend mittels centWave analysiert. Aus der daraus resultierenden Koeffizientenmatrix der Wavelet-Transformation lassen sich die enthaltenen Peaks im Profil bestimmen [58, 245]. Die einzelnen Koeffizienten für jede Skale weisen dabei ein lokales Maximum um die Peakmitte herum auf. Basierend auf diesen lokalen Maxima der einzelnen Skalen wird der Bereich der Skalen definiert, in denen der Peak detektiert wurde. Der Peakmittelpunkt wird anschließend durch den maximalen Wavelet-Koeffizienten innerhalb dieses Skalenbereiches bestimmt. Die lokalen

Koeffizienten der dazugehörigen Skale werden genutzt, um mittels beidseitigem Abstieg auf den Koeffizienten lokale Minima zu identifizieren, die die linke und rechte Grenze des Peaks darstellen [245]. Die Analyse eines Breitenprofils resultiert damit in einer Menge von Peaks, die jeweils eine „Ausbuchtung“ im Breitenprofil entlang des Neuriten kennzeichnen.

Abbildung 2.8 zeigt die centWave basierte Peakdetektion am Beispiel eines langen Neuriten von fast 670 Pixel (ca. 128 μm) mit drei erkennbaren Peaks im Breitenprofil. Abbildung 2.8 (A) zeigt die graphische Darstellung der Koeffizientenmatrix der Wavelet-Transformation in Falschfarben. Koeffizienten mit hohen Intensitäten sind durch weiße, Koeffizienten mit niedriger Intensität durch grüne Bereiche gekennzeichnet. Anhand der Matrix über die verwendeten Skalen von 10 bis 25 sind drei Peaks im Profil erkennbar, wobei sich zwei stärkere Peaks an den Rändern und ein schwächer ausgeprägter Peak vor dem Peak des rechten Randes befindet. Die minimale und maximale Skale wird anhand der vom Nutzer festgelegten Peakbreite (in Pixel) in einem vorgegebenem Intervall der zu erwartenden Breite [$Peak_{min}, Peak_{max}$] ermittelt. Anhand der Basisbreite des Mexican-Hat-Wavelets von $2s$, bei einer Skalierung von s (vgl. Abb. 2.7), lässt sich aus der Peakbreite der entsprechende Skalenbereich der Wavelet-Transformation ableiten:

$$Skale_{min} = Peak_{min} / 2, \quad (2.15)$$

$$Skale_{max} = Peak_{max} / 2. \quad (2.16)$$

Da in aktuell verfügbarer Literatur keine durchschnittliche oder approximative Breite bzw. ein Durchmesser für die Neuritenursprünge oder Wachstumskronen hippocampaler Mausneuronen ermittelt werden konnte, wurden die ermittelten Breitenprofile der hier verwendeten Bilder zur Abschätzung der Breite genutzt. Für die untersuchten Neuriten wurde Breiten im Bereich von 20 bis 50 Pixel (ca. 4 bis 10 μm) für die Bereiche des Neuritenursprungs bzw. der Wachstumskronen im Breitenprofil gemessen, was nach (2.15) und (2.16) zu dem genannten Skalenbereich von 10 bis 25 führt. Neben dem Mittelpunkt und den Grenzen des lokalisierten Peaks ist es in centWave möglich eine Gauß-Funktion an den Peak anzupassen, um Informationen über die Form und die Qualität des Peaks zu erhalten. Für die automatische Identifizierung des Somas und der Wachstumskrone sind die äußeren Peaks an den Enden des Breitenprofils von Interesse. Die rechte Grenze B_N des linken Peaks identifiziert den Ursprung des Neuriten, während die linke Grenze B_W des rechten Peaks den Übergang vom Neuritenschaft zur Wachstumskrone kennzeichnet (vgl. hierzu Abb. 2.9). Die Positionen der Peakgrenzen werden

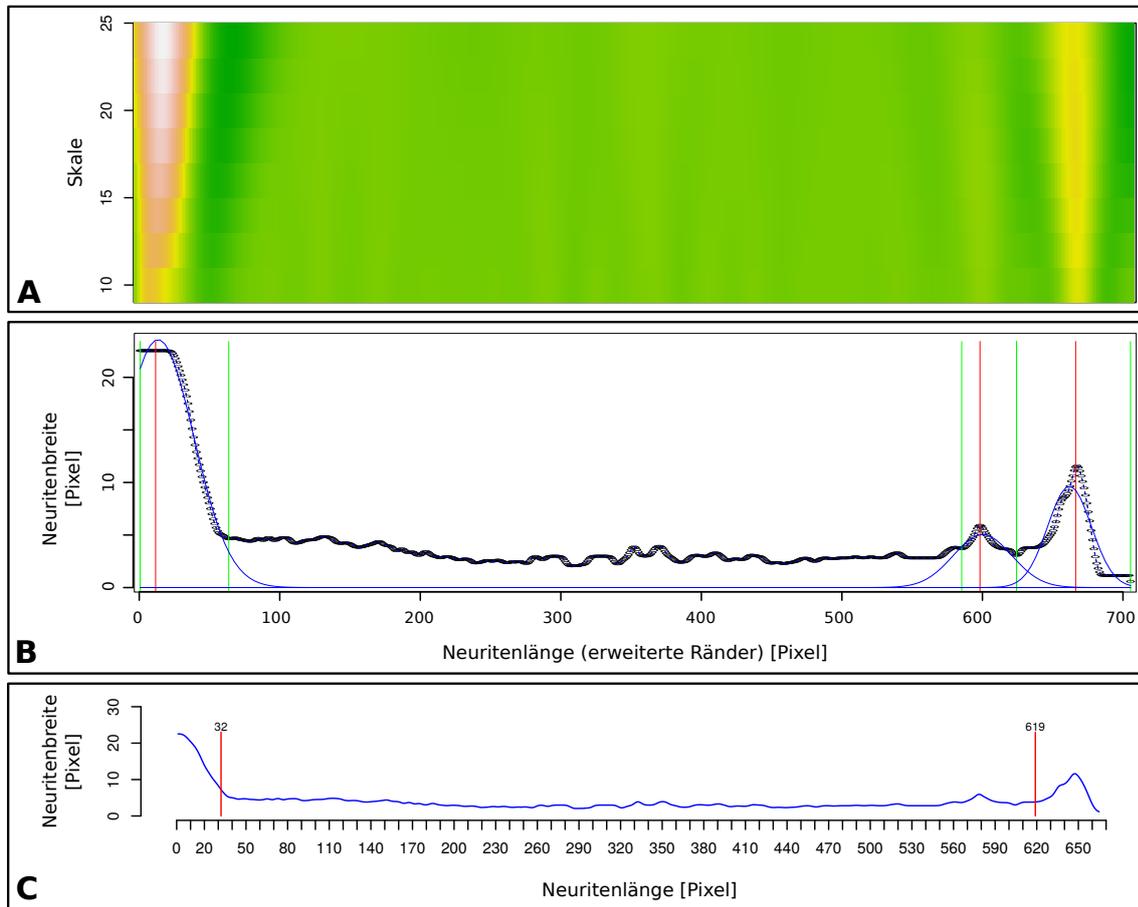


Abbildung 2.8.: Peakdetektion im Breitenprofil eines Neuriten basierend auf dem *contWave* Algorithmus unter Verwendung der kontinuierlichen Wavelet-Transformation. (A) Falschfarbendarstellung der Wavelet-Koeffizientenmatrix für die verwendeten Skalen von 10 bis 25. Die Farben kennzeichnen die Intensität der Koeffizienten, wobei grün für eine niedrige und weiß für eine hohe Intensität steht. (B) Geglättetes Breitenprofil des Neuriten (schwarz) mit den drei detektierten Peakmittelpunkten (rot) und den Peakgrenzen (grün). An die detektierten Peaks angepasste Gauß-Funktionen sind in blau dargestellt. (C) Geglättetes Breitenprofil (blau) mit den ermittelten Grenzen (rot) für das Soma (bei einer Länge von 32 Pixel) und der Wachstumskrone (bei einer Länge von 619 Pixel).

berechnet durch:

$$B_N = C_{links} + 1,5 * \sigma , \quad (2.17)$$

$$B_W = C_{rechts} - 1,5 * \sigma . \quad (2.18)$$

Die Positionen der Mittelpunkte des linken und rechten Peaks sind durch C_{links} und C_{rechts} gekennzeichnet, wobei σ den Parameter der Gauß-Funktion durch den Peak beschreibt und aus der Information über den Gauß in centWave bestimmt wird. Ein Wert von $\pm 1,5 * \sigma$ zur Festlegung der Grenzpositionen hat sich dabei als geeignet erwiesen. Durch den Faktor von 1,5 wird vermieden, dass die Grenzen zu weit innerhalb des Somas oder der Wachstumskrone gesetzt werden, was bedingt durch den sich stark verändernden Breitenunterschied beim Übergang

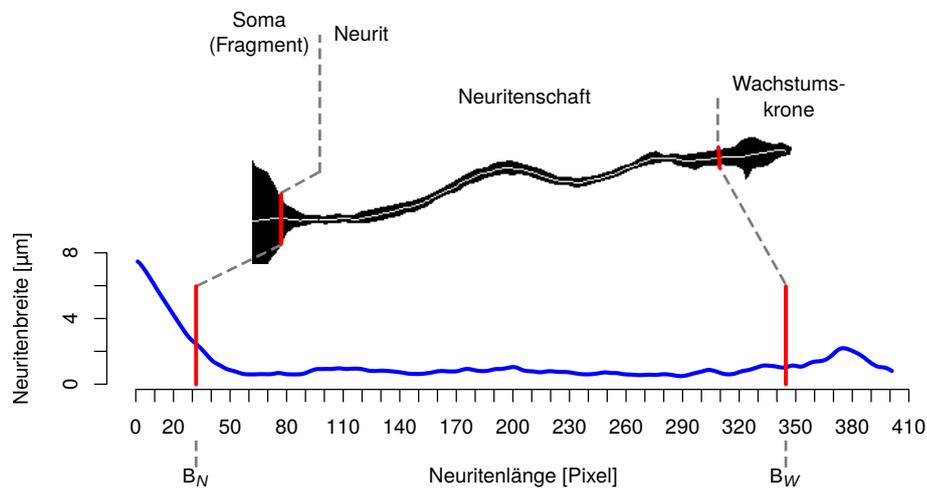


Abbildung 2.9.: Binäres Bild eines Neuriten (schwarz) mit dazugehörigem Breitenprofil (blau) entlang des Neuritenskelettgraphen (graue Linie im Neuriten) vom Soma zur Wachstumskrone. Die Grenzen der identifizierten Neuronenteile sind in rot dargestellt und trennen das Soma vom Neuriten und den Neuritenschaft von der Wachstumskrone.

dieser Strukturen zu den Neuriten auftritt. Basierend auf diesen ermittelten Grenzen wird die segmentierte Neuronenregion anschließend in ihre strukturellen Teile zerlegt, wenn sowohl eine Grenze für das Soma als auch für die Wachstumskrone vorliegen. Abbildung. 2.10 zeigt die Zerlegung eines Neurons in die Somaregion und drei Neuriten mit ihrem Neuritenschaft und den Wachstumskronen. Der rechte Neurit weist als Besonderheit eine Verzweigung des Neuriten auf. Die damit identifizierten neuronalen Strukturen dienen nachfolgend als Grundlage für die Extraktion der Proteinprofile entlang des Neuritenschafts und der Wachstumskrone.

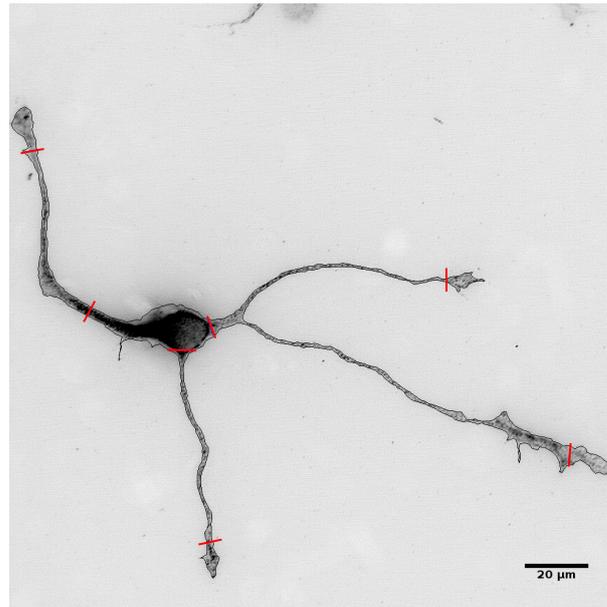


Abbildung 2.10.: Segmentiertes Neuron mit den identifizierten Grenzen (rot) zwischen den verschiedenen strukturellen Teilen des Neurons. Das Bild wurde zu verbesserten Darstellung invertiert und die Helligkeit sowie der Kontrast angepasst.

2.1.4. Extraktion von Proteinverteilungen

In der vierten und letzten Phase der fluoreszenzbasierten Analyse werden anhand der Fluoreszenzintensitäten der markierten Proteine die Proteinprofile extrahiert. Durch diese Extraktion der Fluoreszenzintensität entlang des lokalisierten Neuriten wird das Verteilungsprofil eines analysierten Proteins quantitativ bemessen. Die Intensitätsverteilungen der Proteine werden dabei entlang des Neuritenschafts bis zur Spitze der Wachstumskrone auf den zuvor erzeugten Neuritenskelettgraphen (vgl. 2.1.3) ermittelt. Besteht ein Neurit aufgrund von Verzweigungen des Neuritenschafts aus mehreren Neuritenskelettgraphen, sogenannten Ästen oder engl. *Branches*, so wird zur quantitativen Bestimmung der Verteilung der längste Neuritenskelettgraph dieses Neuriten verwendet. Um die Intensitätsverteilung innerhalb des Neuriten zu bestimmen, wird jedes Pixel der Neuritenregion mithilfe einer euklidischen Distanzmetrik dem nächsten Skelettpixel zugeordnet. Dieses Vorgehen entspricht einer Voronoi-Zerlegung [4] der Region des Neuriten, wobei die Mittelpunkte der Voronoi-Regionen durch die Skelettpunkte des Neuritenskeletts gegeben sind (vgl. Abb. 2.11). Basierend auf der durchschnittlichen Fluoreszenzintensität der Voronoi-Regionen wird anschließend die Intensitätsverteilung des Proteins entlang des Neuriten extrahiert. Für die identifizierten Neuriten des segmentierten Neurons ergibt sich daraus eine Menge von Intensitätsverteilungen

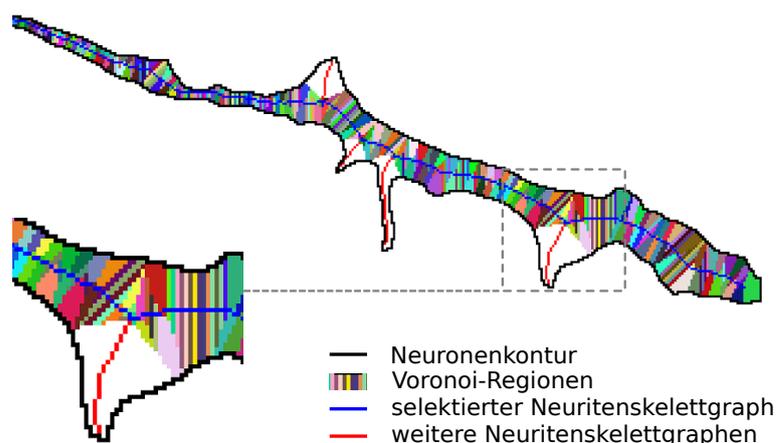


Abbildung 2.11.: Darstellung der Voronoi-Zerlegung der Neuritenregion mithilfe einer euklidischen Distanzmetrik. Anhand der erstellten Voronoi-Regionen wird aus der Fluoreszenzintensität entlang des Neuriten die Intensitätsverteilung des betrachteten Proteins extrahiert. Der markierte Bereich zeigt einen vergrößerten Ausschnitt der Voronoi-Regionen im Neuriten.

für jedes analysierte Protein. Um die unterschiedlichen Längen der Neuriten zu berücksichtigen, wird jeder Neurit in 100 Segmente zerlegt und die Intensitäten innerhalb der Segmente aufsummiert. Für jedes Protein wird pro Segment der Median der Summen in den Segmenten über alle Neuriten berechnet. Die somit entstandene Intensitätsverteilung aus den Medianen jedes Segments eliminiert Variationen zwischen den Intensitätsverteilungen der einzelnen Neuriten und Bilder, um eine für das analysierte Protein charakteristische Verteilung zu erhalten. Neben den Variationen der Fluoreszenzintensität in jedem einzelnen Neuriten weisen die Intensitätsverteilungen häufig ein Rauschen, bedingt z. B. durch die Proteinmenge, auf. Um dem entgegenzuwirken, werden die Intensitätsverteilungen zur Extraktion des finalen *Proteinprofils* approximiert. Die Approximation erfolgt mithilfe einer angepassten polynomialen Kurve unter Verwendung der in R [201] implementierten NLS-Methode (*Nonlinear Least Squares*) [8]. Nach genauerer Betrachtung der Approximationen haben sich hierfür Polynome 2. Ordnung als geeignete Approximation der Intensitätsverteilungen erwiesen. Abbildung 2.12 zeigt exemplarisch die Approximation einer Intensitätsverteilung durch Polynome verschiedener Ordnungen. Polynome höherer Ordnung, deren Ordnung über den Grad 2 hinausgeht, neigen zu einer Überanpassung (*Overfitting*) der approximierten Intensitätsverteilung. Des Weiteren erschweren Polynome höherer Ordnung, durch Schwankungen in der Intensitätsverteilung, eine biologisch eingängige Interpretation des Verlaufes des Proteinprofils entlang der Neuriten. Das heißt es sollen allgemeine Aussagen gemacht werden können, wie beispielsweise, dass eine

Abnahme der Proteinkonzentration vom Soma zur Wachstrumskrone entlang des Neuritenschaftes erfolgt und eine leichte Zunahme des Proteins in der Wachstumskrone verzeichnet werden kann. Schwankungen durch die Approximation höheren Grades, die z. B. innerhalb der Struktur des Neuritenschaftes eine Zunahme mit anschließender Abnahme der Konzentration erkennen lassen, bedürfen einer komplexeren Interpretation und genauere biologischen Untersuchung des sich innerhalb des Neuritenschaftes verändernden Verlaufes (vgl. Abb. 2.12).

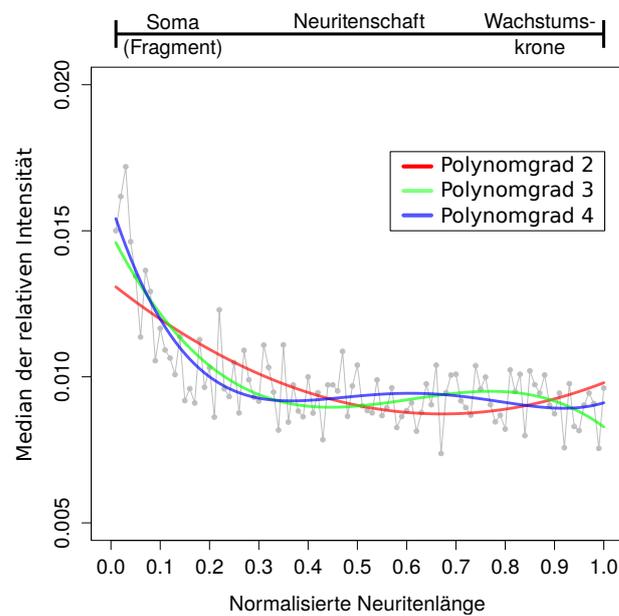


Abbildung 2.12.: Vergleich unterschiedlicher Approximationen der Intensitätsverteilung (grau) mittels Polynomen verschiedener Ordnung (rot, grün, blau).

2.2. Transkriptomanalyse

Dank der heute verfügbaren Sequenzierungsverfahren können, beispielsweise genomische, transkriptomische oder exonische, Abweichungen innerhalb zu untersuchender Gewebe, Zellen, primärer Tumore oder ganzer Organismen in großen Maßstäben, die bis zu mehreren hundert Proben umfassen können, ermittelt und analysiert werden. Insbesondere im klinischen Bereich zur Diagnose, Therapie und Prognose von Krebserkrankungen eröffnen die aktuellen Verfahren neue und bedeutende Möglichkeiten in der Erforschung und zur Behandlung der Erkrankungen. Ein wesentlicher Faktor, der auch in dieser Arbeit im Fokus steht, ist es, die zu einer schnellen Diagnose führenden pathologischen Befunde, wie z. B. die Stadieneinteilung von malignen Tumoren, deren Lokalisation und Verbreitung, um wesentliche molekularbiologische Aspekte, wie beispielsweise Genamplifikationen, Mutationen oder Abweichungen in den Chromosomen oder dem Transkriptom, zu ergänzen. Dabei ist ein Ziel dieser Arbeit, die transkriptomischen Veränderungen innerhalb der klinischen Stadien des Neuroblastoms, speziell die Veränderungen in der mRNA und microRNA Expression, zu analysieren. Dazu wurden in dieser Arbeit verschiedene Methoden entwickelt, die auf den RNA- und microRNA-Sequenzierungen primärer Tumorproben basieren. Die im Folgenden beschriebene Methodik dient zum einen der Ermittlung differentiell exprimierten mRNAs und microRNAs (*DE-mRNAs* und *DE-microRNAs*) zwischen den verschiedenen klinisch relevante Stadien der Neuroblastomproben. Des Weiteren finden die hier entwickelten Methode Anwendung mit dem Ziel der Identifizierung und Analyse stadienspezifischer Gensignaturen auf Basis der hier ermittelten DE-mRNAs der RNA-Sequenzierung sowie putativ neuer microRNA-Target Beziehungen anhand der DE-microRNAs der microRNA-Sequenzierung.

2.2.1. RNA-Sequenzierungsdaten

Die Transkriptomanalyse der Proben des Neuroblastoms wurde in dieser Arbeit auf 66 primären Tumoren durchgeführt. Das Tumormaterial dieser Proben stammt aus der Kölner Tumorbank sowie dem Universitätsklinikum Essen. Im Vorfeld der in dieser Arbeit entwickelten Analysen wurden die Proben von Dr. J. L. Bell⁸ aufbereitet und die für die Sequenzierung benötigte Total-RNA isoliert. Ein bedeutender Aspekt, der in dieser Arbeit herauszustellen ist, ist dadurch geben, dass aus den einzelnen Tumorproben sowohl das Material für die RNA- als auch für

⁸Institut für Molekulare Medizin, Martin-Luther-Universität Halle-Wittenberg

	s1	s2	s3	s4	s4S	MYCN
RNA	14	5	8	25	9	12
microRNA	14	6	9	25	12	12

Tabelle 2.1.: Einteilung der sequenzierten primären Tumorproben für die RNA- und microRNA-Sequenzierung anhand der Stadien nach dem internationalen Einteilungssystem (INSS) des Neuroblastoms. Die verschiedenen Stadien (1, 2, 3, 4, 4S) sind mit den Bezeichnungen (s1, s2, s3, s4, s4S) gekennzeichnet. Die letzte Spalte gibt die Anzahl der Proben an, die eine MYCN Amplifikation aufweisen.

die microRNA-Sequenzierung gewonnen wurde. Dies bietet den entscheidenden Vorteil, dass die Sequenzierungsergebnisse für die weiteren Analysen, hinsichtlich der Veränderungen im Transkriptom zwischen den klinischen Stadien, auf mRNA- als auch microRNA-Ebene, bzgl. der Stärke der Zunahme oder Abnahme der Expression veränderter Transkripte, zu einem Zeitpunkt aus der selben Probe untersucht werden können.

Die isolierte Total-RNA der Proben wurde in der *Core Unit DNA-Technologien* am IZKF⁹ der Medizinische Fakultät an der Universität Leipzig sequenziert. Dabei wurden sogenannte *barcoded libraries* verwendet, die eine Sequenzierung von bis zu 96 Proben gleichzeitig ermöglichen (*Multiplexing*). Bei den verwendeten *Barcodes* handelt es sich um kurze künstliche Sequenzstücke, die jeder Probe hinzugefügt werden. Mit diesen Barcodes können gleichzeitig sequenziert Proben eines Experiments anschließend identifiziert und getrennt voneinander analysiert werden, was zu einer erheblichen Kosten- als auch Zeitersparnis führt. Abhängig von der Art der Sequenzierung, d. h. Total-RNA-Sequenzierung (bezeichnet als RNA-Seq.) für die Analyse der mRNAs und eine Sequenzierung kleiner RNAs (als *small-RNA*- oder microRNA-Seq. bezeichnet) für die Analyse der microRNAs im Transkriptom, wurde die isolierten Total-RNA am IZKF aufbereitet (vgl. hierzu auch 1.3). Für den RNA-Seq. wurde die Total-RNA durch Hitze fragmentiert und es wurden ribosomale RNAs (*rRNA*) mittels Kit zum Abbau von rRNA (*RiboMinus*, Thermo Fisher Scientific) entfernt. Die entstandenen Fragmente wurden anschließend mittels *Paired-End* Verfahren sequenziert. Dabei entstanden Paare von Sequenzen, den sogenannten *Reads*, die eine Länge von 101 Basen aufweisen. Die Sequenzierung der microRNAs erfolgte im Anschluss an die Auftrennung der Total-RNA und einer Größenselektion zwischen 140 und 165 Basen (inklusive Adapter- und Primer-Sequenzen). Hierbei fand keine Fragmentierung der RNA statt und es entstanden einzelne Reads mit einer Länge von 51 Basen. Die Hochdurchsatz-Sequenzierung (vgl. 1.3) der Total-RNA als auch der microRNA erfolgte unter Verwendung

⁹<http://www.uni-leipzig.de/~izkf/>

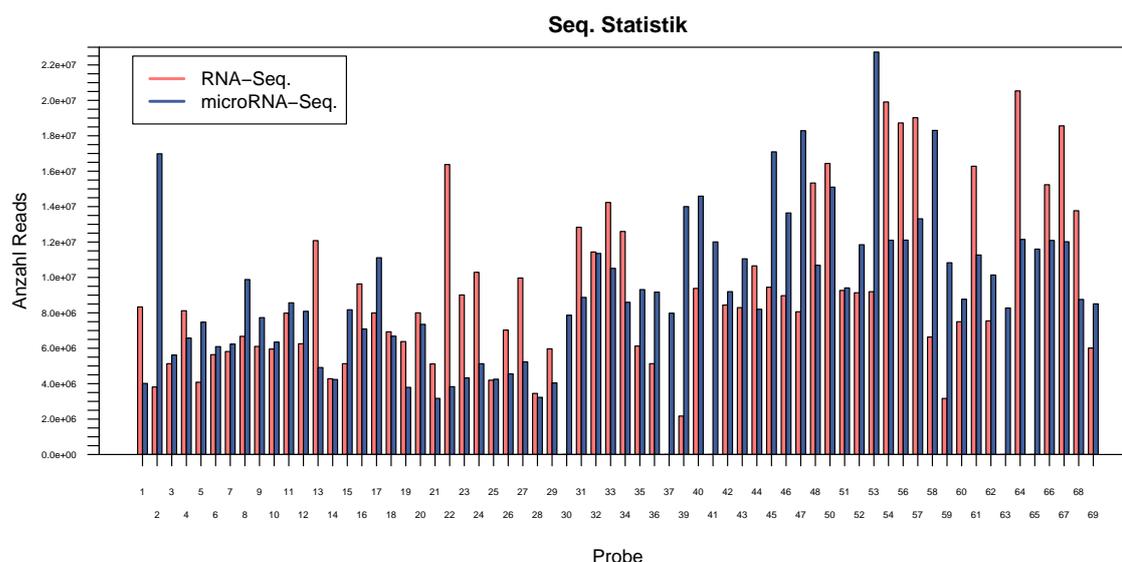


Abbildung 2.13.: Anzahl der sequenzierten Reads der einzelnen Proben der RNA- (rot) und microRNA-Sequenzierung (blau). Die Proben 30, 37, 41, 63 und 65 sind in den Daten der RNA-Sequenzierung nicht enthalten, da diese 5 Proben aufgrund von Sequenzierungsfehlern verworfen wurden und nachfolgend nicht in die Analyse eingegangen sind.

eines Illumina HiScanSQ¹⁰ Sequenzierers. Die aus den Sequenzierungen resultierenden Rohdaten zur weiteren Verwendung der in dieser Arbeit durchgeführten Analysen liegen im FASTQ [43] Format vor, welches ein Standardformat für Sequenzierungsdaten darstellt. Alle in dieser Arbeit analysierten Proben wurden, entsprechend ihrer pathologischen Befunde, von der Tumorbank bzw. Universität in sogenannte Tumorstadien (*Stages*) eingeteilt und sind im Folgenden nach dem, einleitend in dieser Arbeit beschriebenen, *International Neuroblastoma Staging System* (INSS) gekennzeichnet (vgl. 1.2). Einen Überblick über die vorliegende Probenanzahl der jeweiligen Stadien, die für die RNA- und microRNA-Sequenzierung verwendet wurden, ist in Tabelle 2.1 dargestellt. Die vollständige Sequenzierung aller Proben wurde in zwei Sequenzierungsdurchläufen am IZKF realisiert, bei denen pro Durchlauf 28 und anschließend 38 der RNA- bzw. microRNA-Proben gemeinsam sequenziert wurden. Eine detaillierte Übersicht der einzelnen Proben und deren Merkmale befindet sich im Anhang in der Tabelle B.2. Aufgrund von Sequenzierungsfehlern mussten fünf RNA-Seq. Proben verworfen werden, so dass für die weiteren Analysen in dieser Arbeit 61 Proben aus dem RNA-Seq. und 66 Proben aus dem microRNA-Seq. verwendet werden konnten. Darunter befinden sich insgesamt 12 Proben, in denen eine Amplifikation¹¹ des *MYCN* Genes von Dr. J. L. Bell über die DNA-Menge bestimmt

¹⁰https://support.illumina.com/sequencing/sequencing_instruments/hiscansq.html

¹¹Vervielfältigung eines Gens durch DNA-Replikation.

wurde. Diese Proben sind auf verschiedenen Tumorstadien verteilt, wobei vier der Proben aus dem Stadium 3 (s3) und acht der Proben aus dem Stadium 4 (s4) stammen (vgl. Tab. 2.1). Ergänzend hierzu sind in Abb. 2.13 die Anzahlen der sequenzierten Reads der einzelnen Proben für die RNA- und microRNA-Sequenzierung dargestellt¹².

2.2.2. RNA-Sequenzierungsanalyse

Die aus den Sequenzierungen resultierenden Reads (vgl. Abb. 2.13) bilden die gesamtheitliche Datengrundlage für die Transkriptomanalyse. Die im Folgenden beschriebene RNA-Sequenzierungsanalyse stellt den ersten Schritt der Analyse dar und dient im Wesentlichen der Qualitätskontrolle der Sequenzierungsdaten sowie der Annotation der einzelnen Reads, basierend auf dem menschlichen Genom. Abbildung 2.14 zeigt das Datenflussdiagramm der verschiedenen Schritte der vollständigen Transkriptomanalyse wie sie für diese Arbeit entwickelt und angewendet wurde. Die RNA-Sequenzierungsanalyse setzt sich aus den Schritten der Qualitätskontrolle, des Alignments und der Annotation zusammen, die nachfolgend im Detail erläutert werden. Die analysierten RNA-Sequenzierungsdaten sind über die *R2 Genomics Analysis and Visualization Platform*¹³ zugänglich¹⁴.

Qualitätskontrolle

Den ersten Schritt der Analyse stellt die Qualitätskontrolle der sequenzierten Reads dar. Sie dient der Kontrolle der Reads auf Basenebene und kann darüber Aufschluss geben, ob Probleme bei der Sequenzierung (z.B. geringe Mengen an RNA, Störungen im System, zu kurze Reads etc.) aufgetreten sind, welche Qualität die einzelnen Basen im Mittel aufweisen oder ob bestimmte Sequenzen überrepräsentiert werden. Heutige Sequenzierungen mittels Illumina Geräten weisen oftmals sinkende Qualitätswerte der Basen am Ende der Reads auf. Darüber hinaus können Adaptersequenzen in den Reads vorhanden sein, die das spätere Alignment negativ beeinflussen können. Um die Qualität der Daten zu bestimmen und zu verbessern, wurde zunächst eine Qualitätskontrolle mit anschließendem Trimming auf den FASTQ-Rohdaten durchgeführt. Für die Qualitätskontrolle wurde das Programm „FastQC“ (Version 0.11.2)¹⁵

¹²Die Nummerierung der Proben ist nicht lückenlos fortlaufend, da Proben im Vorfeld verworfen wurden, so diese beispielsweise geringe RNA-Mengen oder -Qualitäten aufwiesen.

¹³<http://r2.amc.nl>

¹⁴Registrierung per Mail, unter Angaben von Name und Institution. Der Zugang zu den hier beschriebenen Daten erfolgt auf zusätzliche Anfrage an Jan Koster (jankoster@amc.uva.nl).

¹⁵<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

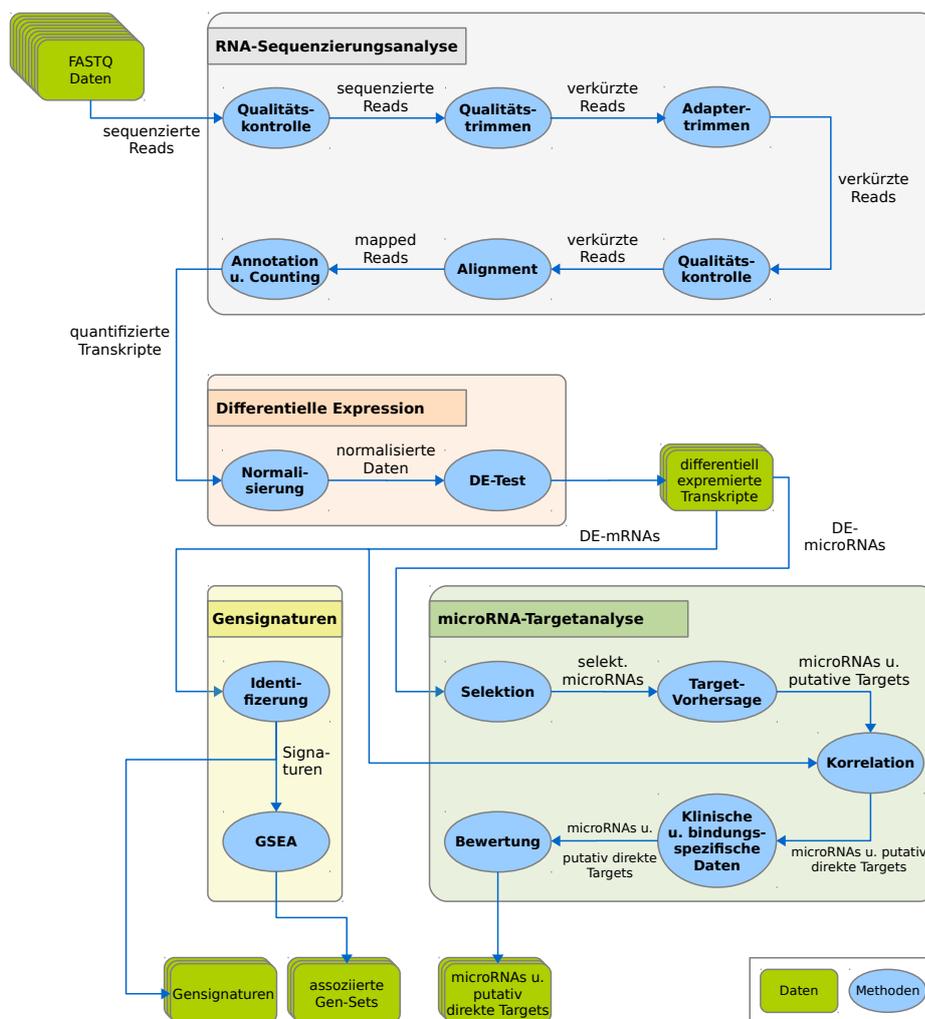


Abbildung 2.14.: Datenflussdiagramm der Transkriptomanalyse. Die Analyse setzt sich aus vier Teilen zusammen, die sequentiell auf den Rohdaten der Sequenzierungen angewendet werden. Mithilfe der darin enthaltenen Schritte werden differenziell exprimierte mRNAs und microRNAs sowie darauf basierende neuroblastomspezifische Gensignaturen und putativ neue microRNA-Target Beziehungen eruiert.

Parameter	Wert		Funktion
	RNA	microRNA	
-q	10	10	Schwellwert des Qualitätstrimmen
-O	7	3	Überlappung zwischen Read u. Adapter
-m	20	15	minimale Readlänge
-a	5'-Adapter ¹⁷ 3'-Adapter ¹⁹	5'-Adapter ¹⁸	Adaptersequenz

Tabelle 2.2.: Verwendete Parameter der Qualitätskontrolle der RNA- und microRNA-Sequenzierung im Programm *cutadapt*. Erklärungen zu den genannten Parametern können aus dem Text entnommen werden.

verwendet. Dieses analysiert die FASTQ-Dateien und gibt einen Überblick auf die Metadaten (Sequenzierungsgerät, Anzahl und Länge der Reads etc.) und veranschaulicht grafisch die Qualität der Sequenzierungsdaten unter verschiedenen Aspekten. Das Programm gibt die basenweisen Qualitätswerte während der Sequenzierung wieder, kann jedoch keinen Aufschluss darüber geben, ob bereits Schritte vor der Sequenzierung, z. B. bei der RNA Isolierung oder der Generierung der Sequenzierungs-Bibliotheken (*Libraries*), fehlerbehaftet waren oder Kontaminationen in den Proben vorlagen. Im Anschluss daran wurde zur Verbesserung der Daten das Programm „cutadapt“ [150] (Version 1.4.2; 1.6)¹⁶ verwendet, welches die Qualität der Reads durch das Abschneiden einzelner Basen (*Trimming*) mit schlechter Qualität (*Qualitätstrimmen*) und dem Entfernen von Adaptersequenzen (*Adaptertrimmen*) verbessern kann. Dabei wurden für den RNA-Seq. und den microRNA-Seq. unterschiedliche Parameter verwendet, welche in Tab. 2.2 zusammengefasst sind.

Das Qualitätstrimmen fand vor dem Adaptertrimmen statt, um auszuschließen, dass vorhandene Adaptersequenzen aufgrund von qualitativ schlecht bewerteten Nukleotiden oder durch mögliche Sequenzierungsfehler nicht erkannt werden. Lag die Gesamtqualität der einzelnen Nukleotide bzw. Basen eines Reads unterhalb der angegebenen Grenze, erfolgt ein Trimming vom 3'- zum 5'-Ende. Dabei wurde der gegebene Schwellwert für das Qualitätstrimmen von den aktuellen Qualitätswerten der einzelnen Basen subtrahiert, die partielle Summe vom 3'-Ende her gebildet und der Read an der Stelle geschnitten, an dem die Summe minimal wurde²⁰. Im Anschluss

¹⁶Da die Sequenzierung in zwei Durchläufen erfolgte, können mehrere Versionen benutzt worden sein. Wenn vorhanden gilt die erste Version für den ersten und die zweite Version für den zweiten Durchlauf der Sequenzierungen.

¹⁷5'-Adapter RNA-Seq.: AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC

¹⁸5'-Adapter microRNA-Seq.: TGGAATTCTCGGGTGCCAAGGAACTCCAGTCAC

¹⁹3'-Adapter RNA-Seq.: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

²⁰<http://cutadapt.readthedocs.org/en/stable/guide.html>

daran erfolgte der Abgleich (*Alignment*) der gegebenen Adaptersequenzen mit den einzelnen Reads. Wurden innerhalb der Reads Adapter mit einer minimalen Überlappung von mindestens sieben Basen (RNA-Seq.) bzw. 3 Basen (microRNA-Seq.) gefunden, so wurden diese entfernt und der Read wurde erneut verkürzt. Die finalen Reads wurden auf eine minimale Länge von 20 Basen (RNA-Seq.) und 15 Basen (microRNA-Seq.) begrenzt. Kürzere Reads wurden verworfen, da später ein möglichst eindeutiges Alignment mit dem Referenzgenom nicht mehr gewährleistet wäre oder die biologisch sinnvolle Größe der zu identifizierenden Transkripte nicht mehr zutreffend ist (microRNAs haben beispielsweise eine Länge von ca. 22 Nukleotiden). Tritt bei den gepaarten Reads des RNA-Seq. der Fall ein, dass ein Read die minimale Länge unterschreitet, so wurden beide Reads verworfen. Alternativ könnte der längere der beiden Reads erhalten bleiben, was jedoch zwei ungleichen Mengen an Reads zur Folge hat und in den nachfolgenden Schritten des Alignments und der Annotation zu Fehlern oder Verlusten von Reads führen kann. Des Weiteren könnten zu stark verkürzte Reads in ihrer vollen Länge erhalten bleiben, was jedoch ebenso zu Problemen oder fehlerhaften Zuordnungen beim Alignment führen kann. Für die spezielle Behandlung der gepaarten Reads wurde zusätzlich zu cutadapt ein Perl-Skript verwendet das diese Sonderfälle berücksichtigt²¹. Neuere Versionen von cutadapt sind bereits in der Lage diese Fälle zu handhaben und bieten verschiedene Einstellungsmöglichkeiten der Vorgehensweise an. Als Ergebnis des Trimmings wurden neue FASTQ-Dateien erzeugt, welche die, hinsichtlich ihrer Qualität, verbesserten Reads enthalten. Es fand anschließend eine erneute Qualitätskontrolle der Reads mittels FastQC statt anhand derer eine Statistik der hier erläuterten Schritte generiert wurde, um einen schnellen Überblick über die Optimierung der Reads zu ermöglichen. Diese Reads bilden die Grundlage für den nachfolgenden Schritt des Alignments der Reads gegen ein gegebenes Referenzgenom.

Alignment

Das Alignment ist der zeitaufwändigste und rechenintensivste Schritt bei der Analyse der Sequenzierungsdaten. Hierbei erfolgte für die einzelnen Reads der Sequenzierungen der Abgleich ihrer Sequenzen mit der Sequenz eines Referenzgenoms. Die Reads wurden mindestens einer Position im Genom zugeordnet, an dessen Stelle die Readsequenz möglichst genau mit der Referenz übereinstimmte. Für das Alignment der Daten der RNA- und microRNA-Sequenzierung kamen unterschiedliche Programme zur Anwendung, deren Algorithmik darauf

²¹Das Skript wurde von <https://www.ecseq.com/support/> zur Verfügung gestellt.

ausgelegt ist, die verschieden langen Reads der zu untersuchenden Transkripte mit der Referenz zu vergleichen und die Position der Sequenz im Genom zu ermitteln.

Die RNA-Seq. Daten wurden mithilfe von „TopHat2“ [108] (Version 2.0.12; 2.0.13) gegen das Humangenom (UCSC hg19 assembly, GRCh37)²² abgeglichen. TopHat2 bietet die Möglichkeit, Reads über Spleißstellen hinweg miteinander zu verknüpfen. Dieser Schritt ist vor allem bei größeren zu untersuchenden Transkripten, wie zum Beispiel mRNAs, nützlich und kommt in dieser Arbeit zur Anwendung. Bei den kurzen, nicht fragmentierten und nicht gespleißten, microRNAs hingegen ist diese aufwendige Form des Alignments nicht von Nöten. Aus diesem Grund wurde für das Alignment der microRNAs das Programm „Bowtie2“ [122] (Version 2.2.3.0; 2.2.4.0) mit der gleichen Referenz des Humangenoms verwendet. Bowtie2 wird intern auch von TopHat2 genutzt, bietet jedoch selbst nicht die Möglichkeit Reads über Spleißstellen hinweg zu ermitteln. Die wichtigsten Parameter, die bei der Verwendung der Programme Anwendung fanden, sind in Tab. 2.3 dargestellt. Zu erwähnen ist dabei die Nutzung mehrfacher Alignments und das Zulassen von Fehlern innerhalb der ersten Basen des Reads. Mehrfache Alignments treten auf, wenn die Sequenz eines Transkripts verschiedenen Stellen im Genom zugeordnet werden kann. Was, neben kurzen Sequenzen, z.B. häufig bei Pseudogenen oder auch microRNAs vorkommt. So tritt beispielsweise der 5'-Arm (hsa-let-7a-5p) der humanen *hsa-let-7a* microRNA an drei Stellen im Genom auf, während der 3'-Arm (hsa-let-7a-3p) nur genau einmal im Humangenom vorhanden ist. Dies sollte bei jeder Analyse von Sequenzierungsdaten beachtet und entsprechend der experimentellen Fragestellung behandelt werden. Standardmäßig werden in vielen Programmen keine mehrfachen Alignments verwendet oder zugelassen, was häufig dazu führt, dass eine der möglichen Positionen, denen der Read zugeordnet werden kann, zufällig bestimmt und als Ergebnis geliefert wird. Innerhalb dieser Arbeit wurden mehrfache Alignments bis zu einer Anzahl an 20 möglichen Positionen zugelassen, die mit in die nachfolgenden Analysen eingehen. Des Weiteren wird eine definierten Anzahl an Basen eines Reads häufig als *Seed-Region* bezeichnet, welche in den Standardeinstellungen der genannten Programme durch die 20 Basen des Reads definiert sind. Bowtie2 extrahiert diese Seeds aus der Sequenz sowie dem reversen Komplement der gegebenen Reads. Die Seeds sind kontinuierliche, lückenlose Sequenzen, die sich gegenseitig überlappen können, da sie standardmäßig aller 10 Basen entlang des Reads extrahiert werden [122]. Die Seeds dienen einem schnellen, sensitiven und speichereffizientem Alignment, wobei in einem ersten

²²<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/>

Schritt die Seeds vollständig mit der Referenz abgeglichen werden, während die Seeds in einem zweiten Schritt, unter Berücksichtigung eines möglichen Fehlers (*Mismatch*), zur Referenz verglichen und deren Positionen im Genom ermittelt werden [122]. Das Zulassen eines Fehlers in dieser Region kann dazu führen, dass Reads mehreren Positionen innerhalb der Referenz zugeordnet werden können. Gleichzeitig ermöglicht das Erlauben eines Fehlers jedoch, dass, beispielsweise durch Punktmutationen, veränderte Reads nicht bereits im Vorfeld verworfen werden, wenn innerhalb des Seeds ein Mismatch auftritt. Final ergibt sich aus jeder Probe das Alignment der Reads im Standardformat BAM [128], in dem die Reads indiziert und sortiert nach den chromosomalen Positionen des Referenzgenoms vorliegen. Die dem Humangenom zugewiesenen Reads, sogenannte *mapped* Reads, wurden im nachfolgenden Schritt der Annotation bekannten Transkripten zugeordnet und gezählt. Die nicht zugewiesenen Reads, sogenannte *unmapped* Reads, wurden ebenfalls im BAM Format gespeichert und können beispielsweise für weiterführende Analysen oder die Ermittlung von Kontaminationen, wie z.B. bakterielle Infektionen mit Mykoplasmen, verwendet werden.

Parameter	Wert		Funktion
	RNA	microRNA	
-r	50	-	Read-Distanz innerhalb eines Paares
-g/-k	20	20	Anzahl mehrfacher Alignments
-b2-N/-N	1	1	Fehler innerhalb der Seed-Region
index	hg19	hg19	Verwendetes Referenzgenom

Tabelle 2.3.: *Verwendete Parameter des Alignments in den Programmen TopHat2 und Bowtie2 für die Daten der RNA- respektive microRNA-Sequenzierung. Für die Anzahl der mehrfach erlaubten Alignments und den Fehler innerhalb der Seed-Region verwenden beide Programme unterschiedliche Parameternamen.*

Annotation und Auszählung

Die Annotation und die Auszählung (*Counting*) der mapped Reads ist der letzte Schritt in der RNA-Sequenzierungsanalyse. Dabei wurden die Reads anhand ihrer ermittelten Position im Referenzgenom mit einer, zu diesem Genom passenden, Annotations-Datei verglichen und bekannten Transkripten zugeordnet. Somit erhält jeder sequenzierte Read eine funktionelle Zuordnung zu einem bereits bekannten Transkript und es ist möglich diese nach ihrem biologischen Typ (mRNA, tRNA, snoRNA etc.) zu unterscheiden und das zugehörige Transkript

Parameter	Wert		Funktion
	RNA	microRNA	
-p	ja	nein	Verwendung von gepaarten Reads
-M	ja	ja	Verwendung mehrfacher Alignments
-t	exon	miRNA	Merkmalstyp der Annotierung
-g	gene_id	ID	Gruppierungsmerkmal (Meta-Feature)
-a	Ensembl GRCh 37.75	miRBase 20	Verwendete Annotation mit Version

Tabelle 2.4.: Verwendete Parameter der Annotation und des Auszählens im Programm *featureCounts* für die Daten der RNA- und microRNA-Sequenzierung.

zu quantifizieren. Annotation und Auszählung erfolgten mithilfe des Programms „*featureCounts*“ [132] (Version 1.4.6). Für die RNA-Seq. Daten wurde die Annotation des Humangenoms (Ensembl release 75, GRCh37.p13)²³ verwendet und es fand eine Annotierung auf codierende Bereiche (Exons) des Genoms statt, während für die microRNA-Seq. die Annotation aus der microRNA-Datenbank miRBase [74–76, 115, 116] (Version 20, 2794 humane microRNAs)²⁴ verwendet wurde. Die Nutzung der miRBase bietet den Vorteil, dass gegenüber der genomischen Annotation mehr annotierte microRNAs enthalten sind. In beiden Fällen der Annotierung ist darauf zu achten, dass die Version der Annotation der Version des Referenzgenoms entspricht, da sich, neben Änderungen der Gen-Namen oder biologischen Typen, chromosomale Positionen zwischen den Versionen verschieben können und demnach die Reads nicht korrekt annotiert werden. Des Weiteren spielen die Parameter der Annotation und des Auszählens eine wichtige Rolle, da diese darüber entscheiden, was und wie annotiert und gezählt wird (Protein-codierende Bereiche, exonische oder intronische Bereiche, spezielle RNA-Typen etc.). Weiterhin muss festgelegt werden, wie mit speziellen Situationen (überlappende Transkripte, Paired-End Reads etc.) umgegangen werden soll. Die verwendeten Parameter, die für die Analyse in dieser Arbeit genutzt wurden, sind dazu in Tab. 2.4 aufgelistet. Zum Abschluss der Annotation und Auszählung wurde eine Liste von Transkripten generiert, anhand derer für jedes annotierte Transkript die ermittelte Anzahl an Reads (*Counts*) der einzelnen Proben entnommen werden kann. Da je nach Annotation die Transkripte meist nur über eine Identifizierungsnummer (Gene ID) gekennzeichnet sind, wurden mithilfe des Bioconductor Paketes „*biomaRt*“ [59] (Version 2.22.0) die Namen der identifizierten Transkripte zusätzlich zur Gene ID ermittelt und gespeichert.

²³ftp://ftp.ensembl.org/pub/release-75/gtf/homo_sapiens/Homo_sapiens.GRCh37.75.gtf.gz

²⁴<ftp://mirbase.org/pub/mirbase/20/genomes/hsa.gff3>

2.2.3. Differentielle Expression

Die differentielle Expression (DE) gibt Auskunft über die transkriptomischen Veränderungen auf Basis der Genexpression mittels Vergleich von Proben unterschiedlicher Bedingungen, wie beispielsweise der Veränderung zwischen Wildtyp und Mutante, behandelter und unbehandelter Probe oder zwischen Tumor- zu Normalgewebe. Mithilfe der Methode der DE lassen sich in den vorliegenden Sequenzierungsdaten der Neuroblastomproben die Unterschiede in der Expression der Transkripte zwischen den verschiedenen Stadien detektieren. Das Ziel ist es, die Gene zu identifizieren, die zwischen den Stadien eine signifikante Veränderung in ihrer Abundanz aufweisen, um infolgedessen stadienspezifische biologische Erkenntnisse aus diesen Daten zu erlangen. Auf Grundlage der Counts der einzelnen Transkripte aus den Sequenzierungsdaten lässt sich ein (annähernd) linearer Zusammenhang zwischen den Counts der Expressionsdaten und der relativen Abundanz der Transkripte herstellen [170]. Eine Approximation der Count-Daten durch eine kontinuierliche Verteilung, wie bei Microarray-Daten, ist jedoch nicht optimal, da die Expressionsdaten in Form der Counts keine kontinuierlichen Intensitäten aufweisen. Dies kann insbesondere bei niedrigen Count-Werten und einer geringen Probenanzahl zu Problemen bei der Analyse und zu fehlerhaften Auswertungen führen. Aus diesem Grund werden die Daten anhand von diskreten Wahrscheinlichkeitsverteilungen modelliert und mittels statistischer Methoden ausgewertet, um Informationen über die differentiell exprimierte Gene (DE-Gene) aus den Daten zu erlangen [185]. Eine häufige angewendete Modellierung der Count-Daten aus RNA-Sequenzierungen stellt die Poissonverteilung zur Analyse der DE-Gene dar [27, 147]. Durch die biologische Variabilität innerhalb und zwischen den Proben kann die Modellierung mithilfe der verteilungsbedingten Annahme, dass Mittelwert und Varianz der Daten gleich sind, eine erhöhte falsch positive Rate an ermittelten DE-Genen zur Folge haben. Zudem ist bekannt, dass für RNA-Sequenzierungen die Varianz meist größer als der Mittelwert ist, auch als *Überdispersion* bezeichnet, und insbesondere bei Genen mit hoher Expressionsrate, eine meist geringere Varianz vorliegt [3, 211, 212]. Zur Berücksichtigung dieser Variabilität wird in neueren Analysen überwiegend die negative Binomialverteilung verwendet. Diese Verteilung stellt eine natürliche Erweiterung der Poissonverteilung dar und weist einen zusätzlichen Streuungsparameter auf, der die Varianz der Daten reguliert [3, 185, 212]. Die Selektion differentiell exprimierter Transkripte, aus den analysierten Sequenzierungsdaten des Neuroblastoms, geht mit der Normalisierung der Count-Daten und dem eigentlichen statistischen Testen auf differentielle Expression einher, was

in den folgenden beiden Abschnitten im Detail erläutert wird. Auf Basis des RNA-Seq. wurden wie bereits erwähnt die DE-mRNAs Protein-codierender Gene ermittelt, während basierend auf dem microRNA-Seq. die DE-microRNAs ermittelt wurden.

Normalisierung

Die Normalisierung der Sequenzierungsdaten ist ein wesentlicher und notwendiger Schritt, um sowohl die Expressionslevel der Gene innerhalb und zwischen den Proben zu vergleichen als auch für die anschließende Analyse der Gene auf differentielle Expression. Ähnlich wie bei Microarray-Daten ist auch hier das Ziel, systematische Variationen, sowohl biologischen als auch technischen Ursprungs, zu minimieren ohne weitere Störungen zu generieren. Eine geeignete Normalisierung sollte sicher stellen, dass die verbleibenden Variationen biologischer Natur sind, um einen geringen Einfluss auf die Ergebnisse nachfolgender Analysen zu haben [55, 211, 241]. Systematische Variationen, die bei Sequenzierungsdaten auftreten, lassen sich in intra- und interlibrary Variationen unterscheiden, d.h. Variationen innerhalb und zwischen den einzelnen Proben der Sequenzierungen. Wobei die Proben oftmals auch als Sequenzierungs-Bibliotheken (*Libraries*) bezeichnet werden (vgl. 1.3). Intralibrary Variationen treten durch genspezifisch Effekte auf, wie beispielsweise Unterschiede in den Längen der Gene. Variationen zwischen den Libraries hingegen sind durch die unterschiedliche Größe der Libraries und somit die Anzahl der sequenzierten Reads pro Probe gekennzeichnet. Aktuell gibt es eine Reihe von Normalisierungsmethoden, die die unterschiedlichen Variationen berücksichtigen, jedoch gibt es keine klaren Angaben, welche Methode sich für bestimmte Daten eignet oder wie diese weitere Analysen, wie z. B. das Testen auf DE, beeinflussen [55, 185, 241].

Für die Analyse der vorliegenden Neuroblastomdaten auf DE-mRNAs und -microRNAs ist eine interlibrary Normalisierung maßgebend, da es das Ziel ist, die differentielle Expression aus den Vergleichen der Proben verschiedener Stadien zu ermitteln. Methoden, die zur Normalisierung der Daten sogenannte Skalierungsfaktoren aus den Daten schätzen können, haben sich hier als geeignet erwiesen, da sich diese Faktoren in die statistischen Modelle der DE-Analyse integrieren lassen und die Count-Daten für nachfolgende Analysen erhalten [185]. Auf Grundlage einer interlibrary Normalisierung und den daraus ermittelten Skalierungsfaktoren wurden die vorliegenden Daten in dieser Arbeit normalisiert, auf denen anschließend die differentiell exprimierten Gene bestimmt wurden. Bei der hier angewendeten Normalisierung handelt es sich um die *Trimmed Mean of M-values* Methode, kurz als *TMM* bezeichnet [211]. Diese Methode

hat sich sowohl für die Normalisierung von RNA- als auch microRNA-Sequenzierungsdaten als praktikabel erwiesen [22, 55, 241]. Während die Gesamtanzahl an Reads pro Probe durch die Sequenzierung bekannt ist, bleibt die reale Menge an vorhandener RNA unbekannt und kann zwischen den Proben, in Abhängigkeit der RNA-Zusammensetzung der Probe, stark variieren. In Proben, die große Mengen RNA beinhalten, kann dies zu einer Unterschätzung vieler Genen führen, relativ gesehen zu einer Proben mit weniger RNA-Gehalt. Da die Expression der einzelnen Gene und deren tatsächliche Transkriptlänge unbekannt sind, kann die vorhandene Menge an RNA eines Gens aus der Probe nicht direkt geschätzt werden. Hier bietet die verwendete TMM-Methode die Möglichkeit, einen globalen *Fold Change* (FC)²⁵ der relativen RNA-Menge zwischen zwei Proben zu bestimmen. Dabei beruht die grundlegende Annahme der Methode auf der Hypothese, dass der Großteil der Gene zwischen den Proben nicht differentiell exprimiert ist, diese somit einen FC von 0 aufweisen. Die Bezeichnung *Trimmed Mean* der Normalisierungsmethode steht hierbei für die Berechnung des Mittelwertes nach der Entfernung der (festgelegten) oberen und unteren Prozente an Genen aus dem Vergleich einer Probe mit der Referenzprobe. Die Referenzprobe kann dabei manuell gewählt oder automatisch bestimmt werden, indem die Probe als Referenz selektiert wird, deren oberes Quartil sich am nächsten zum mittleren oberen Quartil über alle Proben befindet. Für die hier ausgewerteten Daten wurde dabei die automatische Selektion der Referenzprobe gewählt. Unter Anwendung der TMM-Methode findet ein zweifaches Trimming statt. Dabei werden von den logarithmischen FC-Werten ($\log FC$) jedes Gens (als *M-Werte* bezeichnet) die extremsten 30 %, von den absoluten Expressionswerten (als *A-Werte* bezeichnet) die extremsten 5 % entfernt. Das bedeutet, dass die Liste der Gene für die hoch- bzw. herunterregulierten Gene einer Probe, im Vergleich zur Referenzprobe, entsprechend des logarithmischen Verhältnisses und der Expressionswerte jedes Gens getrimmt wird. Die prozentualen Werte des Trimmings lassen sich anpassen, kommen in dieser Arbeit jedoch unverändert in Form der genannten Standardwerte zur Anwendung. Aus den getrimmten $\log FC$ -Werten wird ein gewichteter Mittelwert gebildet, dessen Gewichte aus der inversen Varianz generiert werden, da Gene mit einer erhöhten Read-Anzahl eine geringere Varianz in der logarithmischen Veränderung aufweisen [211]. Basierend auf diesen Gewichten und den zweifach getrimmten $\log FC$ -Werten wird anschließend für jede Probe ein Skalierungsfaktor berechnet. Diese Skalierungsfaktoren können direkt in das statistische Modell integriert werden, welches die Daten hinsichtlich der differentiellen Expression testen soll (siehe

²⁵Der Fold Change (FC) beschreibt das Maß der Veränderung der Expression eines Gens zwischen zwei Proben.

nächster Abschnitt), wobei die Daten der Proben nicht verändert werden. Die Anwendung der Skalierungsfaktoren auf die Daten erzeugt sogenannte „Pseudocounts“, so dass die Anzahl der Reads (*Counts*) und somit die unterschiedlichen Librarygrößen der einzelnen Proben an die Referenzprobe angeglichen werden und eine vollständige Normalisierung der Count-Daten zwischen allen Proben stattfindet. Die Normalisierung mittels TMM-Methode erfolgte für die Sequenzierungsdaten in dieser Arbeit unter Verwendung der *calcNormFactors* Funktion des R Bioconductor Pakets „edgeR“ [210] (Version 3.4.2)²⁶.

Test auf differentiell exprimierter Transkripte

Der Test auf differentielle Expression in den Neuroblastomdaten erfolgt zwischen den Proben klinisch relevanter Stadien. Dazu fand eine Einteilung der Stadien in verschiedene differentielle Vergleiche statt. Die Wahl der Vergleiche zwischen diesen Stadien in dieser Arbeit basiert auf der Aggressivität und der damit verbundenen Prognose der Tumorerkrankung (vgl. 1.2). Aus klinischer Sicht ist es wünschenswert, stadienspezifische Gene zu ermitteln, die zwischen den Entitäten differentiell exprimiert sind und somit als Charakteristikum für die Prognose oder Therapie dieser Stadien fungieren könnten. Bei der Einteilung wurden sowohl metastatische Bedingungen, durch den Vergleich der kaum bis gering metastasierenden Stadien (s1, s2 und s3) zu den stark metastasierenden Stadien (s4 und s4S), als auch die Bedingungen der (hoch)risikoreichen (*high-risk*) Stadien, wie beispielsweise s4 oder mit einer *MYCN* Amplifikation, zu den risikoärmeren (*lowh-risk*) Stadien s1 oder s4S betrachtet. Des Weiteren wurde der Fall der spontanen Regression des Tumors durch den Vergleich s1 zu s4S mit berücksichtigt. Stadien, die mit $+M$ oder $-M$ markiert sind, kennzeichnen die Verwendung von Teilmengen der Proben die eine *MYCN* Amplifikation aufweisen bzw. nicht aufweisen. Die Einteilung der klinisch relevanten Stadien, die für den Test auf differentielle Expression verwendet wurden, sind in Tab. 2.5 zusammengefasst und kurz beschrieben.

Basierend auf den Skalierungsfaktoren der TMM-Normalisierung wurden die Count-Daten der Transkripte normalisiert (Pseudocounts) und zwischen den klinischen Stadien, der in Tab. 2.5 definierten Vergleiche, auf differentielle Expression getestet. Mittels Test auf DE soll ermittelt werden, ob die relative Abundanz einzelner Transkripte zwischen den betrachteten Stadien, der in dieser Arbeit analysierten Neuroblastomproben, signifikant Unterschiede aufweisen. Der Test erfolgt unter Verwendung der *exactTest* Funktion des R Bioconductor Pakets „edgeR“

²⁶<https://bioconductor.org/packages/release/bioc/html/edgeR.html>

Vergleich	Beschreibung
1/2/3 – 4/4S	schwach gegen stark metastasierende Stadien
1 – MYCN	Stadium 1 gegen Stadien mit <i>MYCN</i> Amplifikation
1 – 4	Stadium 1 gegen 4
1 – 4 ^{+M}	Stadium 1 gegen 4 mit <i>MYCN</i> Ampl.
1 – 4 ^{-M}	Stadium 1 gegen 4 ohne <i>MYCN</i> Ampl.
1 – 4S	Stadium 1 gegen 4S
4 – 4S	Stadium 4 gegen 4S
4 ^{+M} – 4S	Stadium 4 mit <i>MYCN</i> Ampl. gegen 4S
4 ^{-M} – 4S	Stadium 4 ohne <i>MYCN</i> Ampl. gegen 4S

Tabelle 2.5.: Einteilung der klinisch relevanten Stadien in die Vergleiche der differentiellen Expressionsanalyse. Stadien, die mit ^{+M} oder ^{-M} markiert sind, kennzeichnen die Verwendung von Proben, die eine *MYCN* Amplifikation aufweisen/nicht aufweisen.

[210] (Version 3.4.2). Eine Normalisierung, wie beispielsweise durch die hier verwendete TMM-Methode, unter der Annahme, dass ein Großteil der Transkripte keine differentielle Expression aufweist, ist für diesen Test notwendig. Durch die normalisierten Daten wird gewährleistet, dass die Pseudocounts in den Libraries annähernd identisch verteilt sind und eventuelle Differenzen in den Abundanzen der Transkripte aus den realen Veränderungen zwischen den Stadien resultieren [212]. Der Test arbeitet paarweise auf den zu vergleichenden Stadien einer zu untersuchenden differentiellen Vergleiche und erfolgt zeilenbasiert auf den Pseudocounts der einzelnen Transkripte. Auf den erzeugten Pseudocounts wird ein exakter Test angewendet, der ähnlich dem exakten Test von Fisher für Kontingenztafeln arbeitet. Die im exakten Test von Fisher verwendete hypergeometrische Verteilung wird jedoch, speziell auf die Analyse von Sequenzierungsdaten angepasst, durch eine negative Binomialverteilung ersetzt. Der zusätzliche Streuungsparameter der negativen Binomialverteilung bezieht die intralibrary Variationen ein und berücksichtigt damit die bekannte Überdispersion in RNA-Sequenzierungsdaten, was bedeutet, dass die Varianz in den Daten meist größer als ihr Mittelwert ist [3, 210, 212]. Dem Test liegt die Annahme zugrunde, dass bei einer identischen Verteilung der Pseudocounts, die Summen der Pseudocounts innerhalb einer Bedingung, hier die klinischen Stadien, ebenfalls einer identischen Verteilung folgen. Unter der Nullhypothese, dass die Transkripte keine Veränderungen in ihrer Abundanz zwischen den Stadien aufweisen, wird die Wahrscheinlichkeit dafür berechnet, dass das Abundanzverhältnis (Verhältnis der Pseudocounts) eines Transkripts zwischen zwei Stadien größer-gleich dem Verhältnis aller weiteren Transkripte ist [212]. Der Test liefert somit eine exakte Methode zur Bewertung der differentiellen Expression durch die Berechnung der

Wahrscheinlichkeit dafür, dass eine höhere Abundanz eines Transkripts in einem Stadium verglichen zu einem anderen Stadium auftritt, gemessen an den normalisierten Count-Daten der Sequenzierung.

Im Anschluss an den Test auf differentielle Expression wird in der hier durchgeführten Analyse eine Liste aller annotierter Transkripte erstellt, welche für jedes einzelne Transkript die Information beinhaltet, ob dieses unter Verwendung eines festgelegten p-Werts von 0,05 eine signifikante Veränderung zwischen den Stadien des Vergleiches aufweist. Liegt eine signifikante Veränderung vor, so wird das Transkript als differentiell exprimiert gekennzeichnet. Die Kennzeichnung erfolgt für differentiell exprimierte Transkripte mit einer +1 oder -1, abhängig vom Fold Change und somit von der Richtung der Expressionsveränderung aus dem Vergleich. Nicht differentiell exprimierte Transkripte werden durch den Wert 0 gekennzeichnet. Der p-Wert wird aufgrund der multiplen Tests durch die FDR²⁷ nach der Benjamini-Hochberg-Prozedur korrigiert [15]. Das Maß der Hoch- oder Runterregulation aus den Abundanzen der Transkripte wird über den logFC-Wert gekennzeichnet. Die identifizierten stadienspezifischen Transkripte der differentiellen Expression werden für die weiterführende Identifizierung von Gensignaturen und die microRNA-Targetanalyse in dieser Arbeit verwendet, wobei aus dem RNA-Seq. ausschließlich die mRNA Transkripte betrachtet werden. Wie bereits erwähnt, werden die aus dem RNA-Seq. stammenden differentielle exprimierten mRNAs als *DE-mRNAs*, die aus dem microRNA-Seq. resultierenden differentiellen microRNAs als *DE-microRNAs* bezeichnet. Als Ergebnis der Analyse der differentiellen Expression wurden, für jeden der neun Vergleiche, Tabellen aller annotierten Gene erzeugt, die die normalisierten CPM-Werte der Genen für die einzelnen Proben sowie deren logFC- und FDR-Werte enthalten, die aus dem Test auf differentielle Expression zwischen den Stadien der jeweiligen Vergleiche resultiert. Zusätzlich wird eine Kennzeichnung auf differentielle Expression, unter dem verwendeten p-Wert, in der DE-Wert Spalte angegeben, der wie beschrieben die Werte ± 1 oder 0 für differentielle bzw. nicht differentiell exprimiert Transkripte annehmen kann. Ein Verweis auf die vollständigen Tabellen der differentiellen Expressionsanalyse für die eingeteilten Vergleiche befindet sich im Anhang C.

2.2.4. Identifizierung von Gensignaturen

Eine der weiterführenden Analysen in dieser Arbeit stellt die Identifizierung vergleichsbasierter *Gensignaturen* zwischen den klinischen Stadien der Neuroblastomproben dar. Gensignaturen

²⁷Die False Discovery Rate (FDR) beschreibt das korrigierte Signifikanzniveau beim Testen multipler Hypothesen.

werden hierbei als Mengen von Genen definiert, die aufgrund gemeinsamer Funktionen, beispielsweise der Regulierung eines Stoffwechselweges, an Prozessen wie Differenzierung, Wachstum etc. beteiligt sind. Neben experimentellen Validierungen, durch gezielte Untersuchung von molekularbiologischen Prozessen, können mithilfe von Expressionsdaten derartige Mengen an Genen ermittelt werden. In dieser Arbeit wurden, auf Grundlage der differentiell exprimierten mRNAs, Gensignaturen identifiziert, die die transkriptomischen Veränderungen zwischen den klinisch relevanten Stadien (vgl. Tab. 2.5) der differentiellen Vergleiche widerspiegeln und charakterisieren. Die hier ermittelten Signaturen sind so definiert, dass sie, auf den Vergleichen basierenden, aus einer Menge von Genen bestehen, die je nach Richtung der Veränderung ihrer Expression in sogenannte *Up-Sets* und *Down-Sets* eingeteilt sind und die Gen-Namen der differentiell hoch- bzw. herunterregulierten mRNAs enthalten. Zur Identifizierung der vergleichsbasierten Gensignaturen der Neuroblastomproben wurden die DE-mRNAs selektiert, die eine signifikante und mindestens 2-fache Veränderung in ihrer Expression ($FDR \leq 0,05$ und $|\log_{2}FC| \geq 1$) zwischen den Stadien des betrachteten differentiellen Vergleiches aufweisen. Für eine minimale Repräsentation der Stadien wurde die Mindestanzahl an Genen, die ein *Up*- bzw. *Down-Sets* beinhalten muss, auf einen Wert von zehn Genen festgelegt. Ein entsprechender Referenzwert für eine Mindestanzahl von Genen, die einer charakteristischen Menge von Genen zugrunde liegen sollten, um eine biologische Funktion durch diese Signatur zu charakterisieren, konnte aus der aktuellen Literatur nicht ermittelt werden. Eine Reduzierung der minimalen Größe scheint jedoch nicht sinnvoll, da eine so geringe Anzahl an Genen kaum die komplexen molekularen Prozesse abbilden kann.

Analyse der identifizierten Signaturen

Zusätzlich zur Identifizierung der Gensignaturen wurden deren enthaltenen Gene (Signaturgene) auf ihre gemeinsamen molekularbiologischen Zusammenhänge analysiert. Diese Analyse erfolgte unter Anwendung der *Gene Set Enrichment Analysis* (GSEA), einer analytischen Methode zur Interpretation von Expressionsdaten [237]. Ursprünglich wurde diese Methode zur Auswertung von Microarraydaten, basierend auf Gen-Set Ebene, entwickelt. Diese Gen-Sets sind öffentlich verfügbare und fassen Mengen von Gene in einem annotierten Set zusammen, die aus publizierten Daten, Experimenten oder computergestützten Vorhersagen stammen können und mit diversen molekularbiologischen Prozessen, Stoffwechselwegen, Krankheiten etc. assoziiert sind. Aktuelle GSEA Versionen sind jedoch in der Lage RNA-Sequenzierungsdaten zu analysieren, so dass

Kategorie	Anzahl		Beschreibung
	Gen-Sets	Unterkategorien	
H	50		- vordefinierte biol. Zustände
C1	326		- chromosomale Lage der Gene
C2	4726		5 kuratierte Gen-Sets
C3	836		2 konservierte cis-regulatorische Elemente
C4	858		2 berechnete Gen-Sets (krebsbezogen)
C5	1454		3 Zuordnung zu GO-Termini ³⁰
C6	189		- onkogene Signaturen
C7	4872		- immunologisch Signaturen

Tabelle 2.6.: Überblick über die acht Hauptkategorien der *Molecular Signatures Database* (MSigDB, Version 5.1) die Gene zu verschiedenen biologischen Sachverhalten in Form von öffentlichen Gen-Sets zusammenfasst.

die in dieser Arbeit ermittelten Signaturgene hinsichtlich ihrer Zusammenhänge untersucht werden konnten²⁸. Für die Analyse der ermittelten Signaturgene wurden diese öffentlich verfügbaren Gen-Sets verwendet, die über die *Molecular Signatures Database* (MSigDB)²⁹ zur Verfügung gestellt werden und in acht Hauptkategorien mit weiteren Unterkategorien eingeteilt sind (vgl. Tab. 2.6). Die Analyse der in dieser Arbeit identifizierten Gensignaturen erfolgte unter Verwendung der öffentlich verfügbaren Gen-Sets der Kategorien C2, C5 und C6. Die Gen-Sets dieser drei Kategorien beinhalten annotierte und validierte Mengen von Genen zu diversen biologischen Prozessen (C2), wie beispielsweise Differenzierung, Wachstum, Migration, Invasivität etc., Einteilungen der Gene entsprechend der verfügbaren GO-Termini (C5) sowie veränderten Prozessen, die im Zusammenhang mit Krebserkrankungen annotiert sind (C6). Für jeder der hier identifizierten Gensignaturen wurden die Gene des *Up-* und *Down-Sets* der Signatur mit ihren logFC-Werte, aus der differentiellen Expressionsanalyse zwischen den Vergleichen der Stadien, zu einer Liste zusammengefügt. Anschließend wurden die Gene dieser Liste absteigend nach ihren logFC-Werte geordnet, so dass eine Rangliste aller Gene der Signatur erzeugt wurde. Diese Ranglisten wurden anschließend mit den genannten öffentlichen Gen-Sets der drei Kategorien (C2, C5, C6) unter Verwendung des Programms „GSEA“ (Version 2.2.1)³¹ analysiert.

²⁸http://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/FAQ#Can_I_use_GSEA_to_analyze_SNP.2C_SAGE.2C_ChIP-Seq_or_RNA-Seq_data.3F

²⁹<http://software.broadinstitute.org/gsea/msigdb/index.jsp>

³⁰Gene Ontology (GO) ist eine biomedizinische Ontologie zur Vereinheitlichung der Darstellung von Genen und Genprodukten.

³¹www.broadinstitute.org/gsea

Ziel der Analyse mittels GSEA Methode ist es, die biologischen Zusammenhänge, der in den hier identifizierten Signaturen enthaltenen Gene, anhand beschriebener und annotierten Daten aus den öffentlichen Gen-Sets zu ermitteln. Dies dient im Weiteren der Betrachtung der hier gewonnenen Signaturen im Hinblick auf bekannte und eventuell unbekannt charakteristische Unterschiede im Transkriptom der analysierten Tumorproben der Neuroblastomstadien. Innerhalb der GSEA Methode wird geprüft, welche Gene der Gen-Sets aus den öffentlichen Daten mit den Genen der erzeugten Ranglisten (der hier identifizierten Signaturen) übereinstimmen und wo diese sich entlang der Rangliste befinden [237]. Da sowohl die hoch- als auch herunterregulierten Gene der DE-mRNAs in den Ranglisten enthalten sind (Up- und Down-Set der Signatur), kann gleichzeitig geprüft werden, welche der öffentlichen Gen-Sets anhand der positiv und negativ regulierten Gene der Rangliste überrepräsentiert werden. Eine Überrepräsentierung bedeutet hier, dass die übereinstimmenden Gene am Anfang oder am Ende der Rangliste mit den Genen des betrachteten Gen-Sets übereinstimmen. Dabei werden Gen-Sets positiv regulierter Gene am Anfang der Rangliste und somit - entsprechend der Ordnung der Rangliste - durch hochregulierte Gene (positiver Fold Change) überrepräsentiert. Gen-Sets negativ regulierter Gene hingegen werden durch die Gene am Ende der Rangliste und somit durch herunterregulierte Gene (negativer Fold Change) der Signatur überrepräsentiert. Die zugrunde liegende Annahme der GSEA Methode beruht darauf, dass bei einer Überrepräsentierung eines Gen-Sets, durch die Gene einer Rangliste, eine Regulierung des assoziierten Prozesses des Gen-Sets vorliegt, wenn Gene der Rangliste mit dem untersuchten Gen-Sets übereinstimmen und diese entweder am Anfang (positive Regulation) oder am Ende (negative Regulation) der Rangliste auftreten [237]. Eine zufällige Verteilung der übereinstimmenden Gene entlang der Rangliste hingegen weist auf keine klare Regulation der Gene und auf keine eindeutige Assoziation zum Gen-Set und des damit verbundenen biologischen bzw. funktionellen Zusammenhanges hin. Um zu bestimmen, wie stark ein Gen-Sets überrepräsentiert wird und ob dieses einer positiven oder negativen Regulation der Gene unterliegt, wird ein sogenannter *Enrichment Score* (ES) auf Basis eines gewichteten Kolmogorov-Smirnov-Tests [34] berechnet. Der ES wird entlang der Rangliste, vom Gen mit dem größten logFC-Wert bis zum Gen mit dem kleinsten logFC-Wert, aufsummiert, wobei eine Übereinstimmung des Vorkommens des aktuellen Genes der Rangliste mit den Genen des Gen-Sets zur Erhöhung, ein Nichtvorhandensein des Genes im Gen-Sets zur Reduzierung des ES führt. Als Resultat ergibt sich der ES aus der größten erreichten Differenz zum Wert Null,

der durch seinen Wert im Bereich $[-1, 1]$ und sein Vorzeichen angibt, zu welchem Grad das untersuchte Gen-Sets am Anfang oder am Ende einer gegebenen Rangliste anhand der positiv oder negativ regulierter Gene überrepräsentiert ist. Der ES ermöglicht somit eine Aussage darüber, zu welchem Grad die Gene der Rangliste das jeweilige Gen-Set überrepräsentieren und der biologische Annotation des Gen-Sets entsprechen. Zusätzlich zum ES wird ein normalisierter Enrichment Score (NES) ermittelt, in dessen Berechnung die Größe des Gen-Sets mit einbezogen wird, und es wird aufgrund von multiplen Tests ein FDR-Wert zur Korrektur des Signifikanzniveaus für den NES ermittelt [237]. Alle Gene der identifizierten Signaturen in dieser Arbeit wurden mittels GSEA Methode hinsichtlich annotierter biologischer und funktionaler Zusammenhänge, auf Basis der öffentlichen Gen-Sets der drei verwendeten Kategorien C2, C5 und C6, getestet. Exemplarisch ist für das Ergebnis dieser Analyse ein Enrichment Plot in Abb. 2.15 dargestellt, der den graphisch Verlauf der sich summierenden Werte, über die Gene einer in dieser Arbeit erzeugten Rangliste, zeigt. Als Grundlage dient hier eines der öffentlichen Gen-Sets der Kategorie C2, welches Gene beinhaltet, die mit pädiatrischen Tumoren im Zusammenhang stehen und aus dem Vergleich von acht unterschiedlichen pädiatrischen Tumoren (auch dem Neuroblastom) mit Normalgeweben stammen [267]. Der resultierende ES-Wert, für die mit diesem Gen-Set analysierte Rangliste, hat einen Wert von 0,68. Dieser weist zum einen, durch seinen von Null verschiedenen Wert, auf einen Zusammenhang der Ranglistengene zu pädiatrischen Tumoren hin. Zum anderen weist der positive ES darauf hin, dass die Gene der Rangliste, die mit dem Gen-Set übereinstimmen, am Anfang der Rangliste auftreten und die Gene der gegebenen Rangliste demnach positiv mit dem Gen-Set bzw. der damit verbundene Annotation korreliert sind. Eine fehlende Übereinstimmung sowie eine zufällige Verteilung von übereinstimmenden Genen würde zu einem geringen ES um den Wert Null führen, wobei es keine allgemeingültige Schwelle für den ES gibt, die ein Mindestmaß der Repräsentation festlegt. Ein ES-Wert von $\geq 0,5$ bzw. $\leq -0,5$ hat sich hierbei jedoch als hilfreiche Orientierung erwiesen, wobei die Standardwerte für den p-Wert von 0,05 und einem FDR-Wert von 0,25 nach unten korrigiert werden können, so eine spezifischere Auswahl von stark signifikanten Gen-Sets erfolgen soll.

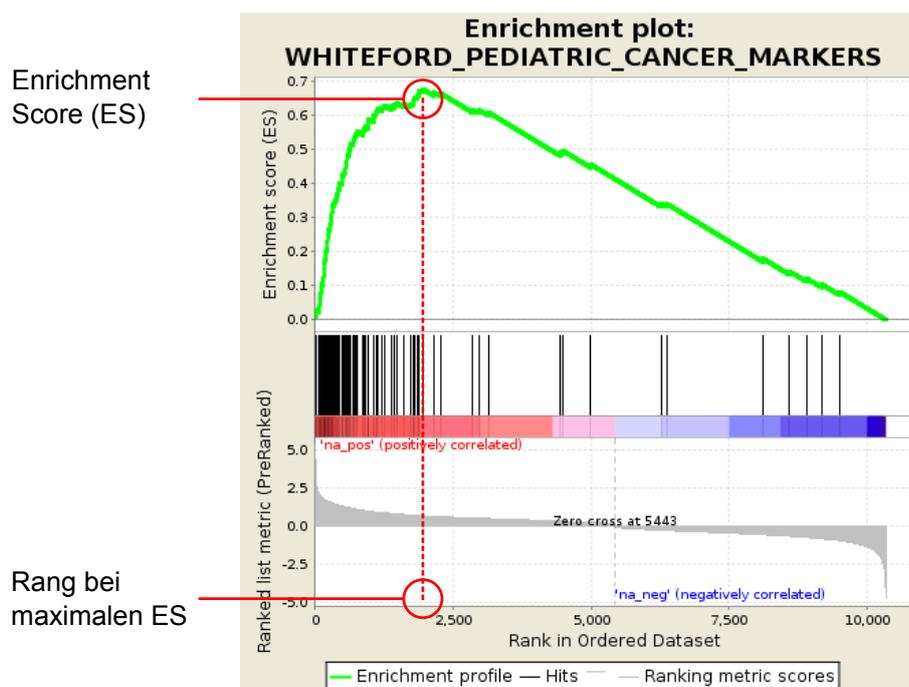


Abbildung 2.15.: Graphische Darstellung des Verlaufes des aufsummierenden Enrichment Scores (ES) aus der Analyse einer in dieser Arbeit erstellten Rangliste aus 10351 differentiell exprimierten Genen des Vergleiches 1 – MYCN mit einem Gen-Set der Kategorie C2. In grün ist der Verlauf des ES entlang der einzelnen Gene der Rangliste dargestellt. Die schwarz gekennzeichneten Linien markieren das Auftreten eines Gens sowohl in der Rangliste als auch im öffentlichen Gen-Set (105 Stück), dessen enthaltenen Gene im Zusammenhang mit verschiedenen pädiatrischen Tumoren ermittelt wurden [267]. Die Sortierung der Gene der Rangliste, nach ihren logFC-Werten, ist durch den Verlauf des farblichen Balkens von Rot nach Blau dargestellt, wobei Rot die Gene mit positivem logFC-Wert und Blau die mit negativen logFC kennzeichnen. Der Übergang zwischen hoch- und runterregulierten Genen, an denen ein Vorzeichenwechsel des Fold Change stattfindet, ist durch den Nulldurchgang (Zero cross) markiert, der in diesem Beispiel bei Position 5443 liegt. Der für die Rangliste von differentiell exprimierten Genen ermittelte ES weist einen Wert von 0,68 an Position 1952 der Rangliste auf (rote Linie). Der ES deutet somit auf eine positive Korrelation der hochregulierten Gene der Rangliste aus dem Vergleich 1 – MYCN der Neuroblastomproben im Zusammenhang mit pädiatrischen Tumoren hin.

Die verwendeten Parameter, die für die Analyse der in dieser Arbeit identifizierten Gensignaturen der Neuroblastomstadien genutzt wurden, sind in Tab. 2.7 aufgelistet. Die identifizierten Gensignaturen wurden als Ergebnis dieser Arbeit im Genmatrix Format (GMT) gespeichert, die daraus erstellten Ranglisten zur Analyse der Gensignaturen wurden im Ranglisten Format (RNK) erstellt. Beide Formate³² können zur Verwendung der identifizierten Signaturen und zu deren Analyse direkt in das GSEA Programm importiert werden.

Parameter	Wert	Funktion
xtools.gsea	GseaPreranked	Analyse einer vordefinierten Rangliste
-gmx	C2, C5, C6	ausgewählte, öffentlich verfügbare Gen-Sets
-nperm	1000	Anzahl Permutationen
-collapse	false	Gensymbole kollabieren
-scoring_scheme	classic	Gewichtungsmethode
-set_min	15	minimale Größe des Gen-Sets
-set_max	500	maximale Größe des Gen-Sets
-norm	meandiv	Normalisierungsmethode, Erzeugen des norm. ES (NES)

Tabelle 2.7.: *Verwendete Parameter der GSEA Methode zur Analyse der Expressionsdaten des Neuroblastoms.*

2.2.5. microRNA-Targetanalyse

Die Untersuchung der differentiell exprimierten microRNAs und deren möglichen Zielgene in den vorliegenden Neuroblastomdaten, stellt eine der weiteren Analysen innerhalb dieser Arbeit dar. Grundlage für diese Analyse bilden die in Abschnitt 2.2.3 ermittelten DE-microRNAs und -mRNAs der differentiellen Expressionsanalyse zwischen den definierten Vergleichen der klinischen Stadien. Bei microRNAs handelt es sich um kurze einzelsträngige RNAs mit einer Länge von ca. 22 Nukleotiden, die eine wesentliche Rolle in der posttranskriptionellen Regulation einnehmen. Durch die Bindung an die mRNAs ihrer Zielgene (microRNA-Targets), können sie die Expression der Targets auf vielfältige Weise beeinflussen, beispielsweise durch translationelle Repression anhand der Bindung der microRNA an den 3'-UTR der mRNA oder den Abbau bzw. die Destabilisierung der mRNA [85]. Die Analyse der microRNA-Targets innerhalb dieser Arbeit dient der Identifizierung von mRNAs Protein-codierender Gene im Neuroblastom, die von den veränderten microRNAs der Tumorstadien beeinflusst werden könnten und

³²http://www.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats

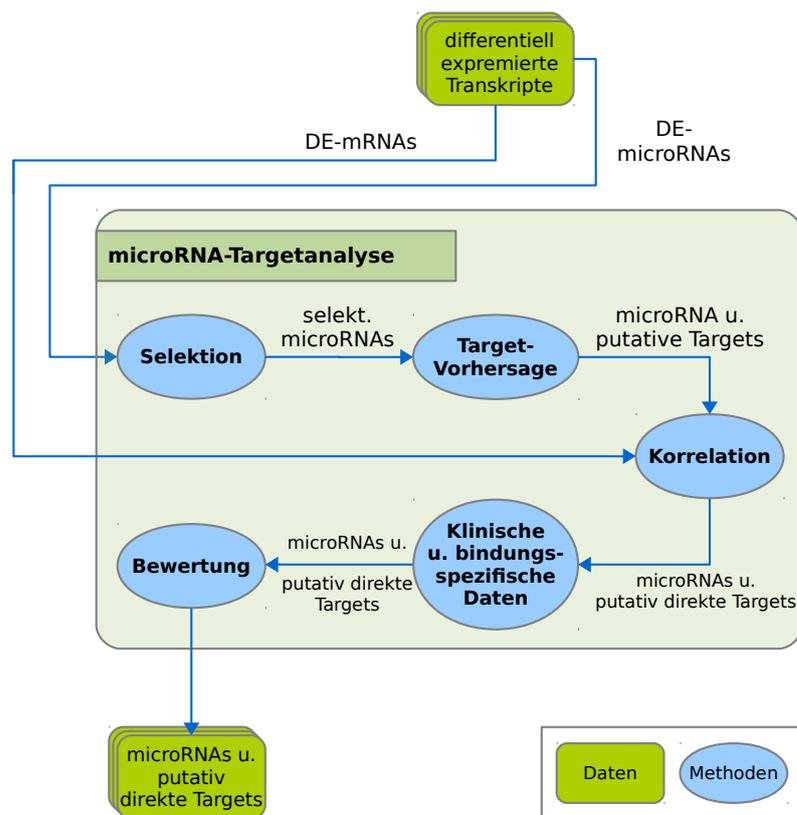


Abbildung 2.16.: Ausschnitt des Datenflussdiagramms der Transkriptomanalyse (vgl. Abb. 2.14) zur Darstellung der in dieser Arbeit entwickelten microRNA-Targetanalyse. Die Methode setzt sich aus mehreren Teilen zusammen, die sequentiell auf den resultierenden Daten der differentiellen Expression ausgeführt werden. Anhand der microRNA-Targetanalyse werden für die Vergleiche zwischen den klinischen Stadien der hier analysierten Neuroblastomproben differentiell veränderte microRNAs sowie deren putative direkte mRNA Targets ermittelt.

somit Hinweise auf putativ neue microRNA-Target Beziehungen, im Zusammenhang mit dem Neuroblastom und dessen Stadien, ermöglichen könnten. Ein bedeutender Aspekt dieser Arbeit liegt darin, dass sowohl das Material für die RNA- als auch für die microRNA-Sequenzierung zusammen aus den jeweiligen Tumorproben gewonnen wurden. Somit können auf Grundlage der Sequenzierungen die transkriptomischen Veränderungen der klinischen Stadien auf mRNA- und microRNA-Ebene miteinander korreliert werden. Die entwickelte Methode der microRNA-Targetanalysen besteht aus der Selektion der microRNAs anhand der differentiellen Expression mit anschließender Vorhersage von putativen microRNA-Targets. Daran schließt sich die eigentliche Analyse der Targets an, die die Korrelation mit den RNA-Sequenzierungsdaten und die Überlebenszeitanalyse beinhaltet. Weiterhin kommt es zu einer Bewertung der ermittelten

microRNAs und ihrer putativen direkten Targets mithilfe bindungsspezifischer Daten und der Einbeziehung der in dieser Arbeit ermittelten Expression der microRNAs und Targets aus den untersuchten Neuroblastomproben (vgl. Abb. 2.16).

Selektion von microRNAs

Im ersten Schritt der Targetanalyse werden signifikant veränderte microRNAs zwischen den klinisch relevanten Stadien der Neuroblastomdaten ausgewählt, um die Veränderungen des Transkriptoms der Tumordaten auf microRNA-Ebene zu ermitteln und in weiteren Analysen zu untersuchen. Basierend auf der differentiellen Expression der microRNA-Seq. Daten erfolgte die Selektion der microRNAs anhand der eingeteilten Vergleiche der klinischen Stadien (vgl. Tab. 2.5). Dabei wurden für jeden der Vergleiche die DE-microRNAs selektiert, die einen mittleren CPM-Wert³³ von 100 aufwiesen sowie eine signifikante 2-fache Veränderung zwischen den untersuchten Stadien eines Vergleiches, bei einem $FDR \leq 0,05$ (vgl. 2.2.3) und einem $|\logFC| \geq 1$. Diese Werte wurden gewählt, um die nachfolgende Target-Vorhersagen auf microRNAs zu beziehen, die eine signifikante Veränderung zwischen den Stadien zeigen und ein gewisses Mindestmaß an Expression bzw. Veränderung innerhalb der Proben aufweisen. Weder in aktueller Literatur, noch durch eigene Analysen am Institut, konnten bisher allgemeingültige Werte für eine minimale Expression oder deren Veränderung in der differentiellen Expression ermittelt und festgelegt werden. Die angewendeten Werte benachteiligen in gewissem Maße wenig abundante microRNAs oder solche mit geringer Veränderung, führen aber zu einer Selektion sicherer Kandidaten. Transkripte mit geringeren Abundanzen oder logFC-Werten bewegen sich hingegen meist an der Grenze der festgelegten Signifikanz und können, gerade bei RNA-Sequenzierungsdaten, starken Varianzen unterliegen, wie bereits im Abschnitt 2.2.3 genauer erläutert wurde.

microRNA-Target Vorhersage

Die sich anschließende Vorhersage putativer Targets basiert auf den selektierten microRNAs unter Verwendung des R-Pakets „multiMiR“ [215] (Version 1.0.1)³⁴. Dieses ermöglicht die Identifizierung von vorhergesagten und validierten Targets unter der Nutzung von 14 externen Datenbanken. Von diesen stammen drei aus experimentell validierten, acht aus

³³Counts per Million (CPM) gibt die Anzahl der annotierten microRNA Reads, normalisiert auf eine Million Reads, an.

³⁴<http://http://multimir.ucdenver.edu/>

Parameter	Wert	Funktion
org	human	Organismus
table	all	Datenbankauswahl
summary	true	Zusammenfassung der Ergebnisse
predicted.cutoff.type	p	Art der Eingrenzung, prozentual (p) oder Anzahl (n)
predicted.cutoff	100	Eingrenzung der Vorhersage
predicted.site	all	Art der Konservierungsinformation

Tabelle 2.8.: Verwendete Parameter im R-Paket *multiMiR* zur Vorhersage von putativen Targets der ermittelten differentiell exprimierten *microRNAs* des Neuroblastoms.

computergestützten Vorhersage- und drei weitere aus krankheits- bzw. medikamentenbezogenen Datenbanken. Die verwendeten Datenbanken zur Vorhersage von *microRNA*-Targets weisen Ähnlichkeiten in der Methodik auf und nutzen bekannte Bindungseigenschaften zwischen *microRNA* und Target mRNA. Eigenschaften, die von allen der genutzten Datenbanken untersucht werden, sind Konservierungsinformationen, die freie Bindungsenergie sowie die Zugänglichkeit der Bindungsstelle des Targets für die *microRNA*. Grundlage zur Bestimmung dieser Eigenschaften ist die Sequenzinformation der *microRNAs* auf Basis der Nukleotide 2 bis 7, welche im Zusammenhang mit *microRNA*-Target Vorhersagen auch als *microRNA Seed-Sequenz* bezeichnet wird [61, 195]. Trotz Ähnlichkeiten in der Vorhersagemethodik, führt die Verwendung der einzelnen Datenbanken zu teilweise unterschiedlichen Ergebnissen, bezogen auf die Anzahl der vorhergesagten Targets der untersuchten *microRNAs* als auch auf die Übereinstimmung der vorhergesagten Targetgene aus den Datenbanken [2, 61]. Zudem kommt kein einheitliches Bewertungssystem der Vorhersagen zur Anwendung, so dass die ermittelten Bewertungen, anhand von sogenannten *Scores* der einzelnen Datenbanken, nicht miteinander verglichen werden können. Da die Ergebnisse der Datenbanken demnach, hinsichtlich ihrer Vorhersagequalität, nicht bewertet werden können und keine Festlegung auf nur eine oder eine kleine Auswahl von Datenbanken erfolgen kann und soll, wurden alle in *multiMiR* verfügbaren Datenbanken zur Vorhersage der putativen *microRNA*-Targets verwendet. Als Grundlage für die Eingabe dienen die in dieser Arbeit selektierten hoch- bzw. herunterregulierten DE-*microRNAs* der eingeteilten Vergleiche der differentiellen Expression. Das Ergebnis der Vorhersage mittels *multiMiR* resultiert in einer Liste der vorhergesagten *microRNA*-Targets. Diese Liste besteht aus den Gen-Namen und -IDs der Targets und der selektierten *microRNAs* sowie aus den Summen der Anzahlen von vorhergesagenden Datenbanken aus den zusammengefassten drei experimentell validierten, den acht computergestützten und den drei medikamentenbezogenen Datenbanken.

Weiterhin sind, wenn eine Vorhersage stattfand, die ermittelten Scores der Vorhersage der einzelnen Datenbanken angegeben. Die Parameter, die bei der Vorhersage mittels multiMiR verwendet wurden, sind in Tab. 2.8 dargestellt.

Korrelation mit RNA-Sequenzierungsdaten

Im Anschluss an die Vorhersage der microRNA-Targets fand eine Korrelation (im Sinne von funktionellem Zusammenhang) der vorhergesagten Targetgene mit den differentiellen Daten des RNA-Seq. statt. An dieser Stelle sei noch einmal darauf hingewiesen, dass es das Ziel dieser Arbeit ist, putativ direkte microRNA-Targets in den Neuroblastomstadien zu identifizieren. Direkte Targets der microRNAs würden somit negativ reguliert werden und sollten bei einer Hochregulierung der microRNA eine entsprechende Herunterregulierung durch die microRNA-vermittelte Regulation erfahren. Auch wenn diese Annahme eine indirekte Regulierung des Targets nicht ausschließen kann, so ist sie doch notwendig, um *in silico* auf eine putativ direkte microRNA-Target Beziehung aus den hier vorliegenden Daten hinzuweisen. Die Bestätigung einer tatsächlichen direkten Bindung aus diesen hier gewonnenen Hinweisen kann nur auf experimentellem Wege in weiterführenden Analysen erfolgen. Die Korrelation für die hoch- und herunterregulierten DE-microRNAs in der Targetanalyse dieser Arbeit wurden getrennt vorgenommen, entsprechend der differentiellen Vergleiche der klinisch relevanten Neuroblastomstadien. Für die hochregulierten DE-microRNAs eines Vergleiches wurden deren ermittelte Targets aus den Tabellen der Analyse auf differentiell exprimierte mRNAs identifiziert, die zum einen eine differentielle Expression aufwiesen ($FDR \leq 0,05$) und sich in der Veränderung ihrer Expression zwischen den betrachteten Stadien umgekehrt zu der Veränderung der microRNA verhielten, somit einen negativen Fold Change aufwiesen. Für herunterregulierte microRNAs wurden dementsprechend die DE-mRNAs aus den vorhergesagten Targets identifiziert, die einen positiven Fold Change aufwiesen. Aufgrund der erwähnten unterschiedlichen Vorhersageergebnisse der genutzten Datenbanken, wurde festgelegt, dass jedes Target von mindestens zwei der Datenbanken vorhergesagt werden sollte. Die resultierenden Tabellen der microRNA-Target Vorhersage wurden dahingehend verändert, dass alle microRNA-Target Beziehungen, die diese Korrelationskriterien nicht erfüllen, entfernt wurden. Zusätzlich wurden die Tabellen um die Informationen aus der differentiellen Expression, in Form der logFC- und FDR-Werte der Targetgene, erweitert.

Einbeziehung klinischer Daten

Auf Basis der klinischen Patientendaten, der in dieser Arbeit analysierten Neuroblastomproben, wurden sogenannte Überlebenszeitanalysen (*Survival Analysis*) durchgeführt. Bei dieser Analyse kommen klinische Daten über den Verlauf der Behandlung bis zum Eintreffen eines definierten Ereignisses in Kombination mit weiteren zu untersuchenden Daten, wie beispielsweise verschiedene Therapieansätze, Medikamente, Medikamentendosen, Expressionsdaten etc., zur Anwendung. Solche definierten Ereignisse können z. B. das Abklingen einer Infektion, die Heilung einer Krankheit oder der Tod des Patienten sein. Für fast alle analysierten Neuroblastomproben liegen diese klinischen Daten in Form der Überlebenszeit von der Diagnose bis zum möglichen Tod der Patienten vor (vgl. Tab. B.2 im Anhang). Dabei muss das Ereignis, dass ein Patient bedingt durch den Tumor stirbt, am Ende der Beobachtungszeit nicht eingetreten sein. Mithilfe der klinischen Daten lässt sich die genannte Überlebenszeitanalyse realisieren, die im hier vorliegenden Fall des Neuroblastoms dazu verwendet wurde, die Überlebenswahrscheinlichkeiten der Patienten in Abhängigkeit der Sequenzierungsdaten zu berechnen. Grundlage der Berechnung bilden, neben den klinischen Informationen, die Expressionswerte der aus den Korrelationen ermittelten microRNAs und ihrer Target mRNAs. Die Überlebenszeitanalysen erlaubt es, basierend auf den Expressionswerten, Aussagen darüber zu treffen, ob sich die vorliegende Expression der mRNA oder microRNA, positiv oder negativ auf die Überlebenswahrscheinlichkeit, des an dem Tumor erkrankten Patienten, auswirkt. Die Einbeziehung der klinischer Daten in Form der Überlebenszeitanalyse dient zusammen mit den im weiteren Verlauf beschriebenen Analyseschritten einer Bewertung der ermittelten microRNA-Target Beziehungen, worauf im letzten Abschnitt detaillierter eingegangen wird.

Zur Berechnung der Überlebenswahrscheinlichkeiten, basierend auf den Expressionen der aus der Korrelation stammenden microRNA-Target Beziehungen, wurde in dieser Arbeit die Kaplan-Meier Methode angewendet [103]. Der Idee dieser Methode liegt darin begründet, dass die Beobachtungsintervalle nicht fest vorgegeben, sondern die Ereignisse selbst die Zeitintervalle definieren. Um die Überlebenswahrscheinlichkeit für einen bestimmten Zeitpunkt zu erhalten, werden zunächst die bedingten Wahrscheinlichkeiten für jedes Zeitintervall berechnet, die angeben, wie wahrscheinlich es ist, dass ein Patient ein Zeitintervall überlebt, wenn er schon bis zum Beginn des Intervalls überlebt hat. Das Produkt der bedingten Wahrscheinlichkeiten bildet die Gesamtwahrscheinlichkeit, aus der ermittelt werden kann, wie wahrscheinlich es ist einen

definierten Zeitpunkt zu überleben. Dabei können zensierte Ereignisse, durch Todesfälle oder das Ausscheiden des Patienten aus der Studie, auftreten. Die Überlebenswahrscheinlichkeiten können zwischen verschiedenen verfügbaren Bedingungen miteinander verglichen werden, wie beispielsweise in dieser Arbeit der Vergleich zwischen einer hohen bzw. niedrigen Expression eines Gens. Als statistischer Test für die Vergleiche kommt in dieser Arbeit der Log-Rang-Test [16] zur Anwendung, welcher ein Standardverfahren in der Überlebenszeitanalyse darstellt. Die Überlebenszeitanalyse der hier analysierten Neuroblastomproben wurde unter Verwendung der Kaplan-Meier Methode der *R2 Genomics Analysis and Visualization Platform*³⁵ durchgeführt, auf der die Neuroblastomdaten aus dieser Arbeit bereits hinterlegt sind. Unter Verwendung der hier vorliegenden Expressionsdaten erfolgte durch die Kaplan-Meier Methode eine Trennung aller Proben in zwei Gruppen. Die Auftrennung erfolgte ausschließlich anhand der Expressionsdaten einer gegebenen mRNA oder microRNA, unabhängig von den klinischen Stadien, um eine generelle Aussage zur Überlebenswahrscheinlichkeit im Neuroblastom für die untersuchte mRNA bzw. microRNA zu ermitteln. Eine der Gruppen umfasst dabei alle die Proben, in denen die mRNA/microRNA eine hohe Expression aufweist, während die zweite Gruppe die Proben umfasst, in denen eine niedrige Expression vorliegt. Die Definition von hoch bzw. niedrig kann auf verschiedenen Wegen, beispielsweise über den Median, Mittelwert oder über einen manuellen Schwellwert, festgelegt oder über eine automatische Analyse der Expressionswerte ermittelt werden. Alle Proben, deren Expressionswerte für die mRNA/microRNA unterhalb dieser Schwelle liegen, werden der Gruppe *low* mit niedrigen, alle Proben deren Werte darüber liegen, werden der Gruppe *high* mit hohen Expressionswerten zugeordnet. Bei der in dieser Arbeit durchgeführten Analyse wurde dieser Schwellwert mithilfe des in R2 enthaltenem *KaplanScans*³⁶ auf automatischem Wege bestimmt. Die angewendete KaplanScan-Methode trennt die Proben anhand ihrer Expressionswerte in die zwei Gruppen (low/high) und testet anschließend die Nullhypothese, wie wahrscheinlich es ist, dass sich die Kurven (als Kaplan-Meier Kurven bezeichnet) der Überlebenswahrscheinlichkeiten in Abhängigkeit der Überlebenszeit nicht voneinander unterscheiden. Die Todesfälle der Patienten somit in zufälliger Reihenfolge gruppenunabhängig und damit unabhängig von der Expression des untersuchten Gens auftreten. Der Test erfolgt mithilfe des erwähnten Log-Rang-Tests, welcher ein nichtparametrisches Verfahren darstellt und als Standardtest in

³⁵<http://r2.amc.nl>

³⁶<http://r2-tutorials.readthedocs.io/en/latest/index.html>

der Überlebenszeitanalyse verwendet wird. Bei diesem Test wird für die einzelnen Zeitpunkte die Summe der beobachteten und der erwarteten Todesfälle in den Gruppen berechnet und zwischen den Gruppen, basierend auf der Log-Rang-Teststatistik, verglichen [16]. Mittels Log-Rang-Test ermittelt der KaplanScan den signifikantesten Expressionswert, der zu einer Trennung der Gruppen führt, bei dem sich die Kaplan-Meier Kurven beider Gruppen am stärksten voneinander unterscheiden. Aus dem Ergebnis der Kaplan-Meier Methode resultiert die Gruppe (low oder high), bei der die geringere Überlebenswahrscheinlichkeit in Abhängigkeit der Expression vorliegt und somit eine schlechte Prognose für den Verlauf der Krankheit zu erwarten ist. Als minimale Probenanzahl für die Gruppen wurde der Standardwert von mindestens acht Proben pro Gruppe beibehalten.

Bei den in dieser Arbeit durchgeführten Überlebenszeitanalysen wurden, unabhängig voneinander, die aus der Korrelation ermittelten DE-microRNAs sowie deren putativ direkten DE-mRNA Targets, für die neun differentiellen Vergleiche der klinischen Stadien, untersucht. Als Ergebnis erhält jede der analysierten microRNAs bzw. mRNAs zwei zusätzliche Einträge in der microRNA-Target Tabelle der Korrelation (vgl. vorheriger Abschnitt). Ein Eintrag gibt die Gruppe (low oder high) an, die zu einer geringen Überlebenswahrscheinlichkeit führt und somit die Information, ob eine niedrige oder hohe Expression zu einer schlechten Prognose der Erkrankung führt. Weiterhin wird der FDR-Wert des Bonferroni-korrigierte Signifikanzniveaus aus dem Vergleich der Überlebenswahrscheinlichkeit der zwei Gruppen (low und high) angegeben. Darüber hinaus kann eine graphische Darstellung generiert werden, die die Überlebenswahrscheinlichkeiten in Form einer Kaplan-Meier Kurve darstellt. Abbildung 2.17 veranschaulicht die Überlebenswahrscheinlichkeit, in Form der berechneten Kaplan-Meier Kurven, für die in dieser Arbeit analysierten Neuroblastomproben, unter Einbeziehung der ermittelten Expression des *MYCN* Gens. Die Kaplan-Meier Kurven zeigen den Verlauf der Überlebenswahrscheinlichkeit für einen Zeitraum von fast 16 Jahren (Abb. links) bei einer niedrigen (rote Kurve) bzw. hohen (blaue Kurve) Expression des *MYCN* Gens, unter Aufteilung der 61 Neuroblastomproben in 47 bzw. 13 Proben. Die Einteilung der Proben in die Gruppen low und high, die für eine niedrige bzw. hohe Expression stehen, wurden wie beschrieben mittels

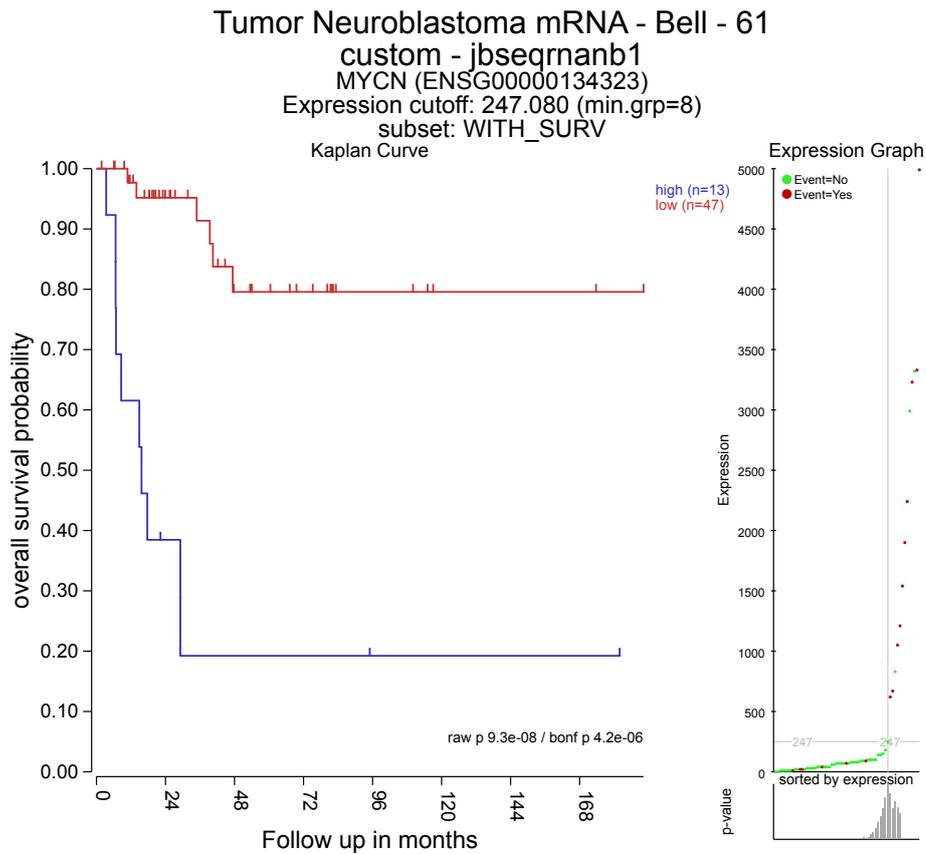


Abbildung 2.17.: Kaplan-Meier Kurve der Überlebenszeitanalyse aus den 61 Proben³⁷ des RNA-Seq., anhand der mRNA Expression des MYCN Gens. (Links) Überlebenswahrscheinlichkeit in den unterteilten Gruppen (low oder high), basierend auf den Expressionswerten der untersuchten Proben. Die blaue Kurve (bestehend aus 13 Proben) kennzeichnet eine schlechte Überlebenswahrscheinlichkeit, bei hoher MYCN Expression. Die rote Kurve hingegen kennzeichnet eine hohe Überlebenswahrscheinlichkeit, bei einer geringen Expression des Gens, basierend auf 47 Proben. (Rechts) Graphische Darstellung der Probenaufteilung anhand der in dieser Arbeit ermittelten MYCN Expressionswerte des RNA-Seq. unter Verwendung des KaplanScans zur automatischen Schwellwertbestimmung. Proben deren Expression unterhalb der Schwelle liegen, werden der Gruppe low (rote Kurve), die deren Expression oberhalb der Schwelle liegen, der Gruppe high (blaue Kurve) zugeordnet.

³⁷In die Analyse gehen nur 60 der 61 Proben ein, da für einen Patienten keine Informationen zu den klinischen Daten vorhanden ist (vgl. Probe JB_48 in Tab. B.2 im Anhang).

KaplanScan von R2 vorgenommen, welcher eine automatische Schwelle der Expression zur Trennung der Gruppen ermittelt (Abb. rechts). Anhand der Kaplan-Meier Kurven lässt sich die mittlere Überlebenszeit (*overall survival probability* = 0,50) entnehmen, die in diesem Beispiel bei einem Wert von 16 Monaten erreicht und es ist ersichtlich, dass eine geringe Überlebenswahrscheinlichkeiten mit einer hohen *MYCN* Expression korreliert.

Als Ergebnis der in dieser Arbeit durchgeführten Überlebenszeitanalysen wurden die Korrelationstabellen erweitert und als neue Tabellen gespeichert. Jede enthaltene microRNA-Target Beziehung aus der Korrelation erhält zusätzlich, für die enthaltene microRNA und deren putativ direktes Target, die erwähnten Einträge der Expressionen (low oder high), die zu einer schlechten Prognose führen, sowie den FDR-Werten, aus dem Vergleich der ermittelten Überlebenswahrscheinlichkeiten.

Einbeziehung bindungsspezifischer Daten

Neben den in der Vorhersage durch die Datenbanken genutzten Bindungseigenschaften, wie z.B. Sequenzinformationen und freie Bindungsenergie, wurde innerhalb dieser Arbeit die generelle Anzahl ermittelter microRNA Bindestellen für ein ermitteltes putativ direktes Target betrachtet. Wie einleitend erwähnt beruht die microRNA vermittelte Regulierung auf der Bindung der microRNA Seed-Region an die komplementäre Sequenz im 3'-UTR der Target mRNA. Diese Bindestelle wird als *MRE* (*microRNA recognition/response element*) bezeichnet [7, 64, 229]. Die Datenbanken zur microRNA-Target Vorhersage ermitteln die putativen Bindungsstellen der microRNA basierend auf der Sequenzinformation der Target mRNAs. Diese Anzahl, an möglichen MREs für die einzelnen Targets, geht in dieser Arbeit als zusätzliches Bewertungskriterium der hier ermittelten microRNA-Target Beziehung ein. Die hierzu verwendeten Daten stammen aus „TargetScan“ [127] (Version 7.1), einer öffentlichen Datenbank zur microRNA-Target Vorhersage, welche zudem in der hier verwendeten multiMiR Version genutzt wurde. TargetScan stellt Tabellen zur Verfügung, die die Positionen der Bindestellen in den vorhergesagten Targets sowie deren Überlagerung mit den 3'-UTR Koordinaten der Targets beinhalten³⁸. Eine Zusammenfassung dieser Informationen erfolgte durch Dr. M. Glaß³⁹. Darauf basierend wurde, entsprechend der in dieser Arbeit benötigten Daten, eine Tabelle in der Form extrahiert, so dass für jedes gelistete Target der Datenbank der

³⁸http://www.targetscan.org/cgi-bin/targetscan/data_download.vert71.cgi

³⁹Institut für Molekulare Medizin, Martin-Luther-Universität Halle-Wittenberg

Gen-Name, die Gen-ID des Transkripts mit dem längsten 3'-UTR, die Anzahl der vorhergesagten MREs sowie die auf die 3'-UTR Länge normalisierte Anzahl an MREs enthalten ist (siehe Verweis auf die Daten im Anhang F.1).

Bewertung der ermittelten microRNA-Target Beziehungen

In Kombination stellen die differentielle Expression, die Targetvorhersage sowie die klinischen und bindungsspezifischen Daten, die Grundlage zur Bewertung der hier entwickelten microRNA-Targetanalyse dar. Die finale Bewertung der in dieser Arbeit ermittelten microRNA-Target Beziehungen integriert alle bisherigen Informationen der Targetanalyse und erweitert die zuvor generierten Ergebnistabellen. Die Notwendigkeit eines Bewertungsmaßes liegt darin begründet, dass die Ergebnisse der Vorhersagedatenbanken, wie bereits erwähnt, unterschiedliche Resultate liefern und zum anderen weitere Faktoren in Betracht gezogen werden sollen, die zusätzliche Hinweise auf eine putativ direkte Bindung erkennen lassen. Nichtsdestotrotz können anhand der hier durchgeführten *in silico* Analysen nur Hinweise auf eine direkte Regulierung eines Targets durch eine microRNA geben werden. Eine reale Bindung bzw. Regulierung kann jedoch nur durch weiterführende laborbasierte Experimente bestätigt werden, wie beispielsweise durch *miTRAP* Bindungsexperimente [22]. Auf der genannten Grundlage erfolgte eine unabhängige Bewertung der identifizierten microRNA-Target Beziehungen zwischen den Stadien der Neuroblastomproben. Die Bewertung in Form eines Scores soll eine Aussage darüber treffen, wie wahrscheinlich die ermittelten putativen Bindungen innerhalb der betrachteten differentiellen Vergleiche, also dem Vergleich zwischen zwei Stadien des Neuroblastoms, ist. Dabei soll ein hoher Score vorliegen, wenn alle aus den Analysen stammenden Daten auf eine putativ direkte Bindung zwischen microRNA und Target hinweisen, somit eine Regulierung des Targets durch die microRNA wahrscheinlich ist. Für die hier durchgeführten Analysen bedeutet dies, dass im idealen Fall, einer vorliegenden direkten Regulierung, alle Datenbanken die mRNA als Target vorhersagen, eine unterschiedliche Expression in der Überlebenszeitanalyse zwischen microRNA und Target vorliegt und mehrere microRNA-Bindungsstellen im 3'-UTR des Targets vorhanden sind. Der hierfür entwickelte Score setzt sich somit aus mehreren Faktoren zusammen, die ein normalisiertes Maß zur Bewertung der ermittelten putativen direkten Bindung (pdB) der microRNA und des Targets bilden. Im weiteren Verlauf dieser Arbeit wird dieser Score als $Score_{pdB}$ bezeichnet. Der $Score_{pdB}$ sei im Folgenden für eine microRNA m und deren putatives

mRNA Target t definiert als

$$Score_{pdB}(m, t) = DB(m, t) + KM(m, t) + MRE(t), \quad (2.19)$$

wobei $DB(m, t)$, $KM(m, t)$ und $MRE(t)$ die normalisierte Werte der einbezogenen Analysen aus der microRNA-Target Vorhersage, den klinischen und bindungsspezifischen Daten darstellen, welche nachfolgend detaillierter betrachtet werden. Die normalisierte Anzahl der Datenbanken, welche eine mRNA als Target t für eine aktuell betrachtete microRNA m vorhersagen, wird durch den Wert $DB(m, t)$ angegeben. Dieser Wert wurde, mittels der maximal genutzten Anzahl an Vorhersagedatenbanken, auf den Bereich $[0, 1]$ normiert, so dass gilt:

$$DB(m, t) = \frac{\text{Anzahl Datenbanken, die } t \text{ als Target von } m \text{ vorhersagen}}{14 \text{ Datenbanken}}. \quad (2.20)$$

Der KM -Wert kennzeichnet die Beziehung der vorliegenden Gruppen der Expression (low/high), die sich aus der Analyse der klinischen Daten von microRNA und Target ergibt und für die sich eine schlechten Prognose der Erkrankung anhand der Kaplan-Meier Methode in der Überlebenszeitanalyse nachweisen lässt. Hierbei werden die Einträge (low oder high) aus der Überlebenszeitanalyse verwendet. Wie im Vorfeld beschrieben, sollte bei einer direkten Regulierung des Targets durch die microRNA eine unterschiedliche Expression zu erkennen sein. Weist eine microRNA beispielsweise eine hohe Expression auf (high), so sollte das putativ direkte Target durch die microRNA vermittelte Regulierung (Abbau, Destabilisierung etc.) eine geringe Expression (low) aufweisen. Gleiches gilt für den umgekehrten Fall, dass eine microRNA eine geringe Expression aufweist, was zu keiner negativen Regulierung der Expression des Targets - durch die microRNA - führen sollte. Der Wert $KM(m, t)$ in Abhängigkeit der Überlebenszeitanalyse (hier mit *surv* von *Survival Analysis* bezeichnet) ist definiert als:

$$KM(m, t) = \begin{cases} 1, & \text{falls } surv(m) \neq surv(t), \text{ mit } surv \in \{low, high\} \\ 0, & \text{sonst} \end{cases}. \quad (2.21)$$

Ein Wert $KM(m, t) = 1$ liegt somit vor, wenn für eine schlechte Prognose, d. h. für eine geringe Überlebenswahrscheinlichkeit in Abhängigkeit der Expression, eine unterschiedliche Expression (low/high oder high/low) zwischen m und t ermittelt wurde. Ein Wert von 0 liegt hingegen vor, wenn keine Unterschiede in der Expression ermittelt werden konnten. Dies könnte womöglich

darauf hindeuten, dass keine direkte Regulation der mRNA durch die microRNA erfolgt, was jedoch ohne weiterführende Experimente nicht nachgewiesen werden kann.

Der Wert $MRE(t)$ stellt die normalisierte Anzahl an bekannten microRNA Bindestellen (MREs) im 3'-UTR des Targets t dar, wie sie anhand von TargetScan ermittelt wurden. Für den hier verwendeten $MRE(t)$ Wert erfolgte eine Normalisierung der MRE-Anzahl in Abhängigkeit der jeweiligen 3'-UTR Länge des Targets. Diese Normalisierung wirkt einem möglichen Längenbias entgegen, da auf kürzeren Sequenzen eine geringere Anzahl an möglichen Bindestellen entfallen könnten, was somit eine Benachteiligung von Targets mit kürzerem 3'-UTR innerhalb der Bewertung zur Folge hätte. Anhand aktueller Literatur ist keine signifikante Korrelation der 3'-UTR Länge mit der Anzahl an MREs nachgewiesen, wenn auch vereinzelte Annahmen und Daten darauf hindeuten, dass längere 3'-UTRs mehr microRNA Bindestellen aufweisen und dadurch ein erhöhtes Potenzial für eine microRNA vermittelte Regulation zugrunde liegen könnte [40, 91]. Um den möglichen Bias der Länge und demnach einer Beeinträchtigung der Bewertung in dieser Arbeit entgegenzuwirken, wurde die Normalisierung anhand der 3'-UTR Länge der Targets angewendet. Aus der aktuell verwendeten Version von TargetScan ergab sich anhand dieser Längennormalisierung ein maximaler Wert von 0,1745, welcher zur zusätzlichen Normierung des $MRE(t)$ auf einen Bereich von $[0, 1]$ genutzt wurde, so dass gilt:

$$MRE(t) = \frac{\text{Anzahl MREs für Target } t}{\text{Länge 3'-UTR von } t \text{ in Nukleotiden} \cdot 0,1745} \quad . \quad (2.22)$$

Vorangegangene Betrachtungen haben gezeigt, dass, entsprechend der bisher beschriebenen Bewertung, die Anzahl der vorhersagenden Datenbanken für eine microRNA-Target Beziehung einen großen Einfluss auf den $Score_{pdB}$ der Gleichung 2.19 hat. Hingegen gehen die Veränderungen in der Anzahl der möglichen MREs im 3'-UTR der Target mRNA nur mit einem sehr geringen Einfluss in den $Score_{pdB}$ ein, da anhand der vorliegenden Daten nur durchschnittlich ca. 0,6 MREs auf eine Länge von 100 Nukleotiden entfallen. Aus diesem Grund wurde eine Wichtung innerhalb der Berechnung des Scores vorgenommen, mit dem Ziel, dass unterschiedliche Expressionen in der Überlebenszeitanalyse und die Zunahme an microRNA Bindestellen es ermöglichen sollen, einen besseren Score beim Vergleich zweier Beziehungen zu erzielen, wenngleich geringfügig weniger Datenbanken eine der betrachteten Beziehungen vorhersagen. Zum anderen soll der $Score_{pdB}$ auf einen einheitlichen Bereich von $[0, 1]$ normiert werden, um den Vergleich der microRNA-Target Beziehungen zwischen den stadienbasierten

Vergleichen der Targetanalyse zu ermöglichen. Die Anzahl der Vorhersagedatenbanken bildet innerhalb dieser Arbeit das wesentliche Maß zur Bewertung, da die dabei genutzten Eigenschaften wichtige Faktoren für eine mögliche Bindung enthalten, jedoch sollen die zusätzlichen Hinweise, durch die klinischen und bindungsspezifischen Daten, die Bewertung mit beeinflussen können. Die Kompensation soll dabei nicht allein auf einem der Werte KM oder MRE beruhen, da beispielsweise die klinischen Daten unvollständig sein können oder deren genaue Erhebung nicht (vollständig) dokumentiert wurde. Des Weiteren stellen die *in silico* vorhergesagte Bindestellen keinen Nachweis für eine reale Regulation dar, noch ist aktuell die Bedeutung der Anzahl an Bindestellen für die Regulation geklärt. Für die Bewertung, der ermittelten microRNA-Target Beziehungen in dieser Arbeit, wurde festgelegt, dass bei einem KM -Wert von 1, also dem Vorhandensein eines Unterschiedes in der Expression von microRNA und Target in der Überlebenszeitanalyse, und mehr als 10 vorhergesagten MREs im 3'-UTR des Targets, die Kompensierung von mindestens einer vorhergesagten Datenbank gegeben sein soll. Beispielhaft angedeutet hieße dies, dass eine ermittelte putativ direkte microRNA-Target Beziehung, in der das Target von 14 Datenbanken vorhergesagt wird, beide Bindungspartner jedoch keine unterschiedliche Expression in der Überlebenszeitanalyse aufweisen und das Target nur eine microRNA Bindestelle besitzt, einen schlechtere Score erhält als die Beziehung, in der das Target von 13 Datenbanken vorhergesagt wird, jedoch eine unterschiedliche Expression in der Überlebenszeitanalyse vorliegt und das Target mehr als zehn MREs aufweist.

Um dies zu realisieren, wurde eine Gewichtung über den KM -Wert eingeführt, welcher bisher den Wert 0 oder 1 annehmen kann. Findet keine Wichtung des KM -Wertes statt, so führt dieser bereits zu starken Schwankungen in der Bewertung und der (oftmals geringe) MRE -Wert wird fast vollständig vernachlässigt. Liegt keine unterschiedliche Expression der Bindungspartner in der Überlebenszeitanalyse vor, kann die Kompensation einer Datenbank, beim Vergleich zweier Beziehungen und der Normalisierung auf die 3'-UTR Länge des Targets, nur über eine starke Zunahme der Anzahl an MREs oder einen kürzeren 3'-UTR erfolgen. Bei unterschiedlicher Expression der Bindungspartner hingegen wird durch die hier vorgenommene Wichtung ermöglicht, dass eine Differenz von mindestens zehn MREs bereits zur Kompensation von wenigstens einer Datenbankvorhersage führen kann. Entsprechend der Gleichung 2.19 weist die Vorhersage von einer Datenbank einen DB -Wert von $1/14 = 0,07143$ auf. Damit der KM und der MRE -Wert gemeinsam mindestens eine Vorhersage durch eine Datenbank kompensieren können, muss die Summe beider normalisierten Werte größer als 0,07143 sein.

Dazu muss zum einen ermittelt werden, welcher *MRE*-Wert bei mindestens zehn vorhandenen Bindestellen theoretisch erreicht werden kann und zum anderen, welche Wichtung sich damit für den *KM*-Wert ergibt, wenn eine unterschiedliche Expression der beiden Bindungspartner vorliegt. Da die 3'-UTR Längen der Transkripte stark variieren können, ist es nur möglich einen ungefähren Wert für das Vorhandensein von zehn Bindestellen unter Einbeziehung der Längennormalisierung zu bestimmen. Für humane Transkripte ist aus der Literatur bekannt, dass deren 3'-UTR eine Länge von mehr als 500 bzw. einer durchschnittlichen Länge von ca. 800 Nukleotiden aufweist [153, 161]. In den hier vorliegenden TargetScan-Daten beträgt die durchschnittliche Länge 898 Nukleotide, für den jeweils längsten 3'-UTR eines bekannten Transkriptes. Auf der Grundlage dieser durchschnittlichen 3'-UTR Längen, ergibt sich für die zehn geforderte MREs eine längennormalisierter Wert von 0,0125 bzw. 0,0111, so dass für den weiteren Verlauf in dieser Arbeit ein minimaler Wert von gerundet 0,01 festgelegt wurde. Mittels des bereits ermittelten maximalen Wertes aus den längennormalisierten TargetScan Daten von 0,1745 (siehe Normierung des *MRE(t)* oben), resultiert daraus somit ein auf den Bereich $[0, 1]$ normierter Wert von 0,0573, bei einer durchschnittlichen 3'-UTR-Länge von ca. 800 Nukleotiden. Damit folglich die Summe aus *KM*- und *MRE*-Wert in der Lage ist, beim Vergleich zweier Beziehungen, mindestens einen Datenbankvorhersage zu kompensieren, muss der *KM*-Wert mit $1/14 - 0,0573 = 0.0141$ gewichtet werden. Liegt somit eine unterschiedliche Expression der Bindungspartner vor und sind mindestens zehn MREs pro 800 Nukleotide 3'-UTR vorhanden, kann diese microRNA-Target Beziehung einen besseren Score erhalten als eine vergleichbare Beziehung, die von einer Datenbank mehr vorhergesagt wird, jedoch keine unterschiedliche Expression der Bindungspartner in der Überlebenszeitanalyse aufweist und/oder über weniger MREs verfügt. Weist die Beziehung eventuell mehr als zehn MREs auf, kann es durch diese Bewertung möglich sein, dass eine größere Anzahl von Vorhersagedatenbanken kompensiert werden kann.

Unter den vorherigen Betrachtungen und der Normierung des Gesamtscores auf einen Bereich von $[0, 1]$, ändert sich die Gleichung 2.19 für das finale Bewertungskriterium der hier entwickelten Targetanalyse wie folgt:

$$Score_{pdB}(m, t) = \frac{DB(m, t) + KM(m, t) \cdot 0,0141 + MRE(t)}{2,0141}. \quad (2.23)$$

Die Gleichung 2.23 beschreibt dabei den in dieser Arbeit angewendeten $Score_{pdB}$ zur Bewertung der identifizierten microRNA-Target Beziehungen unter Einbeziehung der differentiellen Expression, der klinischen als auch der bindungsspezifischen Daten der jeweiligen microRNA und des putativ direkten Targets. Gemäß dieser Gleichung fand für alle ermittelten Beziehungen eine unabhängige Bewertung statt. Die bisherigen Ergebnistabellen aus den Vergleichen der Neuroblastomstadien wurden dabei um zusätzliche Spalten erweitert, welche die längennormalisierten MREs sowie die berechneten Werte des $Score_{pdB}$ enthalten, auf dessen Basis die darin enthaltenen Beziehungen geordnet wurden.

3. Ergebnisse

3.1. Neuronale Zellmorphologie und Proteinverteilung

Neben molekularbiologischen Veränderungen, wie beispielsweise in RNA- oder Proteinabundanzen, sind morphologische Änderungen einzelner Zellen und Veränderungen in der Lokalisation darin enthaltener Proteine ebenso von biologischer Relevanz in menschlichen Erkrankungen. Diese Veränderungen spielen z. B. eine wesentliche Rolle bei der Differenzierung von Zellen oder bei der Fähigkeit zur Invasivität und Migration von Tumorzellen, was signifikante Auswirkungen auf die Prognose und den Verlauf einer Tumorerkrankung nach sich ziehen kann. Im Allgemeinen lassen sich morphologischen Veränderungen relativ einfach mittels visueller Techniken wie Hell- oder Dunkelfeldmikroskopie bewerten. Für detaillierte und quantitativ aussagekräftige Betrachtungen kommen jedoch immer häufiger hochauflösende Hochdurchsatztechniken unter Verwendung oder in Kombination mit der Fluoreszenzmikroskopie zum Einsatz. Proteinverteilungen innerhalb untersuchter Zellen werden oftmals nur visuell betrachtet oder deren Lokalisation in festgelegten Kompartimenten der Zelle mit Hilfe anderer Proteine verglichen, z. B. mittels Koloalisationsanalysen. Quantitative Befunde und aussagekräftige Ergebnisse zur vollständigen Verteilung der Proteine bleiben hierbei ungeklärt. Vor diesem Hintergrund wurde in der vorliegenden Arbeit ein Tool zur automatischen Detektion neuronaler Zellen entwickelt, mit dem Ziel Daten der zellulären bzw. neuronalen Morphologie sowie die Verteilung von markierten Proteinen entlang der Neuritenstrukturen zu extrahieren. Der nachfolgende Abschnitt 3.1.1 befasst sich mit den Ergebnissen der automatischen Detektion zur Berechnung der morphologischen Daten. Dabei werden die Resultate der Detektion anhand der Ergebnisse verfügbarer Anwendungen zur Erkennung neuronaler Zellen validiert. Im Abschnitt 3.1.2 werden die Ergebnisse der ermittelten Proteinverteilungen innerhalb der detektierten Neuritenstrukturen betrachtet. Die Verteilungen einzelner Proteine werden anhand bekannter Lokalisationen bewertet. Zudem wird die

Lokalisation des mRNA bindenden Proteins ZBP1 erfasst und dessen Proteinprofil entlang der Neuriten wird erstmals vollständig dargestellt und beschrieben.

3.1.1. Automatische Detektion der Zellmorphologie

Eine automatische Analyse zur Quantifizierung morphologischer Eigenschaften und darüber hinaus auf einer aussagekräftigen Menge an Bilddaten, schafft wesentliche Vorteile im Hinblick auf das experimentellen Design, den zeitlichen Aufwand und die statistische Signifikanz der Ergebnisse gegenüber einer manuellen Auswertung. Eine automatische Detektion von vollständigen Zellen, wie sie in dieser Arbeit entwickelt wurde, bildet dabei die Grundlage zur Extraktion diverser morphologischer Daten, die - einmal ermittelt - zur Beantwortung unterschiedlicher Fragestellungen und als Grundlagen für weiterführende Hypothesen genutzt werden können. Das Ergebnis der automatischen Detektion des entwickelten *NeuronAnalyzer2D* (vgl. Abschnitt 2.1) besteht aus der vollständigen Morphologie der Neuronen unter Verwendung von aktiven Konturen (vgl. 2.1.2), deren Neuritenregionen, strukturellen Teile sowie morphologischer Daten, wie das Skelett entlang der Neuriten, Verzweigungspunkte, diverse Längen etc. durch Anwendung der waveletbasierter Detektion (vgl. 2.1.3). In Abb. 3.1 ist ein Ausschnitt einer Maximalprojektion eines Mehrkanalfluoreszenzbildes dargestellt (vgl. 2.1.2). Die Abbildung zeigt das segmentierte Neuron mit der detektierten Kontur (schwarz), dem Soma und zwei Neuriten, wovon sich der obere Neurit in zwei Neuritenäste verzweigt. Die Neuritenäste weisen eine Länge von $88,24 \mu\text{m}$ (462 Pixel) und $127,40 \mu\text{m}$ (667 Pixel) auf, der untere Neurit hat eine Länge von $71,05 \mu\text{m}$ (372 Pixel). Die Grenzen der strukturellen Teile sind durch gelbe Linien markiert. Weiterhin sind 10 Verzweigungs- (blaue Punkte) und 12 Endpunkte (rote Punkte) gekennzeichnet, sowie 9 filopodienartige Vorstülpungen (rot) mit einer Länge von maximal $10 \mu\text{m}$ (ca. 52 Pixel). Des Weiteren wird die Fläche der Somaregion ($342,16 \mu\text{m}^2$), die durchschnittliche Breite der Neuritenäste ($0,79 \mu\text{m}$; $1,12 \mu\text{m}$; $0,88 \mu\text{m}$), die Fläche der Wachstumskrone ($33,02 \mu\text{m}^2$; $68,73 \mu\text{m}^2$; $32,10 \mu\text{m}^2$) und die Anzahl der filopodienartige Vorstülpungen innerhalb der Wachstumskronenregion (2; 0; 2) ermittelt.

Abbildung 3.2 veranschaulicht weitere Ergebnisse zur waveletbasierten Detektion der verschiedenen strukturellen Teile des Neurons. Zusätzlich zu den Neuriten in Abb. 3.1 zeigen die hier abgebildeten Neuriten die berechneten Grenzen (vgl. 2.1.3). Alle drei Bilder zeigen Fragmente des Somas auf der linken Seite und die ermittelten Grenzen (gelb), die das Soma vom Neuriten und den Neuritenschaft von der Wachstumskrone abgrenzen. Die Bereiche des

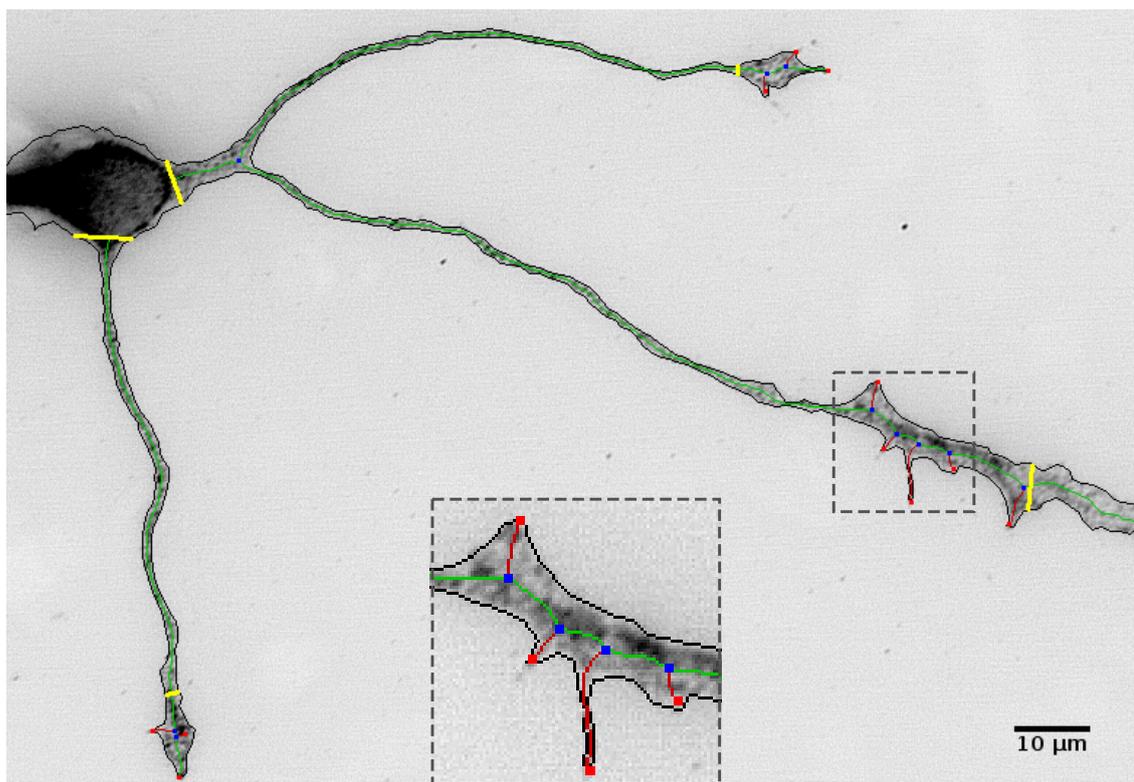


Abbildung 3.1.: Invertiertes Grauwertbild der Maximalprojektion eines Neurons. Schwache Fluoreszenzintensitäten im Originalbild sind durch helle Pixel gekennzeichnet, während hohe Intensitäten durch dunkle Pixel gekennzeichnet sind. Der Ausschnitt zeigt die Kontur des Neurons (schwarz), sowie die Grenzen (gelb) zwischen den strukturellen Neuronenteilen (Soma, Neuritenschaft, Wachstumskrone). Ebenso ist das Neuritenskelett (grün), dessen Verzweigungspunkte (blaue Punkte), Endpunkte (rote Punkte) sowie filopodienartige Vorstülpungen im Neuriten (rot) dargestellt. Der markierte Bereich zeigt einen vergrößerten Neuritenausschnitt mit gekennzeichneten Kontur, dem Skelett und den Verzweigungs- bzw. Endpunkten der Vorstülpungen.

Somas können dabei gut von den Neuriten getrennt werden (vgl. Abb. 3.2 (A-C)), während bei der Abgrenzung des Neuritenschafts von der Wachstumskrone Variationen auftreten können. Maßgeblich für diese Variationen ist die Form der Wachstumskronen, wobei längliche oder deformierte Wachstumskronen zu einer frühzeitigen Abtrennung der Wachstumskrone vom Neuritenschaft führen und die Wachstumskrone somit verlängert wird (vgl. Abb. 3.2 (A, B)). Abbildung. 3.3 stellt exemplarisch eine Ergebnistabelle von ermittelten Morphologiedaten nach der Anwendung des *NeuronAnalyzer2D* dar. Die Tabelle zeigt die Ergebnisse für drei verschiedene Fluoreszenzbilder, auf denen jeweils ein Neuron abgebildet ist. Die Nummerierung der Dateien resultiert aus den Dateinamen (vgl. Abschnitt 2.1.1), welche hier aus Platzgründen verkürzt dargestellt werden. Jedes Neuron im Bild erhält eine *neuronID*, wobei nachfolgend jedes Neuron aus mehreren Neuriten (*neuriteID*) und diese wiederum aus mehreren Neuritenästen

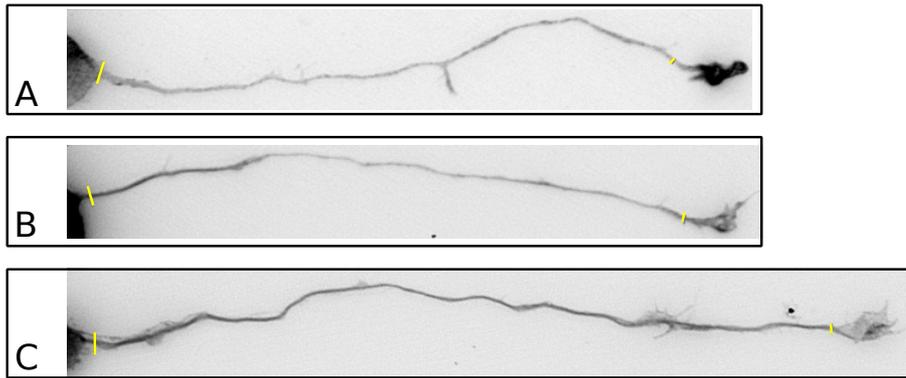


Abbildung 3.2.: Invertierte Maximalprojektion von Neuriten. Die Grenzen (gelb) zwischen den strukturellen Teilen des Neurons wurden aus dem waveletbasierten Ansatz berechnet. Langgestreckte oder deformierte Wachstumskronen der Neuriten führen zu einem vorzeitigen Abschneiden des Neuritenschafts, wodurch die Wachstumskronen verlängert werden (rechte Grenze in A und B). Alle gezeigten Ausschnitte wurden an der Somaregion abgeschnitten.

(*nBranches*) bestehen können. Eine detaillierte Beschreibung der einzelnen morphologischen Daten kann aus Tab. A.1 im Anhang entnommen werden. Die resultierenden Daten der automatischen Detektion sind darüber hinaus als SWC-Datei verfügbar. Das SWC Format [30] erlaubt in einem kompakten Textformat die Repräsentation des Neuritenskeletts mithilfe einer Baumstruktur. Jede Zeile steht dabei für die einzelnen Punkte eines Graphen des Neurites mit sieben Datenpunkten, die Auskunft über den Strukturtypen geben, d.h. seine Lage im Soma, Axon etc., sowie die Koordinaten des Punktes. Das Format kann sowohl für 2D- als auch 3D-Daten genutzt werden. Eine vollständige Spezifikation ist online verfügbar⁴⁰. Die Verwendung dieses Formates ermöglicht die Nutzung der extrahierten Morphologiedaten in anderen verfügbaren Anwendungen, wie beispielsweise Vaa3D⁴¹ [192] oder dem Simple Neurite Tracer⁴² [140]. Die Daten können mittels SWC-Datei importiert werden, um anwendungsabhängig graphisch dargestellt und/oder weiterführend analysiert zu werden.

Validierung der automatischen Detektion

Um die Genauigkeit der beschriebenen Detektionsmethode des *NeuronAnalyzer2D* zu bestimmen, wurden die abgebildeten Neuronen (vgl. Abschnitt 2.1.1) auf 20 zufällig ausgewählt Bilder von einem Lebenswissenschaftler manuell analysiert, was die sogenannte *Ground Truth* der Daten bildet. Die Segmentierung dieser Daten umfasst die vollständigen Konturen der

⁴⁰<http://www.neuronland.org/NLMorphologyConverter/MorphologyFormats/SWC/Spec.html>

⁴¹<http://www.vaa3d.org/>

⁴²http://fiji.sc/Simple_Neurite_Tracer

file	neuronID	neuronArea	neuriteID	nArea	nBranches	nSpines	nEnds	branchID	bLength	bWidth	bsLength	bsWidth	bsArea	bcLength	bcWidth	bcArea	bcSpines
/,60/0	11341	0	649	0	0	1	0	88	3.49	31	5.14	331	57	2.57	295	0	
/,60/0	11341	1	5687	1	1	2	0	1073	2.49	1035	2.4	5200	38	4.91	467	1	
/,57/0	8948	0	5011	2	2	3	0	632	3.59	586	3.34	4144	46	6.62	841	2	
/,57/0	8948	1	1612	1	1	2	0	141	5.15	106	5.09	1159	35	5.26	435	1	
/,53/0	15973	0	3965	2	1	3	0	279	4.44	159	5.64	2012	120	2.83	700	0	
/,53/0	15973	0	3965	2	1	3	1	146	5.34	36	6.75	2012	110	4.66	1202	1	
/,53/0	15973	1	6438	4	2	5	0	529	3.62	493	3.66	4649	36	3.12	252	0	
/,53/0	15973	1	6438	4	2	5	1	663	3.23	438	3.88	4649	225	1.97	830	0	
/,53/0	15973	1	6438	4	2	5	2	381	4.08	330	3.9	4649	51	5.21	649	0	
/,53/0	15973	2	1964	0	0	1	0	176	4.94	139	4.42	1268	37	6.81	668	0	

Abbildung 3.3.: Ausschnitt der Ergebnistabelle wie sie nach Anwendung des NeuronAnalyzer2D generiert wird, um die ermittelten Morphologiedaten der Detektion aufzulisten. Es werden Angaben über Längen, Flächen und Breiten der im Bild detektierten Neuronen, deren Neuriten sowie der Neuritenschäfte und Wachstumskronen aufgeführt.

Neuronen mit ihren Neuriten, auf deren Basis die automatisch segmentierten Neuronenregionen mit den manuell erzeugten Regionen verglichen wurden. Als Maß für die Vollständigkeit und die Genauigkeit der entwickelten Detektionsmethode wurden der *Recall* und die *Precision* auf den Pixeln der zu vergleichenden Regionen berechnet. Diese pixelbasierten Maße berechnen sich dabei wie folgt:

$$Recall = \frac{TP}{TP + FN}, \quad Precision = \frac{TP}{TP + FP},$$

wobei *TP* die Anzahl an richtig positiven (*true positives*), *FN* die falsch negativen (*false negatives*) und *FP* die Anzahl der falsch positiven (*false positives*) Pixel angibt. Pixel die als *FN* bezeichnet werden, wurden fälschlicherweise als Hintergrundpixel detektiert und nicht dem Neuron zugeordnet, während *FP* fälschlich als Neuronenpixel detektiert wurden, diese jedoch zum Hintergrund gehören. Der Recall-Wert gibt somit den Anteil der Ground Truth Pixel an, die korrekt von der Methode als Neuronenpixel detektiert wurden (Vollständigkeit der Detektion). Der Anteil der detektierten Pixel, die innerhalb der Ground Truth Regionen liegen, wird durch den Precision-Wert angegeben (Genauigkeit der Detektion). In Abb. 3.4 sind die Ergebnisse der ermittelten Maße der evaluierten Neuronen als Boxplots dargestellt. Der Median des Recall liegt bei einem Wert von 0,76 und weist ein Minimum von 0,45 sowie ein Maximum von 0,91 auf. Für die Precision wird ein Median von 0,85 bei einem Minimum von 0,74 und einem Maximum von 0,94 erreicht. Die geringen Interquartilabstände der einzelnen Boxplots zeigen eine kleine Varianzen zwischen den Recall- und Precision-Werten. Im Recall-Plot sind zwei Ausreißer erkennbar, welche durch Neuriten entstehen die eine sehr dünn Struktur sowie eine schwache Färbung aufweisen. Dies führt zu einer nicht korrekt Segmentierung dieser Stellen im Neuriten. Für Details zu dieser Problematik sei hier auf die Evaluierung der Segmentierungsergebnisse

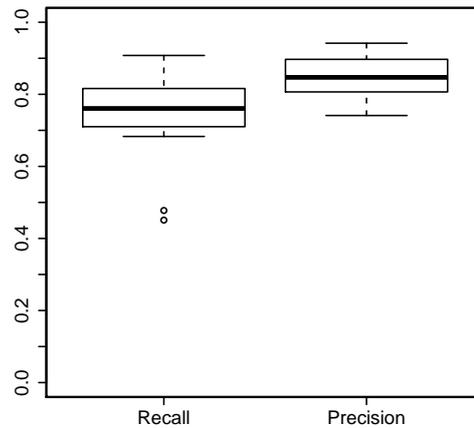


Abbildung 3.4.: Boxplots von Recall und Precision, berechnet anhand der pixelbasierten Vergleiche zwischen den resultierenden Neuronenregionen des NeuronAnalyzer2D und der Ground Truth. Die Grundlage bilden zwanzig zufällig ausgewählten Neuronen. Eine geringe Varianz der Recall- und Precision-Werte wird durch die kleinen Interquartilabstände der Boxplots angezeigt, was die Gesamtqualität der Detektion des Algorithmus verdeutlicht. Zwei Ausreißer (kleine Kreise) im Recall entstehen durch schwach gefärbte dünne Neuriten.

im nachfolgenden Text verwiesen. Ein exemplarisches Segmentierungsergebnis ist in Abb. 3.5 dargestellt und zeigt die Region der Ground Truth sowie des automatisch detektierten Neurons. Anhand der Darstellung ist erkennbar, dass kleine filopodienartige Vorstülpungen (rote Pfeile) nicht vollständig und der untere Neurit (grüner Pfeil) zwar in seiner Länge, aber nicht in der Breite vollständig korrekt segmentiert wurden. Ursache dafür ist der erwähnte geringe Kontrast durch eine schwache Färbung des Neuriten. Des Weiteren weist die detektierte Somaregion eine Vergrößerung (blauer Pfeil) im Vergleich zur Ground Truth auf, was durch eine perinukleäre Akkumulation des Proteins bedingt ist und folglich eine Überstrahlung der Fluoreszenzintensität im Bereich des Somas induziert. Bei dem Vergleichen der Neuronenregionen mit den manuell segmentierten Regionen der Ground Truth weist die automatische Detektion insgesamt jedoch eine hohe Qualität auf, was sich anhand der ermittelten Maße belegen lässt. Für eine Evaluierung der Segmentierungsergebnisse des NeuronAnalyzer2D wurden diese mit den Ergebnissen des NeuriteTracer [196] und Vaa3D-Neuron [192] verglichen. Beide Anwendungen stellen automatische Tools zur Verfolgung (*Tracing*) von Neuriten dar und generieren wie der NeuronAnalyzer2D vollständige Regionen der detektierten Neuronen. Die resultierenden Neuronenregionen aller drei Tools wurden hierzu, ergänzend zu Recall und Precision, mit der Ground Truth unter Anwendung des sogenannten *F-Scores* [256] verglichen. Dieser Score ist

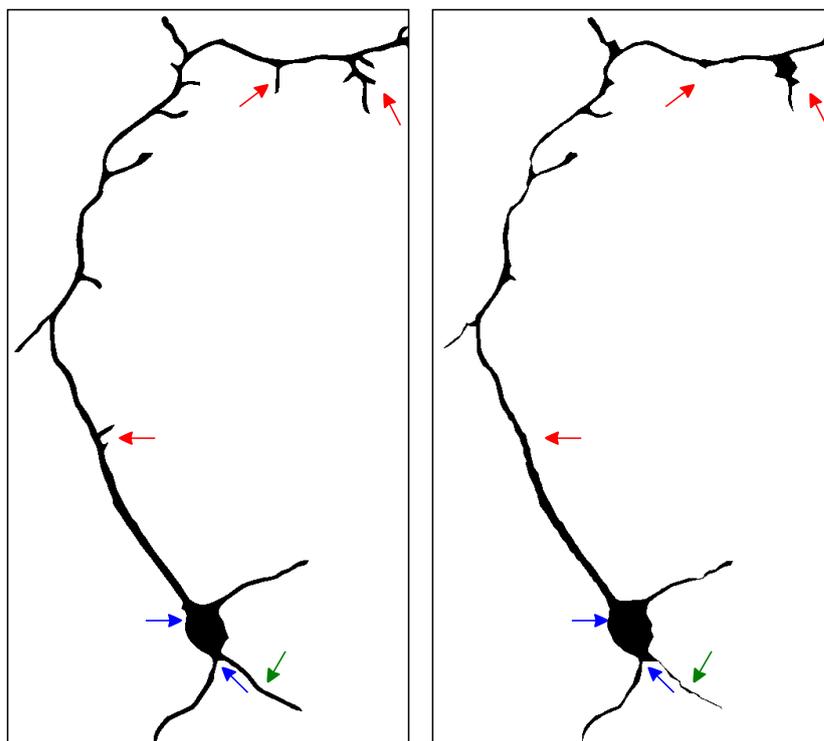


Abbildung 3.5.: Beispiel einer *NeuronAnalyzer2D* basierten Segmentierung eines Neurons. Die Abbildung zeigt den Vergleich zwischen der Ground Truth einer manuellen Segmentierung (links) und der automatischen Segmentierung anhand der beschriebenen Methode des *NeuronAnalyzer2D* (rechts). Durch den schwachen Kontrast und die geringe Breite mancher Regionen, werden der untere rechte Neurit (grüner Pfeil) und einige filopodenähnliche Vorstülpungen (rote Pfeile) nicht vollständig segmentiert. Die Somaregion weist eine Vergrößerung (blaue Pfeile), bedingt durch die perinukleäre Akkumulation des Proteins, auf.

ein Maß für die Detektionsleistung der Methode und dient der Beurteilung deren Güte durch die Kombination von Recall und Precision zu einem einzelnen Gütekriterium. Der F-Score ist definiert als das harmonisches Mittel aus Recall und Precision:

$$F_{Score} = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision}.$$

Ein Score von 1 entspricht dabei einer Detektion von 100 %, d. h. es werden genau die Pixel anhand der Methode detektiert, die mittels Ground Truth die Neuronenregion markieren. Der NeuriteTracer ist als ImageJ Plugin verfügbar und wurde für die automatische Detektion von nicht überlappenden Neuronen in zweidimensionalen Fluoreszenzmikroskopbildern entwickelt. Das Tool benötigt benutzerdefinierte Schwellwerte zur Erkennung der Zellkerne und der markierten Neuronen sowie einen vorgegebenen Bereich für eine zulässige Zellkerngröße.

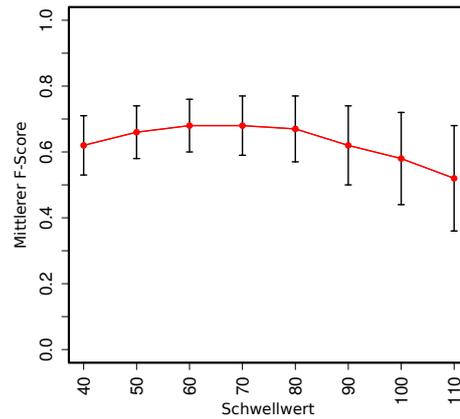


Abbildung 3.6.: Zur Ermittlung eines geeigneten neuronalen Schwellwertes für das Tool NeuriteTracer wurden die segmentierten Neuronenregionen mit der Ground Truth verglichen und die mittleren F-Scores mit den dazugehörigen Standardabweichungen für verschiedene Schwellwerte berechnet.

Alle Schwellwerte und Parameter sollten, nach [196], für die Analyse von mehreren Bildern angemessen eingestellt werden. Als Schwellwert für die Kerne wurde für den hier durchgeführten Vergleich ein mittlerer Intensitätswert von 130 und für die mittlere Größe ein Bereich von [100, 4000] Pixel festgelegt. Die Bestimmung dieser Werte erfolgte durch eine manuelle Binarisierung und Vermessung der Zellkerne auf den zwanzig gewählten Bildern. Auf den vorliegenden fluoreszenzbasierten Bildern ist es, im Gegensatz zu den Zellkernen, nicht trivial einen angemessenen Schwellwert für die Regionen aller markierten Neuronen zu bestimmen. Varianzen im Kontrast der Bilder als auch strukturbedingte Veränderungen der Fluoreszenzintensitäten in den Neuronen sind dabei relevant. Um jedoch den Vergleich mit dem NeuriteTracer zu ermöglichen, wurden verschiedene Schwellwerte im Bereich [40, 110], in Zehnerschritten, für die Detektion der Neuronen in den Testbildern verwendet. Darauf basierend wurde ein Wert von 60 als neuronaler Schwellwert für den NeuriteTracer angenommen. Dieser Wert erreicht zusammen mit dem Schwellwert 70 den höchsten mittleren F-Score (0,68), weist jedoch zum Wert 70 (Standardabweichung 0,09) eine geringere Standardabweichung (0,08) für die getesteten Bilder auf (vgl. Abb. 3.6). Das zweite Tool, Vaa3D-Neuron, ist ein Modul für die automatische Detektion und Rekonstruktion von Neuronen, welches in dem Vaa3D Software-Paket für die Visualisierung und Analyse von biologischen Bildern enthalten ist. Dieses Tool wurde für die Auswertung von 3D-Bildern entwickelt, ist jedoch auch auf 2D-Daten anwendbar. Auf Anfrage wurde eine Version⁴³ zur Verfügung gestellt, die ein automatisches Tracing und die Rekonstruktion von Neuriten ermöglicht [193]. Für die verbleibenden einzustellenden

⁴³Vaa3D-Neuron2 v2.608 (Vaa3D v2.813), Linux Redhat and Fedora 64-bit

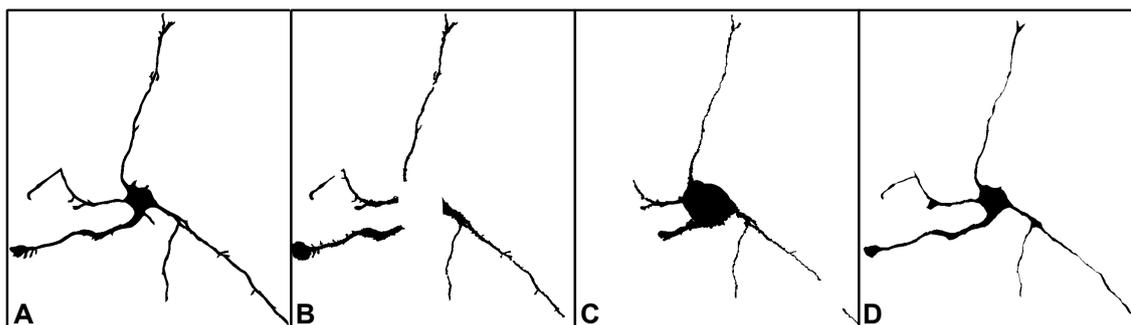


Abbildung 3.7.: Binäre Regionen eines detektierten Neurons. (A) Manuell gekennzeichnete Ground Truth. (B-D) Detektionsergebnisse des NeuriteTracer, Vaa3D-Neuron und des hier entwickelten NeuronAnalyzer2D.

Parameter wurden die empfohlenen Standardwerte verwendet, sowie die verfügbare automatische Schwellwertbestimmung genutzt.

Die Abbildungen 3.7 und 3.8 zeigen die Ergebnisse aus dem Vergleich des NeuronAnalyzer2D mit den beiden genannten Tools. Abbildung 3.7 stellt die binären Regionen eines detektierten Neurons der Ground Truth und der Tools gegenüber. Die F-Scores der Detektionsergebnisse aller zwanzig evaluierten Neuronen sind ergänzend hierzu in Abb. 3.8 (A) für den NeuriteTracer (rot), Vaa3D-Neuron (blau) und den in dieser Arbeit entwickelten NeuronAnalyzer2D (schwarz) dargestellt. Auf der Abszissenachse ist die Nummerierung der einzelnen Neuronen, wie sie aus den Dateinamen resultieren, angegeben (vgl. Abschnitt 2.1.1). Abbildung 3.8 (B) fasst die zur Bewertung der Detektionsergebnisse berechneten Maße, Recall, Precision und den F-Score, in Boxplots zusammen. Die Ergebnisse aus den Vergleichen der ermittelten Maße zeigen, dass der NeuronAnalyzer2D bessere Resultate bei der Detektion der Neuronen erzielt und die beiden anderen Tools im Mittel übertrifft. Der NeuriteTracer erreicht für die drei betrachteten Maße Medianwerte von 0,69; 0,65 und 0,68. Das Tool Vaa3D-Neuron erzielt Mediane von 0,63; 0,55 und 0,56. Der in dieser Arbeit entwickelte NeuronAnalyzer2D erreicht für Recall, Precision und F-Score Mediane von 0,76; 0,85 und 0,81. Die Interquartilabstände der Boxplots in Abb. 3.8 (B) sind beim NeuriteTracer sowie NeuronAnalyzer2D gering und kennzeichnen eine kleine Varianz in den ermittelten Scores, während Vaa3D-Neuron eine höhere Varianz in den Scores aufweist. Diese erhöhten Varianzen deuten auf Schwankungen der Detektionsergebnissen in den verschiedenen Bildern hin.

Ausnahmen in den Resultaten des NeuronAnalyzer2D bilden die Neuronen 14 (vgl. Abb. 3.9 (G)) und 39 (vgl. Abb. 3.9 (I)), welche als die Neuronen mit dem geringsten F-Score in Abb. 3.8 (A)

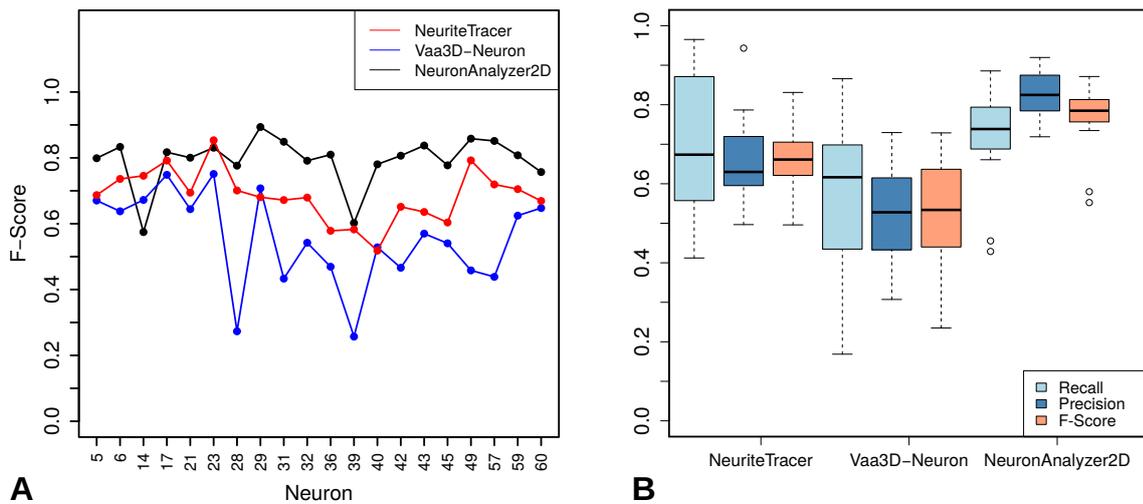


Abbildung 3.8.: Validierung des NeuronAnalyzer2D im Vergleich mit den Ergebnissen zweier öffentlich verfügbarer Tools. (A) Berechnete F-Scores aus dem Vergleich der Neuronenregionen der Ground Truth mit den Ergebnissen des NeuriteTracer (rot), Vaa3D-Neuron (blau) und dem NeuronAnalyzer2D (schwarz). Der F-Score veranschaulicht die hohe Güte der Detektion durch Anwendung des entwickelten NeuronAnalyzer2D gegenüber den andere Tools. Ausnahmen bilden die Neuronen 14 und 39, welche einen dünnen Neuriten mit sehr geringer Fluoreszenz aufweisen und die bereits als Ausreißer im Recall-Boxplot (vgl. Abb. 3.4) erkennbar waren. (B) Zusammenfassung der berechneten Validierungsmaße (Recall, Precision und F-Score), welche sich aus den pixelbasierten Vergleichen der ermittelten Neuronenregionen jedes der drei Tools mit der Ground Truth ergeben. Die beiden in (A) genannten Neuronen 14 und 39 sind neben dem Recall auch als Ausreißer im F-Score erkennbar.

und als Ausreißer in den Boxplots (vgl. Abb. 3.4 und 3.8 (B)) erkennbar sind. Eine genauere Betrachtung verschiedener Detektionsergebnisse in Abb. 3.9 macht deutlich, dass die fehlerbehaftete Detektionen im NeuronAnalyzer2D an sehr langen und dünnen Neuriten mit geringer Fluoreszenzintensität auftritt (vgl. Abb. 3.9 (G, I)). Derartige Strukturen enthalten weniger markierte Proteine, die zusätzlich über eine lange Distanz entlang der Neuriten verteilt sind. Als Konsequenz weisen diese Strukturen eine schwache bis mäßige Fluoreszenz auf sowie meist inhomogen markierte Bereiche. Abbildung 3.9 (H) zeigt jedoch, dass der NeuronAnalyzer2D in der Lage ist sowohl lange als auch dünne Neuriten mit geringer Fluoreszenzintensität zu detektieren.

Verglichen mit dem NeuronAnalyzer2D, detektiert der NeuriteTracer das Neuron 14 besser (vgl. Abb. 3.8 (A) und 3.9 (A)), wobei die Neuronen 17, 23 und 39 bei Betrachtung des F-Scores gleich gut detektiert werden. Für das Neuron 39 zeigt sich anhand der Abb. 3.9 (C) jedoch, dass die Ergebnisse trotz ähnlicher F-Scores von unterschiedlicher Qualität sind. Die Kontur der detektierten Region des NeuriteTracer erfasst den Neuriten besser, als das Ergebnis des NeuronAnalyzer2D. Der Neurit wird jedoch vom Zellkörper getrennt und seine Region wird

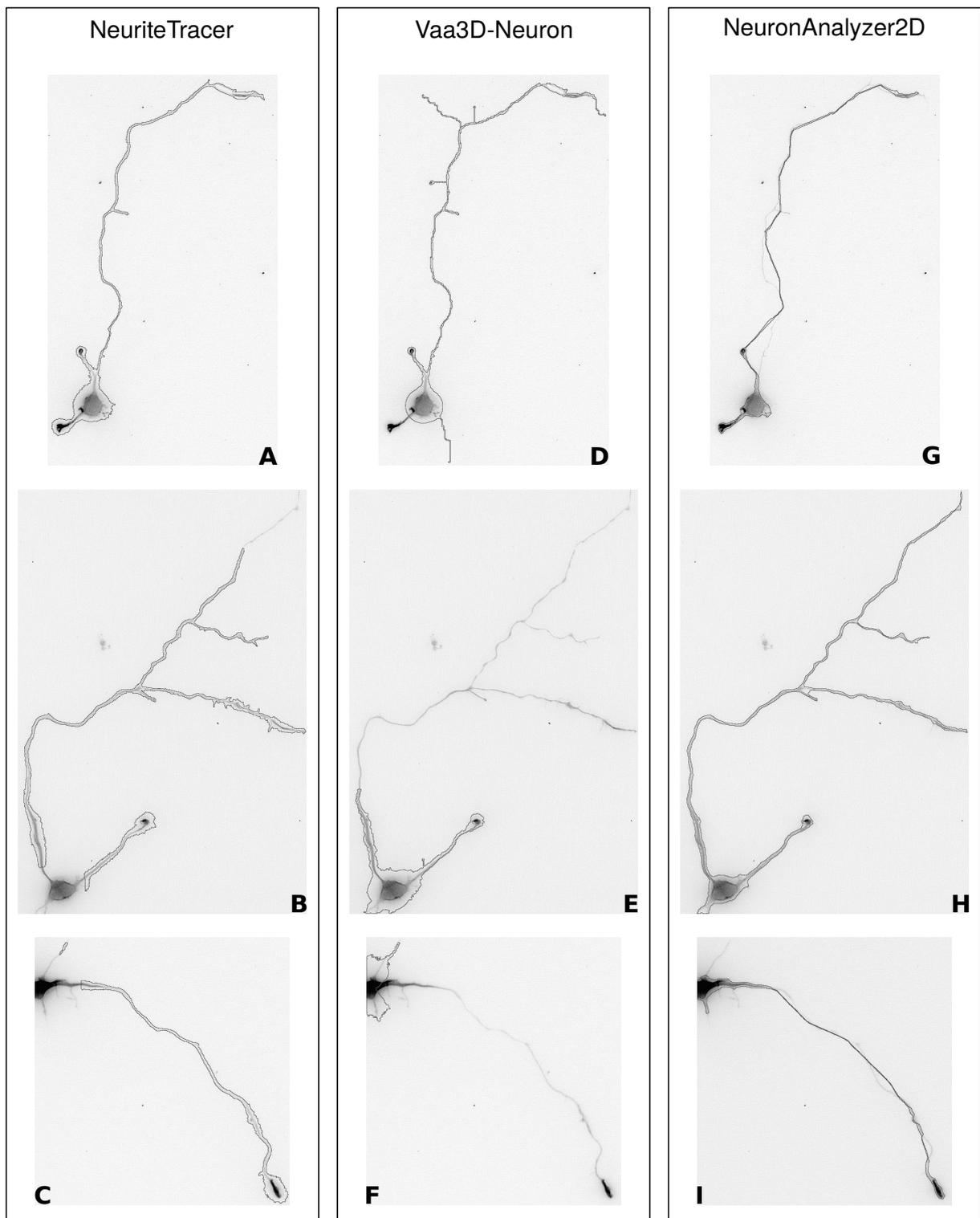


Abbildung 3.9.: Invertierte Maximalprojektionen der Neuronen 14 (A, D, G), 31 (B, E, H) und 39 (C, E, I). Gezeigt sind die Konturen der detektierten Neuronenregionen (schwarz) der drei verwendeten Tools. Die Bilder (G) und (I) zeigen die vom NeuronAnalyzer2D fehlerhaft detektierten Neuronen 14 und 39. Die Fehler treten entlang der dünnen Neuriten auf, bedingt durch eine schwache Fluoreszenzfärbung. (H) zeigt die vollständig vom NeuronAnalyzer2D detektierte Fläche des Neurons 31. Weitere Details siehe Text.

breiter detektiert als die Färbung der Maximalprojektion zeigt. Die kleineren Neuriten am Zellkörper werden, wie auch beim *NeuronAnalyzer2D*, nicht korrekt erkannt. Für alle weiteren Neuronen scheint der eingestellte neuronale Schwellwert, trotz des im Vorfeld durchgeführten Tests auf den Bildern (vgl. Abb. 3.6), nicht für die Detektion aller abgebildeten Neuronen geeignet zu sein. In einigen Fällen gehen Teile des Somas verloren, Neuriten werden fragmentiert und Teilstücke werden im Bereich des Somas und der Neuriten vergrößert (vgl. Abb. 3.7 (B) und 3.9 (A - C)).

Eine bessere Detektion des Neuron 14, im Vergleich zum *NeuronAnalyzer2D*, wird ebenso durch *Vaa3D-Neuron* erzielt (vgl. Abb. 3.8 (A) und 3.9 (D)). Verglichen zur Ground Truth sind die detektierten Regionen der weiteren Neuronen im Bereich des Somas jedoch häufig größer, Neuriten werden fragmentiert oder Teile gehen verloren (vgl. Abb. 3.7 (C)), so dass *Vaa3D-Neuron* meist unterhalb der Detektionsergebnisse der beiden anderen Tools liegt (vgl. dazu Abb. 3.8). Dies zeigt sich auch bei Betrachtung der Ergebnisse in Abb. 3.9 (D - F). Bereiche des Somas und Neuriten in dessen direkter Nähe werden oftmals falsch detektiert. Während in Abb. 3.9 (D), trotz vergrößertem Somabereich, der längste Neurit im Vergleich mit dem *NeuriteTracer* und dem *NeuronAnalyzer2D* wesentlich besser detektiert wird, fallen die Ergebnisse in (E) und (H) schlechter aus. Die langen Neuriten werden gar nicht oder nur ansatzweise detektiert. Ein Grund dafür könnte durch die Charakteristik der Bilder gegeben sein, die zum einen in 2D vorliegen und zum anderen häufig eine Überstrahlung der Somaregion durch die bereits erwähnte perinukleäre Akkumulation der Proteine aufweisen.

Das Gesamtergebnis der hier ermittelten Maße und der abgebildeten Beispielsegmente zeigt die bessere Detektionsleistung des *NeuronAnalyzer2D* im Vergleich mit anderen Anwendungen, die zur Detektion von Neuronen auf Basis der hier vorliegenden Bilder zur Verfügung stehen. Folglich eignet sich die in dieser Arbeit vorgestellte Methode unter Verwendung der hier entwickelten externen Energie der aktiven Konturen und waveletbasierten Erkennung neuronaler Strukturen gut zur Detektion von neuronalen Zellen aus Fluoreszenzmikroskopbildern und extrahiert eine Vielzahl morphologischer Daten für weiterführende Analysen. Neben biologisch relevanten Parametern, wie beispielsweise der Zellkerngröße (vgl. 2.1.2)), der Größe der Somaregion und der maximalen Länge filopodienartiger Vorstülpungen (vgl. 2.1.3), müssen zudem nur wenige Parameter zur Detektion der Neuronen in den einzelnen Bildern manuell und damit subjektiv festgelegt werden.

3.1.2. Proteinverteilung in neuronalen Zellen

Neben den ermittelbaren morphologischen Daten dient die Extraktion von Proteinprofilen (vgl. 2.1.4) der quantitativen Darstellung der Proteinverteilung innerhalb der Neuronen. Basierend auf den identifizierten strukturellen Teilen des Neurons, unter Anwendung der automatischen Detektion (vgl. Abschnitt 2.1), wurden in dieser Arbeit die Proteinprofile der drei markierten Proteine F-Aktin, α -Tubulin und ZBP1 entlang der Neuriten extrahiert. Die Proteine F-Aktin und α -Tubulin dienen dabei zum einen als zellulärer Marker, um ein möglichst vollständiges Neuron, mit Neuriten, Spines sowie Wachstumskronen, zu färben und zum anderen als Kontrollen für die hier extrahierten Proteinprofile, da deren Lokalisation und Verteilung bereits in der Literatur beschrieben ist. Die Verteilung von ZBP1 ist hingegen innerhalb von Neuriten nicht vollständig beschrieben und dessen Lokalisation ist nur für vereinzelte Regionen wie z. B. Wachstumskronen oder Spines untersucht, ohne jedoch quantitative Befunde zur vollständigen Verteilung zu liefern. In der hier vorliegenden Arbeit wird der Verlauf von ZBP1 innerhalb der Neuriten vom Soma bis zur Wachstumskrone erstmals vollständig dargestellt und beschrieben.

Das Ergebnis der Profilextraktion besteht aus der mittleren Intensitätsverteilung der analysierten Proteine und der daraus resultierenden polynomialen Interpolation, deren Verlauf die Verteilung und damit das sogenannte Proteinprofil entlang der Neuriten beschreibt (vgl. 2.1.4). In Abbildung 3.10 (A) sind dazu die Ergebnisse der Extraktion für die drei genannten Proteine dargestellt. Anhand der Proteinprofile ist erkennbar, dass die Konzentration von F-Aktin entlang der Neuriten vom Soma zur Wachstumskrone insgesamt gesehen abnimmt. Etwa in der Mitte des Neuritenschafts beginnt eine leichte Zunahme bis hin zum Ende der Wachstumskrone, wobei dort letztlich eine verringerte Konzentration im Vergleich zur Somaregion vorliegt. Im Gegensatz dazu nimmt die α -Tubulin Konzentration im Verlauf vom Soma zur Wachstumskrone stark ab und es liegt innerhalb der Wachstumskrone eine verringerte Konzentration an α -Tubulin vor. Aus unterschiedlichen Publikationen, beispielsweise aus [53, 141, 232, 259], kann entnommen werden, dass die Beweglichkeit und das Auswachsen der Wachstumskronen stark von der Aktindynamik in der Zelle abhängig ist. Speziell das F-Aktin spielt dabei eine maßgebliche Rolle, wodurch es hauptsächlich in der Peripherie der Wachstumskrone lokalisiert ist und dort eine erhöhte Konzentration zu anderen Bereichen aufweist. Die Konzentration innerhalb der Neuriten ist dabei verhältnismäßig

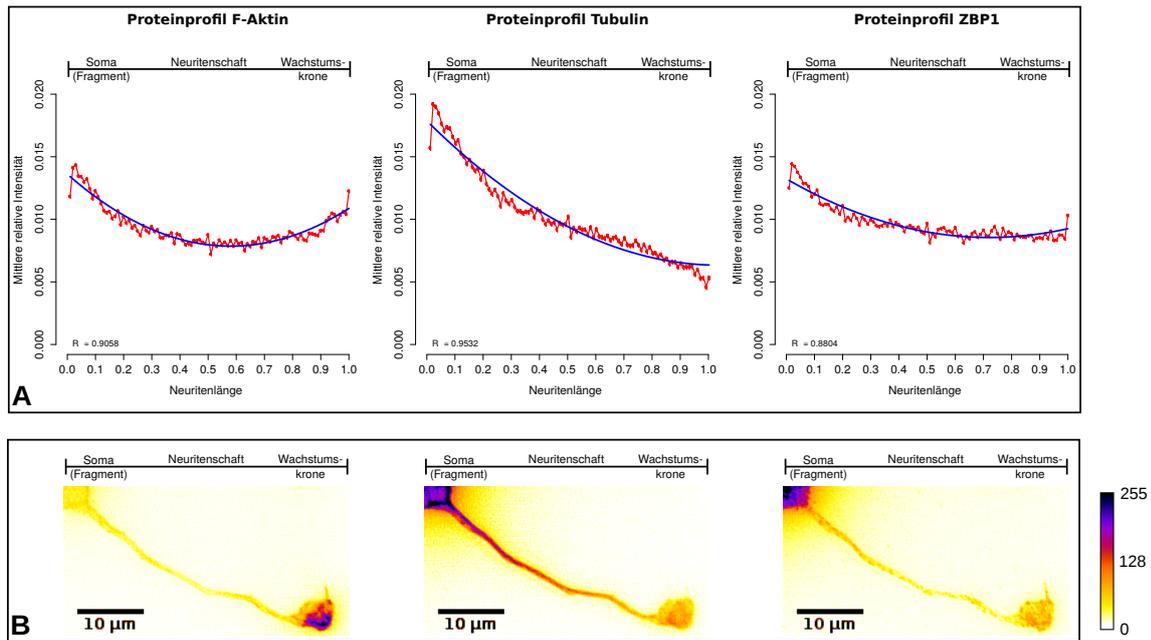


Abbildung 3.10.: (A) Extrahierte Proteinprofile des F-Aktin, α -Tubulin und ZBP1 vom Soma entlang des Neuriten zur Wachstumskrone unter Verwendung des NeuronAnalyzer2D. Die roten Linien stellen das mittlere Intensitätsprofil des dazugehörigen Proteins dar. Die blauen Linien kennzeichnen die polynomiale Interpolation des mittlere Intensitätsprofils. (B) Falschfarbenbilder eines Beispielneuriten unter Verwendung des fire-color LUT. Die Bilder zeigen die Verteilung der Proteine entlang eines einzelnen Neuriten, entsprechend der Proteine der gezeigten Profile in (A). Auf 0 skalierte Pixel kennzeichnen geringe Fluoreszenzintensitäten, während Pixel die auf 255 skaliert sind hohe Fluoreszenzintensitäten und damit eine hohe Proteinkonzentration kennzeichnen.

gering. Zur Strukturierung und Stabilisierung der Neuriten hingegen werden Tubuline für den Mikrotubuliumaufbau vom Zellkörper in die Neuriten transportiert. Dies geht mit einer erhöhten Konzentration an α -Tubulin innerhalb des Neuritenschafts einher. Im Vergleich zur Peripherie der Wachstumskrone liegt somit eine erhöhte Konzentration an Protein entlang der Neuriten vor. Sowohl die Lokalisationen der beiden Proteine als auch deren Verteilungen, wie sie in der Literatur beschrieben sind, lassen sich anhand der in dieser Arbeit extrahierten Proteinprofile von F-Aktin und α -Tubulin bestätigen.

Proteinprofil des Zipcode Binding Protein 1 (ZBP1)

Für das mRNA bindende Protein ZBP1 (*Zipcode Binding Protein 1*) wurde aktuell noch kein Verlauf der Konzentration innerhalb der Neuriten primärer hippocampaler Mauseurone beschrieben oder dargestellt. Es ist bekannt, dass ZBP1 in sich entwickelnden Neuronen an die β -Aktin mRNA bindet und zu Bereichen mit hoher Aktindynamik transportiert werden kann,

wie beispielsweise den Wachstumskronen, wo es durch die lokale Translation der β -Aktin mRNA zu einer ZBP1 vermittelten Aktivierung der β -Aktin mRNA kommen kann (vgl. 1.1.2). ZBP1 steht dabei in räumlicher Nähe zu β -Aktin, was sowohl im Zellkern, in Wachstumskronen als auch in Dendriten und Spines gezeigt werden konnte [57, 95, 248, 283].

Mithilfe der Extraktion von Proteinverteilungen des **NeuronAnalyzer2D** lässt sich aus den vorliegenden Fluoreszenzbildern ermitteln, dass ZBP1 entlang der Neuriten fast gleichmäßig lokalisiert ist. Im Bereich des Somas liegt die höchste Konzentration des Proteins vor, die dann in Richtung Wachstumskrone entlang des Neuritenschafts geringfügig abnimmt. Die hier berechneten Daten zeigen eine leichte Zunahme des geglätteten Profils (blaue Kurve) im peripheren Bereich des Neuritenschaftes in Richtung Wachstumskrone (vgl. Abb. 3.10 (A)). In diesem Bereich tritt eine erhöhte Aktindynamik auf, was sich anhand des zuvor beschriebenen F-Aktin Profils zeigt. Insgesamt zeigt ZBP1 somit eine geringere Konzentration am endständigen Bereich der Wachstumskronen als im Bereich des Neuritenursprung im Soma auf (vgl. Abb. 3.10 (A)). Durch die hier gewonnenen Ergebnisse bestätigt sich das bekannte Bild über die Verteilung von ZBP1 in sich entwickelnden Neuronen auch innerhalb primärer hippocampaler Mausneuronen. Das Protein zeigt dabei einen ähnlichen Verlauf der Proteinverteilung wie das hier gezeigt F-Aktin und scheint an Orten verstärkter Aktindynamik in erhöhter Konzentration vorzukommen, um die ZBP1 vermittelte Aktivierung von β -Aktin an dynamischen Bereichen des Zellwachstums zu gewährleisten. In Abbildung 3.10 (B) ist hierzu, im Vergleich mit den extrahierten Proteinprofile in Abb. 3.10 (A), die Intensitätsverteilungen der markierten Proteine entlang eines Beispielneuriten dargestellt. Der Vergleich zeigt, dass die extrahierten Proteinprofile den dargestellten Intensitätsverteilungen der Proteine in den dazugehörigen Fluoreszenzaufnahmen entsprechen. Aus den ermittelten Profilen sowie der Übereinstimmung mit den beschriebenen Lokalisationen und Verteilungen der beiden Proteine F-Aktin und α -Tubulin wird deutlich, dass sich der in dieser Arbeit entwickelte **NeuronAnalyzer2D** zur Extraktion von Proteinprofilen anhand der Intensitätsverteilungen in Neuronen eignet und zu einer quantitativen Bemessung der Proteinverteilung herangezogen werden kann. Die Methode bietet über die Beschreibung der Proteinverteilung hinaus die Möglichkeit, diese auch graphisch darzustellen und die Verteilungen verschiedener Proteine untereinander zu vergleichen sowie unbekannte Proteinverteilungen bzw. Lokalisationen zu definieren.

3.2. Transkriptomanalyse des Neuroblastoms

Die in dieser Arbeit vorliegenden Tumorproben der Patienten (vgl. 2.2.1) wurden, wie einleitend in Abschnitt 1.2 beschrieben, nach dem *International Neuroblastoma Staging System* (INSS) entsprechend ihrer Charakteristiken in die Stadien des Neuroblastoms eingeteilt. Diese Einteilung beruht im Wesentlichen auf pathologischen Befunden ohne Einbeziehung molekularbiologischer Eigenschaften. Ziel der Analysen in dieser Arbeit ist es, die Proben hinsichtlich ihrer molekularbiologischen Veränderungen zwischen den Stadien zu untersuchen. Die Erfassung des Transkriptoms der einzelnen Proben mittels RNA- und microRNA-Sequenzierung (vgl. 3.2.1) dient dabei als Grundlage für die Ermittlung der stadienbasierten transkriptomischen Veränderungen (vgl. 3.2.2). Darauf aufbauend erfolgen in dieser Arbeit weiterführende Untersuchungen, um die pathologischen Stadien anhand verschiedener molekularbiologischer Aspekte, wie stadienbasierte Gensignaturen und microRNA-Targets, zu erweitern (vgl. Abschnitt 3.2.3 und 3.2.4). Diese Aspekte können neue Erkenntnisse für die Diagnose, Prognose und/oder Therapie der Tumorerkrankung liefern und als Grundlagen für weiterführende Arbeiten Anwendung finden.

3.2.1. Sequenzierung der Neuroblastomproben

Die RNA- und microRNA-Sequenzierung der untersuchten Neuroblastomproben erzeugte durchschnittlich 9,3 Millionen Reads. Die Verteilung der einzelnen Reads jeder untersuchten Probe ist in der Übersicht in Abb. 3.11 dargestellt und zeigt die Anzahl der zugrunde liegenden und prozessierten Reads aus den Proben der RNA- und microRNA-Sequenzierung (vgl. 2.2.2). Alle folgenden Angaben ergeben sich aus der gerundeten durchschnittlichen Anzahl an Reads, prozentuale Anteile beziehen sich immer auf die Gesamtanzahl der aus den Sequenzierungen stammenden Reads.

Bei der Qualitätskontrolle wurden $3 \cdot 10^5$ (3 %) der RNA-Reads bzw. $4,9 \cdot 10^5$ (5 %) der microRNA-Reads durch Qualitäts- und Adaptertrimmen entfernt. Dieser sehr geringe Teil an entfernten Reads lässt - mit wenigen Ausnahmen - auf einen geringen Adaptergehalt und eine fehlerfreie Sequenzierung schließen. Das Alignment der verbleibenden $9 \cdot 10^6$ (97 %) der RNA-Reads und $8,8 \cdot 10^6$ (95 %) der microRNA-Reads ergab $7 \cdot 10^6$ (75 %) und $8,7 \cdot 10^6$ (94 %) Reads der RNA- und microRNA-Sequenzierung, die dem Humangenom zugeordnet werden konnten. Davon wurden im nachfolgenden Schritt $2,9 \cdot 10^6$ (31 %) und $1,4 \cdot 10^6$ (15 %) der RNA- bzw.

microRNA-Reads mittels Humangenom und miRBase annotiert. Verursacht wird der geringere Anteil an annotierten Reads dadurch, dass die Annotation der Transkriptomdaten sich auf codierende Bereiche (Exons) des Genoms bezieht (vgl. Abschnitt. 2.2.2), die nur einen geringen Teil ($\sim 1,5\%$) des menschlichen Genoms ausmachen [96]. Somit werden Reads intronischer und intergenischer Bereiche nicht annotiert und gezählt, was gleichermaßen auf Reads in Bereichen aktuell noch nicht annotierter Transkripte zutrifft. Die Rate der annotierten Reads der microRNA-Sequenzierung fällt insgesamt deutlich geringer aus. Grund hierfür ist die Annotation der microRNA-Reads auf bekannte microRNAs und nicht wie bei den RNA-Reads auf das gesamte annotierte Transkriptom. Die Annotation erfolgt mittels miRBase, die entgegen der Annotation des Humangenoms keine weiteren RNA-Typen außer microRNAs enthält. Um alle derzeit bekannten microRNAs in der Transkriptomanalyse zu annotieren, wurde der Schritt der Annotierung über die miRBase Datenbank gewählt. Diese umfasst in der hier verwendeten Version 2794 humane microRNAs (vgl. Abschnitt 2.2.2). Die für die RNA-Reads verwendete Annotation des Humangenoms beinhaltet hingegen nur rund 1400 gelistete microRNAs.

Zusammenfassend lässt sich in den Daten beider Sequenzierungen eine starke Varianz in der Readanzahl der einzelnen Proben erkennen. An dieser Stelle sei nochmals an die Notwendigkeit einer nachfolgenden Normalisierung der Daten erinnert, die grundsätzlich erfolgen sollte, um weiterführende Analysen zu ermöglichen, wie beispielsweise die Ermittlung differentiell exprimierten Transkripten anhand der klinischen Stadien. In der Tabelle 3.1 sind zudem die Daten der prozessierten Reads über alle Proben mit Informationen zu den jeweiligen Stadien zusammengefasst. Angegeben ist die durchschnittliche Anzahl an Reads, die nach jedem Prozessierungsschritt erhalten bleiben. Die RNA-Sequenzierungsdaten zeigen nach der vollständigen Prozessierung (vgl. dazu Abschnitt 2.2.2) im Vergleich zu den microRNA-Daten eine größere Schwankung in den fertig prozessierten Libraries.

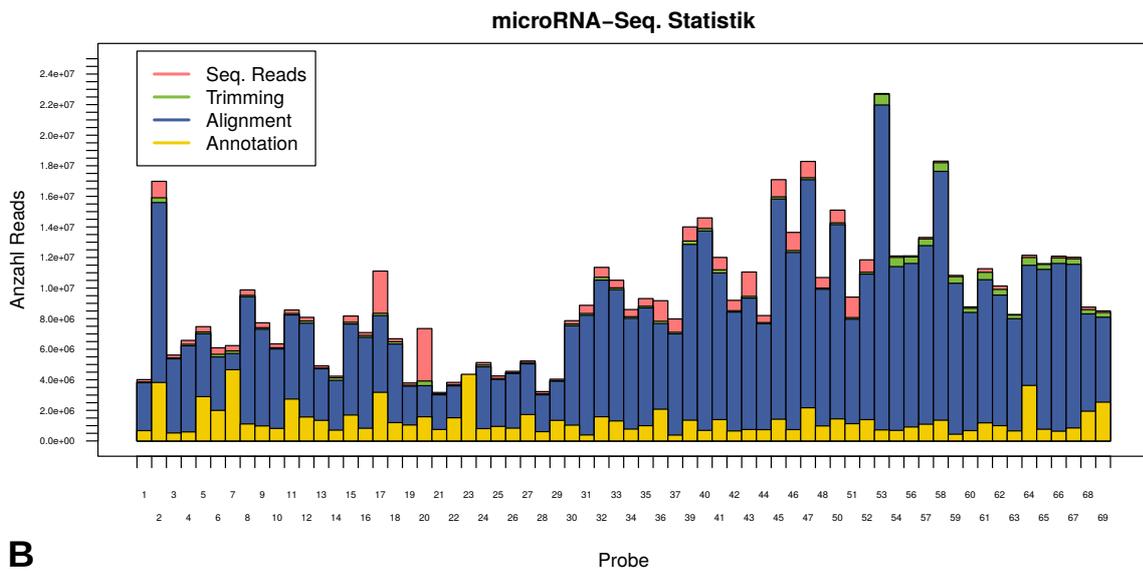
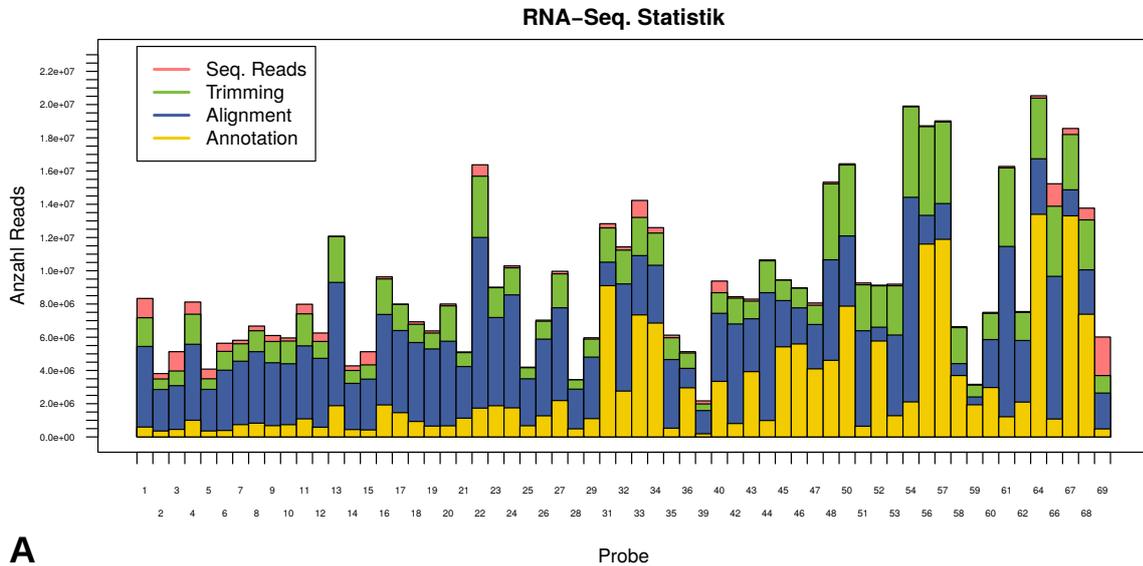


Abbildung 3.11.: Übersicht über die prozessierten Reads der sequenzierten primären Tumorproben der RNA- und microRNA-Sequenzierung. Die überlagerten Balken des Diagramms geben die Anzahl der Reads an, die eingangs für die Sequenzierungsanalyse zur Verfügung standen (rot) sowie die Anzahl der Reads, die nach dem Trimming (grün), dem Alignment (blau) und der Annotation (gelb) erhalten bleiben.

	gesamt	s1	s2	s3	s4	s4S	MYCN
RNA-Seq.	9,3·10 ⁶	7,6·10 ⁶	8,4·10 ⁶	7,2·10 ⁶	10,4·10 ⁶	11,2·10 ⁶	7,8·10 ⁶
Trimming	9,0·10 ⁶	7,0·10 ⁶	8,2·10 ⁶	6,9·10 ⁶	10,3·10 ⁶	10,6·10 ⁶	7,7·10 ⁶
Alignment	7,0·10 ⁶	5,6·10 ⁶	6,6·10 ⁶	5,8·10 ⁶	7,8·10 ⁶	8,4·10 ⁶	6,3·10 ⁶
Annotation	2,9·10 ⁶	1,9·10 ⁶	2,6·10 ⁶	1,9·10 ⁶	3,2·10 ⁶	4,5·10 ⁶	2,6·10 ⁶
miRNA-Seq.	9,3·10 ⁶	8,3·10 ⁶	8,0·10 ⁶	10,9·10 ⁶	10,2·10 ⁶	8,2·10 ⁶	11,2·10 ⁶
Trimming	8,9·10 ⁶	7,9·10 ⁶	7,4·10 ⁶	10,2·10 ⁶	9,6·10 ⁶	8,1·10 ⁶	10,4·10 ⁶
Alignment	8,7·10 ⁶	7,8·10 ⁶	7,3·10 ⁶	10,1·10 ⁶	9,4·10 ⁶	7,9·10 ⁶	10,2·10 ⁶
Annotation	1,4·10 ⁶	1,7·10 ⁶	1,2·10 ⁶	1,1·10 ⁶	1,3·10 ⁶	1,4·10 ⁶	1,3·10 ⁶

Tabelle 3.1.: Tabellarische Übersicht über die prozessierten Reads der sequenzierten primären Tumorproben der RNA- und microRNA-Sequenzierung. Angegeben ist die gerundete mittlere Anzahl an Reads die sequenziert wurden, die nach dem Trimming erhalten bleiben sowie die Reads, die dem Humangenom zugeordnet und anschließend annotiert wurden. Die Werte beziehen sich auf die verschiedenen Stadien (s1, s2, s3, s4, s4S) der 61 RNA bzw. 66 microRNA Proben. Die letzte Spalte kennzeichnet die Neuroblastomproben, die eine MYCN Amplifikation aufweisen.

Abbildung 3.12 veranschaulicht für einen Überblick der sequenzierten Transkripte der untersuchten Neuroblastomproben die durchschnittliche Verteilung der annotierten Reads der Total-RNA-Sequenzierung auf diverse bekannte RNA-Typen. Die Bezeichnungen der einzelnen Typen resultieren aus der verwendeten Annotation des Humangenoms. Der größte Anteil der Reads von rund 64 % entfällt auf mRNAs Protein-codierender Gene (*protein_coding*), gefolgt von 9,85 % kleine nukleäre RNAs (*snRNA*), 9,20 % unterschiedlicher RNAs (*misc_RNA*) wie beispielsweise kleine nicht-codierende YRNAs, 7SK RNAs oder intronische Transkripte. Einen geringen Teil nehmen 6,82 % Pseudogene (*pseudogene*) und 4,75 % Antisense-RNAs (*antisense*) ein. Ribosomale RNA (*rRNA*) liegt in sehr geringem Anteil vor, da im Vorfeld der Sequenzierung eine Entfernung der rRNAs mittels eines Kits zum Abbau ribosomaler RNAs stattfand (vgl. Abschnitt 2.2.1). Die analysierten mRNA- und microRNA-Daten aus der RNA-Sequenzierungsanalyse sind über die *R2 Genomics Analysis and Visualization Platform*⁴⁴ öffentlich verfügbar⁴⁵. Die Plattform bietet verschiedene Möglichkeit, die Daten weiterführend zu analysieren, beispielsweise Vergleiche von Expressionen verschiedener Gene zu erstellen oder Überlebenszeitanalysen durchzuführen. Die Daten können in vielfältiger Weise visualisiert und in tabellarischer Form vollständig oder benutzerdefiniert exportiert werden.

⁴⁴<http://r2.amc.nl>

⁴⁵Registrierung per Mail, unter Angaben von Name und Institution. Der Zugang zu den hier beschriebenen Daten erfolgt auf zusätzliche Anfrage an Jan Koster (jankoster@amc.uva.nl).

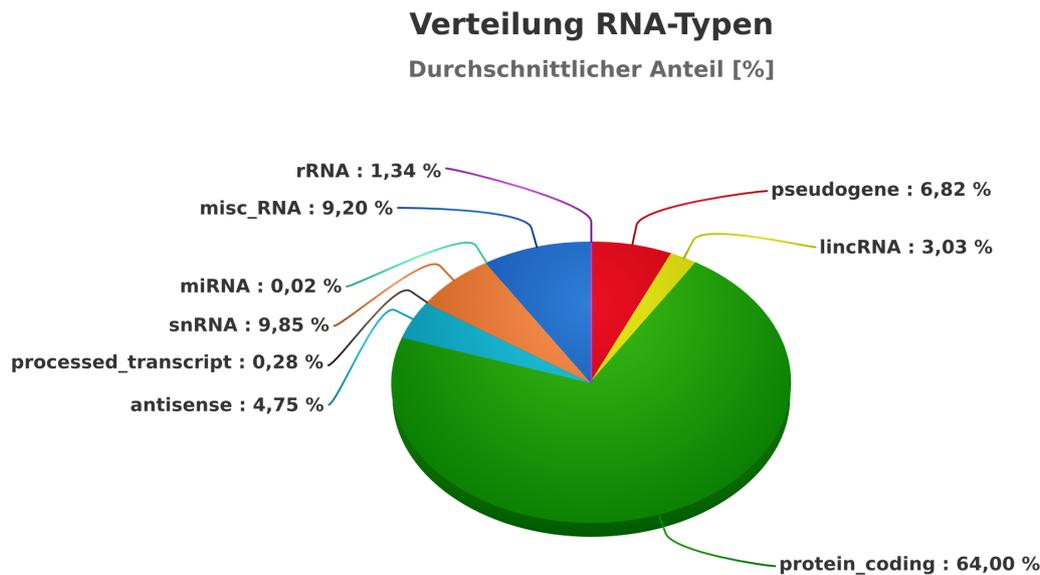


Abbildung 3.12.: Durchschnittliche Verteilung der annotierten Reads aller Proben der RNA-Sequenzierung auf unterschiedliche RNA-Typen, entsprechend ihrer Bezeichnungen in der verwendeten Annotation.

3.2.2. Differentielle Expression im Neuroblastom

Die differentielle Expression der Neuroblastomdaten wurde für neun Vergleiche (vgl. 2.2.3) zwischen den Proben der verschiedenen klinisch relevanten Tumorstadien ermittelt (vgl. hierzu Abschnitt 1.2). Der Test auf differentielle Expression erfolgte dabei auf Grundlage der TMM normalisierten Daten unter Anwendung eines FDR-Wertes $\leq 0,05$ (vgl. 2.2.3). Ein Verweis auf die Ergebnistabellen der sich daraus ergebenden differentiiell exprimierten mRNAs (DE-mRNAs) und microRNAs (DE-microRNAs) kann aus dem Anhang C entnommen werden. Tabelle 3.2 gibt eine Übersicht über die Anzahl der ermittelten DE-mRNAs und DE-microRNAs, die in den einzelnen Vergleichen differentiiell hoch- bzw. runterreguliert sind. Bei den mRNAs weist der Vergleich $1 - 4^{-M}$ die größte Anzahl an DE-mRNAs auf, während $1 - 4S$ die geringste Anzahl an DE-mRNAs aufweist. Der Vergleich $1 - MYCN$ weist bei den DE-microRNAs die größte Anzahl an differentiiell exprimierten microRNAs auf, wobei die geringste Anzahl an DE-microRNAs in den Vergleichen $1 - 4S$ und $4^{-M} - 4S$ auftritt. Die Vergleiche $1 - 4^{-M}$ und $1 - MYCN$ mit den meisten differentiiell exprimierten Transkripten sind mit einer schlechten bis sehr schlechten Prognose und einem ungünstigen Verlauf der Krankheit assoziiert. Das sogenannte (hoch)risikoreiche Neuroblastom (*high risk*) ist durch die Amplifikation des *MYCN* Gens gekennzeichnet. Im Vergleich $1 - MYCN$ werden Proben des risikoarmen Stadiums 1

Vergleich	mRNAs			microRNAs		
	DE \uparrow	DE \downarrow	\sum DE	DE \uparrow	DE \downarrow	\sum DE
1/2/3 – 4/4S	27	15	42	16	21	37
1 – MYCN	1220	481	1701	126	89	215
1 – 4	1155	202	1357	81	71	152
1 – 4 ^{+M}	983	527	1510	110	51	161
1 – 4 ^{-M}	2009	138	2147	52	55	107
1 – 4S	20	1	21	5	6	11
4 – 4S	88	13	101	20	32	52
4 ^{+M} – 4S	522	153	675	36	80	116
4 ^{-M} – 4S	25	4	29	2	9	11

Tabelle 3.2.: Differentielle Expression (DE) der mRNAs Protein-codierender Gene und microRNAs im Neuroblastom. Angegeben sind die Anzahl der hoch- bzw. runterregulierten (\uparrow/\downarrow) mRNAs und microRNAs sowie die Gesamtanzahl der differentiellen Transkripte, entsprechend der vorgenommenen Einteilungen der Vergleiche zwischen den klinisch relevanten Stadien.

(*low risk*) mit denen des risikoreichen Neuroblastoms verglichen. Die differentielle Expression zwischen diesen Stadien kann entscheidende Transkripte liefern, die für die Prognose und den Verlauf sowie für die Behandlung der Krebserkrankung eine wesentliche Rolle spielen (vgl. Abschnitt 1.2). Unter Betrachtung bekannter mRNAs Protein-codierender Gene und microRNAs aus der Literatur kann anhand des Vergleiches 1 – MYCN exemplarisch gezeigt werden, dass die differentielle Expression der Sequenzierungsdaten valide Ergebnisse liefert. Im Folgenden soll dieser Vergleich zur Validierung detaillierter betrachtet werden. Auf eine genauere Untersuchung der anderen Vergleiche zwischen den oftmals weniger gut beschriebenen Stadien soll hier verzichtet werden, da es den Rahmen der vorliegenden Arbeit überschreiten würde. Für das aggressive risikoreiche Neuroblastom ist bekannt, dass aufgrund der *MYCN* Amplifikation Gene innerhalb der chromosomalen Region 2p24 co-amplifiziert werden. Dies geschieht durch die Nähe der Gene zur amplifizierten Region, was zur Folge hat, dass die flankierenden Regionen ebenfalls vervielfältigt werden und die Amplifikation des *MYCN* Gens mit der Überexpression benachbarter Gene korreliert. Bekannte *MYCN* co-amplifizierten Gene sind beispielsweise *DDX1* und *ODC1* [123]. Des Weiteren weisen direkt mit dem risikoreichen Neuroblastom assoziierte Gene im Vergleich zu anderen Stadien oder Normalgeweben starke differentielle Expressionen auf und zeigen ein hohes onkogenes oder tumorsuppressives Potential. Onkogene beschreiben dabei Gene, die in der Lage sind Zellen zu transformieren und dadurch Krebs im Organismus auszulösen, indem deren Genprodukte Signalwege des Zellwachstums oder der Apoptose kontrollieren. Häufig werden Onkogene als Ergebnis von Mutationen, Genfusionen

oder durch Amplifikation und den damit einhergehenden strukturellen Veränderungen aktiviert. Tumorsuppressoren zeichnen sich im Wesentlichen dadurch aus, dass sie die Zellteilung und das Wachstum verlangsamen und Prozesse der Apoptose und DNA-Reparatur positiv beeinflussen. Die Inaktivierung tumorsuppressiver Gene trägt dabei ebenso zur Krebsentstehung bei, wie die Aktivierung von Onkogenen [265]. Die Abbildungen 3.13 und 3.14 geben hierzu einen Überblick über die ermittelten Veränderungen der mRNAs bzw. microRNAs innerhalb des betrachteten Vergleiches 1 – MYCN. Die Veranschaulichung mittels Vulkandiagramm (auch als *Volcano Plot* bezeichnet) fasst die Daten der ermittelten differentiellen Expression der einzelnen Transkripte anhand des logarithmischen Fold Changes und der False Discovery Rate zusammen. Bereits aus der Literatur bekannte, Neuroblastom assoziierte, onkogene (grün) und tumorsuppressive (rot) mRNAs/microRNAs dienen dabei zur Überprüfung der Validität der komplexen Expressionsdaten aus den Sequenzierungen. Blau gekennzeichnete Transkripte stellen putativ neue und noch nicht mit dem Neuroblastom assoziierte mRNAs bzw. microRNAs dar, welche aus den 20 signifikantesten hoch- und den 20 signifikantesten runterregulierten DE-mRNAs und DE-microRNAs ermittelt wurden ($FDR \leq 0,05$, vgl. 2.2.3) und auf die im späteren Verlauf noch detaillierter eingegangen wird.

Aus Abbildung 3.13 des Vergleiches 1 – MYCN geht für die mRNA Expressionen hervor, dass die mit *MYCN* co-amplifizierten Gene *DDX1* und *ODC1* eine ähnlich hohe und signifikante differentielle Expression wie *MYCN* aufweisen, wie es durch die Co-Amplifikation zu erwarten ist. Des Weiteren lassen sich bekannte Onkogene identifizieren, die im Vergleich 1 – MYCN differentiell exprimiert sind. Als Beispiele seien hier genannt *PLK1* (*Polo Like Kinase 1*), das als mögliches therapeutisches Target angesehen wird [1, 203], *E2F3* (*E2F Transcription Factor 3*), welches wesentlich bei der Aktivierung und der Regression von *MYCN* im Neuroblastom beteiligt ist [235] oder das Gen *ALK* (*Anaplastic Lymphoma receptor Tyrosine Kinase*), welches in diversen Tumoren mutiert und/oder amplifiziert ist und therapeutisch eingesetzt werden kann, um das fortgeschrittene Neuroblastom zu behandeln [24, 39]. Auffällig in Abb. 3.13 ist der kaum veränderte anti-apoptotische Faktor *BCL2* (*B-Cell Lymphoma 2*), welcher als Onkogen bekannt und im Neuroblastom beschrieben ist [119]. Im Vergleich 1 – MYCN zeigt dieser Faktor jedoch eine geringe und nicht signifikante Veränderung. Grund hierfür ist, dass die Expression des Faktors generell im Neuroblastom erhöht ist, wobei vor allem Tumoren eine hohe *BCL2*-Expression aufweisen, sich in den meisten Zelllinien jedoch nur eine geringe bis gar keine Expression des Faktors nachweisen lässt [119]. Im hier gezeigten Vergleich liegt somit im

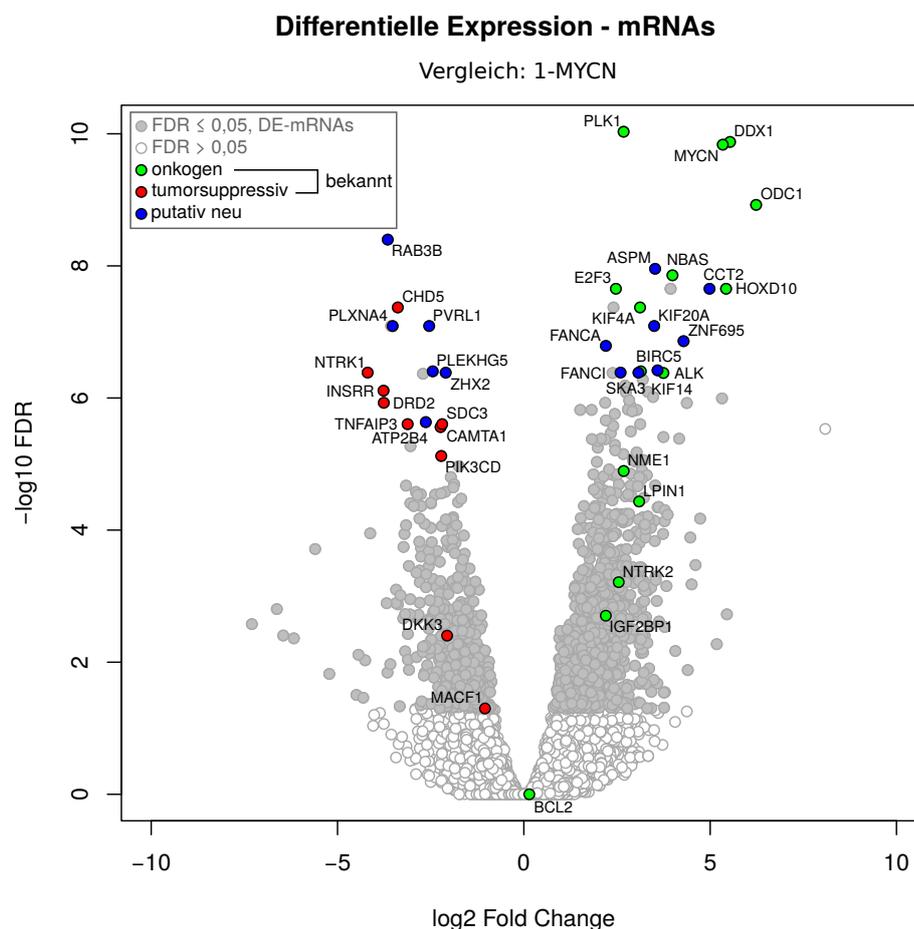


Abbildung 3.13.: Vulkandiagramm der Gene der differentiell exprimierten mRNAs (grau) aus dem hier betrachteten Vergleich 1 – MYCN zwischen dem wenig aggressiven risikoärmeren Stadium 1 und dem risikoreichen und aggressivsten Neuroblastom mit MYCN Amplifikation. Hierfür bereits bekannte Gene mit onkogener Wirkung sind grün, solche mit tumorsuppressiver Wirkung rot gekennzeichnet. Mögliche neue und bis jetzt nicht direkt mit dem hochrisikoreichen Neuroblastom assoziierte Gene sind blau markiert. Die Gene sind logarithmisch entsprechend des Fold Changes und der False Discovery Rate der DE-mRNAs zwischen den Stadien aufgetragen.

Stadium 1 und Proben mit *MYCN* Amplifikation eine ähnliche Expression des Faktors vor. Die ermittelten Mediane der normalisierten Expressionswerte in den hier analysierten Proben liegen bei 105,67 bzw. 127,94 FPM⁴⁶ für Proben des Stadium 1 respektive *MYCN* amplifizierte Proben.

Veränderungen bekannter tumorsuppressiver Gene spielen ebenso wie onkogene Faktoren eine wesentliche Rolle bei der Prognose und Therapie der Tumorerkrankung. Für das risikoreiche Neuroblastom ist bekannt, dass die schwache Expression von *DKK3* (*Dickkopf WNT Signaling Pathway Inhibitor 3*) mit einer schlechten Prognose des Krankheitsverlaufes einhergeht. *DKK3* ist an der Embryonalentwicklung beteiligt und steht in Wechselwirkungen mit dem Wnt-Signalweg, der zur Reaktion der Zelle auf äußere Einflüsse mittels Signaltransduktion dient. Die *DKK3* Expression ist in diversen Krebszelllinien vermindert und ist negativ mit der *MYCN* Expression korreliert, da die Expression von *DKK3* durch den *MYCN* induzierten miR-17~92 Cluster reguliert wird [51, 114]. Weiterer bekannte Tumorsuppressoren, die der differentiellen Expression in 1 – *MYCN* entnommen werden können, sind beispielsweise *CHD5* (*Cadherin 5*), das eine geringe bis gar keine Expression im risikoreichen Neuroblastom aufweist und dessen Expression mit einem günstigen Verlauf der Krankheit in Verbindung steht [69]. Das Gen *NTRK1* (*Neurotrophic Receptor Tyrosine Kinase 1*), welches mit diversen Krebserkrankungen assoziiert ist und dessen Expression durch die direkte Bindung von *MYCN* an den *NTRK1* Promoter negativ beeinflusst werden kann. *NTRK1* ist an der Zelldifferenzierung beteiligt und ist in der Lage die Immunogenität der Zellen zu erhöhen, was zu einem verbesserten Verlauf der Krankheit führen oder eine spontane Regression auslösen kann [188, 222]. Neben diesen Beispielen treten noch weitere bekannte, auch mit dem Neuroblastom assoziierte, suppressive Faktoren auf, die sich anhand der hier gewonnenen Daten als Gene mit suppressivem Potential bestätigen lassen, wie z. B. *INSSR* [264], *CAMTA1* [87], *SDC3* [112] und *PIK3CD* [67].

Im Review von [155] werden charakteristische Neuroblastom assoziierten microRNAs und ihre wesentlichen Funktionen sowie Wirkungsweisen als Onkogene oder Tumorsuppressoren zusammengefasst. Der onkogene miR-17~92 Cluster des risikoreichen Neuroblastoms ist weithin gut untersucht und setzt sich aus den microRNAs *hsa-miR-17-5p*, *-18a*, *-19a*, *-19b-1*, *-20a*, *-92a-1* zusammen. Die diversen microRNAs des Cluster können durch Amplifikation eine

⁴⁶Fragments per Million (FPM) gibt die Anzahl der annotierten Paired-End mRNA Reads, normalisiert auf eine Million Reads, an.

erhöhte Expression aufweisen oder direkt durch das *MYCN* Gen transaktiviert werden. Der miR-17~92 Cluster kennzeichnet einen ungünstigen Verlauf und eine schlechte Prognose der Krankheit [155,224]. Im risikoreichen Neuroblastom sind Teile des Clusters stark hochreguliert, was sich ebenfalls anhand des betrachteten Vergleiches 1 – MYCN der analysierten Daten zeigen lässt (vgl. Abb. 3.14). Die microRNAs *hsa-miR-17-5p*, *-18a-5p*, *-18a-3p*, *-19a-3p*, *-19b-3p*, *-20a-5p* weisen in den Daten eine starke differentielle Expression auf, welche dem Großteil der Komponenten des Clusters entsprechen [46, 159, 165]. Die ebenfalls zum Cluster gehörende *hsa-miR-92a-1* weist eine geringere differentielle Expression auf, während die microRNAs *hsa-miR-19a-5p*, *-19b-1-5p* und *-20a-3p* in den vorliegenden Daten nicht signifikant exprimiert sind. Mögliche Ursache hierfür könnte eine Teilregulation des Clusters sein, die zu unterschiedlichen Expressionen einzelner Komponenten führen kann, oder aus Unterschieden im eigentlichen Reifungsprozess sowie der Stabilität der reifen microRNA resultieren [46].

Weiterhin sind microRNAs der *miR-181* Familie bekannt, welche eine regulatorische Funktion auf posttranskriptioneller Ebene ausüben und in *MYCN* amplifizierten Tumoren sowie Tumoren mit ungünstiger Prognose hochreguliert sind [38,155]. Von dieser Familie zeigen die microRNAs *hsa-miR-181a-3p*, *-181a-5p*, *-181b-3p*, und *-181b-5p* eine differentielle Expression im Vergleich 1 – MYCN, wobei die microRNAs in den Proben des risikoreichen Neuroblastoms eine hochregulierte Expression aufweisen. Des Weiteren sind die microRNAs *hsa-miR-21 (-3p/-5p)*, *hsa-miR-380-5p* und *hsa-miR-15a* als onkogene microRNAs bekannt [155]. Die *hsa-miR-21* hat eine proliferierende Wirkung und weist in diversen Krebserkrankungen eine erhöhte Expression auf, im Neuroblastom vorwiegend in Tumoren mit einem ungünstigen Verlauf. Bei der *hsa-miR-380-5p* handelt es sich um eine microRNA, welche die p53 vermittelte Apoptose der Zellen unterdrückt und die eine erhöhte Expression in *MYCN* amplifizierten Tumoren aufweist. Die microRNA *hsa-miR-15a* zeigt hingegen eine Hochregulierung im metastasierenden Neuroblastom. Alle drei microRNAs sind in den analysierten Proben der Stadien des Vergleiches 1 – MYCN exprimiert, weisen jedoch keine signifikante Veränderung zwischen den Stadien und somit keine differentielle Expression auf. Die erhöhte Expression der microRNAs deutet zwar auf ein onkogenes Potential hin, jedoch würde man in diesem risikoreichen und aggressiven Stadium mit *MYCN* Amplifikation eine viel stärkere und signifikante Veränderung im Vergleich zum Stadium 1 erwarten. Anhand der hier vorliegenden Daten kann somit ein onkogenes Potential dieser microRNAs, wie es in der Literatur beschrieben ist, basierend auf deren Expression nicht bestätigt werden. Möglicherweise könnten sich Beeinflussungen in der microRNA Biogenese

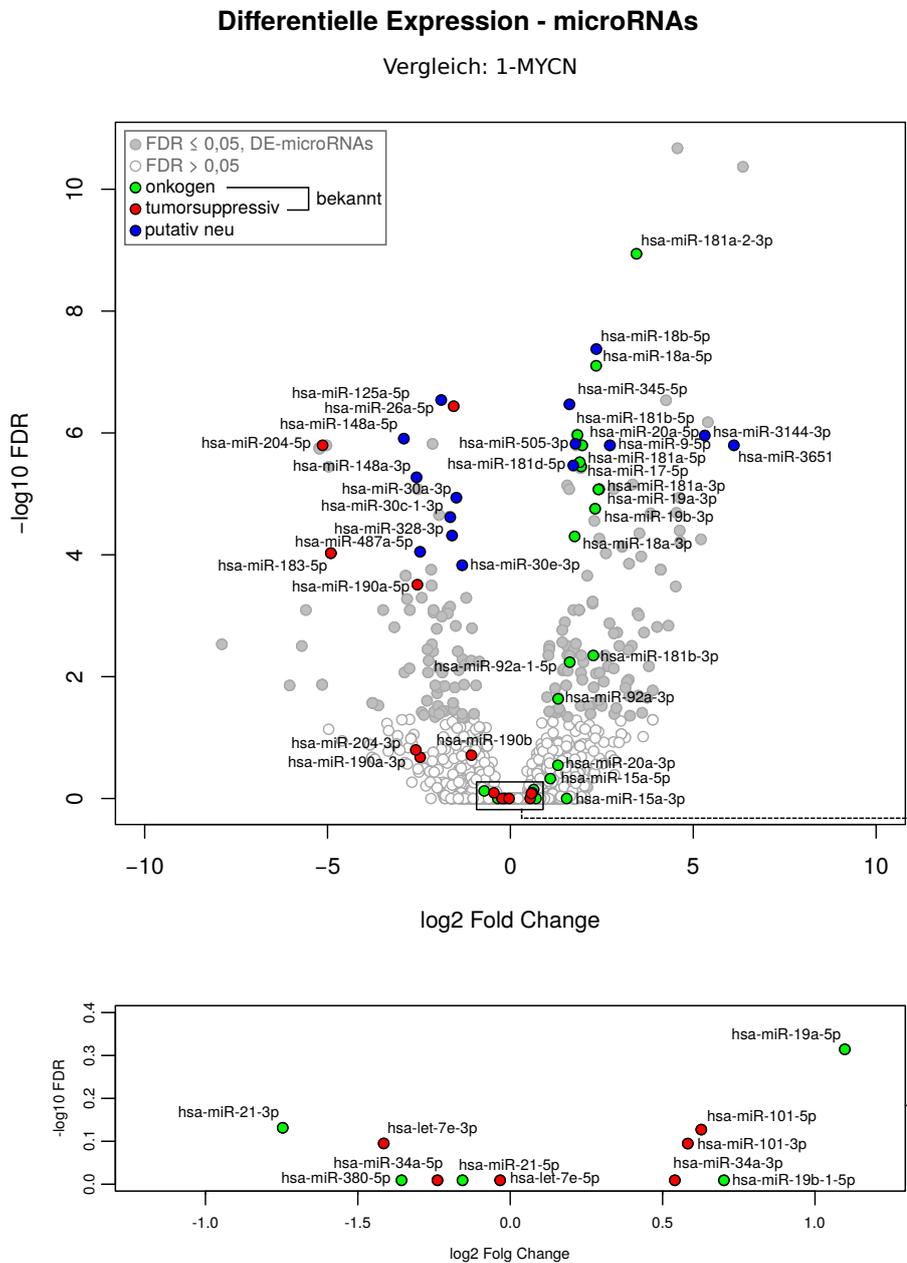


Abbildung 3.14.: Vulkandiagramm der differentiell exprimierten microRNAs (grau) aus dem Vergleich 1 – MYCN des risikoreichen Neuroblastoms. Bereits bekannte microRNAs mit onkogenen Wirkung sind grün, solche mit tumorsuppressiver Wirkung rot gekennzeichnet. Mögliche bis jetzt nicht direkt mit dem risikoreichen Neuroblastom assoziierte microRNAs sind blau markiert. Die microRNAs sind logarithmisch entsprechend des Fold Changes und der False Discovery Rate der DE-microRNAs zwischen den Stadien aufgetragen. Zur besseren Darstellung einzelner microRNAs zeigt der untere Bereich die Vergrößerung der oben markierten Region.

negativ auf die Expression auswirken und/oder die Heterogenität der Sequenzierungsdaten innerhalb der einzelnen Proben beeinflusst das Ergebnis der differentiellen Expression. Beide Punkte können an dieser Stelle nicht eindeutig geklärt werden und die Expression der microRNAs sollten hierfür nachträglich anhand quantitativer Verfahren, wie z. B. mittels quantitativer PCR-Methoden, experimentell überprüft werden.

Mithilfe der differentiellen Expression der microRNAs lassen sich auch bekannte tumorsuppressive microRNAs identifizieren, deren Expression im Vergleich 1 – MYCN runterreguliert ist. Beispielsweise stellt die microRNA *hsa-miR-190* eine der häufigsten signifikant runterregulierten microRNAs im aggressiven risikoreichen Neuroblastom dar und zeigt in den analysierten Daten eine starke differentielle Expression zwischen den beiden Stadien. Durch das microRNA Target *NEUROD1* erfolgt bekanntermaßen eine indirekte Aktivierung des TrkB Tyrosinkinase-Rezeptors, welcher bei der Entwicklung und Funktion von Neuronen beteiligt ist [155, 230]. Eine weitere bedeutende microRNA mit tumorsuppressiver Wirkung ist die *hsa-miR-204*. Sie dient der Vorhersage des Verlaufs der Krankheit und funktioniert in Teilen sensitivitätssteigernd für den Zellwachstumshemmer Cisplatin durch eine direkte Runterregulation der anti-apoptotischen Faktors *BCL2*. Darüber hinaus hat die microRNA eine direkte Wirkung auf das starke *NTRK2* Onkogen, welches in die Chemotherapieresistenz des Neuroblastoms involviert ist [216]. Im Vergleich 1 – MYCN weisen die drei microRNAs *hsa-miR-34a*, *hsa-miR-101* und *hsa-let-7* eine schwache Expression auf, zeigen jedoch keine signifikante Veränderung in den hier analysierten Daten. Die microRNA *hsa-miR-34a* ist im risikoreichen Neuroblastom runterreguliert und inhibiert die Zellproliferation, induziert Apoptose und scheint eine direkte Runterregulation des *MYCN* Onkogens zu bewirken [155, 224]. Die tumorsuppressiven microRNAs *hsa-miR-101* und *hsa-let-7* beeinflussen ebenfalls direkt die *MYCN* Expression und hemmen die Proliferation im *MYCN* amplifizierten Neuroblastom [11, 12, 26, 155].

Generell spiegeln die Daten der differentiellen Expressionsanalyse dieser Arbeit für die mRNA- und microRNA-Sequenzierung valide Ergebnisse für beschriebene Transkripte der Literatur im Vergleich 1 – MYCN wieder. Über die Bestätigung bekannter Transkripte hinaus können anhand der differentiellen Expression auch putativ neue mRNAs bzw. microRNAs identifiziert werden. Wobei diese in einem möglichen Zusammenhang mit dem risikoreichen Neuroblastom stehen könnten, aber diese aktuell noch nicht oder nicht nachgewiesenermaßen direkt in Verbindung mit dem risikoreichen Neuroblastom gebracht werden. Zur Identifizierung putativ neuer

Transkripte wurden aus den signifikantesten 20 hoch- und 20 runterregulierten DE-mRNAs und DE-microRNAs die ausgewählt, die anhand von Literaturrecherchen keine direkte Assoziation mit dem risikoreichen Neuroblastom in der Literatur aufweisen, jedoch im Zusammenhang mit anderen Tumorerkrankungen publiziert sind. Die putativ neuen Transkripte aus dem Vergleich 1 – MYCN sind in den Abbildungen 3.13 und 3.14 blau markiert und mit ihren bisher bekannten biologischen Assoziationen und/oder ihrem Vorkommen in Tab. 3.3 aufgeführt. Unter den identifizierten Transkripten befinden sich sowohl putativ neue onkogen als auch tumorsuppressiv wirkende, wie beispielsweise das hochregulierte *ASPM* (*Abnormal Spindle Microtubule Assembly*), welches eine zentrale Rolle bei der Zellteilung der Gehirnentwicklung einnimmt. Das Gen weist eine erhöhte Expression in diversen Krebserkrankungen, wie dem Glioblastom oder dem hepatozelluläres Karzinom auf, und ist mit dem Fortschreiten, der Metastasenbildung und der Wiederkehr der Erkrankungen assoziiert [82, 89, 134]. Das Gen *CCT2* (*Chaperonin Containing TCP1 Subunit 2*) weist ebenfalls eine erhöhte Expression in den Daten auf. Es ist Teil des TCP1 Chaperonin-Komplexes (TRiC Ring-Komplex) und ist an der ATP-abhängigen Faltung von Polypeptiden beteiligt. *CCT2* kann als therapeutisches Target beim Mammakarzinom eingesetzt werden und ist mit dem Adenokarzinom assoziiert [77, 290]. Gene der mikrotubuli-bindenden Kinesine *KIF14* und *KIF20A* (*Kinesin Family Member 14/20A*) weisen ebenso eine starke differentielle Expression im Vergleich 1 – MYCN auf. Sie haben eine regulatorische Funktionen bei der Zellteilung und weisen eine onkogene Wirkung sowie mögliche Eigenschaften für die Krebstherapie auf. Die Kinesine sind mit diversen Tumorerkrankungen assoziiert, wie Bauchspeicheldrüsenkrebs, Hautkrebs und dem Adenokarzinom der Lunge [94, 138, 207, 233]. *KIF14* könnte dabei möglicherweise mit dem Neuroblastom assoziiert sein, wobei bisher kein Nachweis der biologischen Relevanz erfolgte [207] und es aus diesem Grund genannt sei. Einen Hinweis auf ein mögliches neues tumorsuppressives Gen im risikoreichen Neuroblastom könnte das am stärksten differentiel exprimerte Gen *RAB3B* (*Ras-related protein Rab-3B*) liefern. Das Gen ist Mitglied der RAS Onkogen-Familie und ist an der Regulation der Exocytose beteiligt. *RAB3B* Steht in Verbindung mit dem Pankreaskarzinom und der megakaryoblastischen Leukämie. Während der Zelldifferenzierung weist *RAB3B* eine erhöhte Expression auf und es gibt bereits Anzeichen, dass es möglicherweise auch mit dem Neuroblastom assoziiert sein könnte [123].

Neben den differentiel exprimierten mRNAs können auch microRNAs identifiziert werden, für die noch keinen direkten Zusammenhang mit dem risikoreichen Neuroblastom beschrieben

wurde. Onkogen wirkende microRNAs sind beispielsweise die *hsa-miR-18b-5p* oder *hsa-miR-345-5p*. Beide sind wichtige biologische Marker bei der Teilung und dem Wachstum von Zellen. Die Expression der *hsa-miR-18b* ist für die Vorhersage der klinischen Prognose und für die Diagnose im hepatozelluläres Karzinom entscheidend [172]. Die *hsa-miR-345-5p* ist mit der Proliferation und der Tumordinvasion des kolorektalen Karzinoms assoziiert und hat eine wesentliche Bedeutung bei der antineoplastisch Wirkung als Wachstumsinhibitor während der Karzinomentwicklung [243]. Die microRNA *hsa-miR-3144-3p* weist eine erhöhte Expression im Adenokarzinome des Dickdarms, dem hepatozellulären Karzinom und bei der Leberzirrhose auf [83, 152, 270]. Durch die Assoziation mit diversen Tumoren könnte die differentielle Expression im Vergleich 1 – MYCN auf eine ebenfalls vorhandene Wirkung im risikoreichen Neuroblastom hinweisen. Auch die microRNA *hsa-miR-9-5p* weist eine hohe Expression zwischen diesen Stadien auf. Sie ist innerhalb des Gehirns stark exprimiert und reguliert die neuronale Differenzierung. Sie wird zudem mit einem erhöhten Risiko für Schizophrenie in Verbindung gebracht, wobei sie im Gehirn regulatorisch auf die Neurogenese neuronaler Stammzellen wirkt [251]. Im kolorektalen Karzinom kommt dieser microRNA während der Krebsentstehung eine wesentliche Bedeutung in der Entwicklung des Darmkrebses zu, bedingt durch die Förderung des Zellwachstums und dem Fortbestand der Krebszellen [33]. Für die microRNA *miR-3651-5p* gibt es aktuell nur wenig Daten für deren Rolle in humanen Krebserkrankungen. Es wurde eine erhöhte Expression im kolorektalen Karzinom ermittelt, die es ermöglicht zwischen Normal- und Tumorgewebe zu unterscheiden, was einen möglichen Hinweis auf eine Krebsassoziation der microRNA liefert [191, 219]. Die microRNA kann als potentieller biologischer Marker im Plattenepithelkarzinom dienen, und es konnte gezeigt werden, dass eine erhöhte Expression eine schwache Korrelation mit Lymphknotenmetastase und dedifferenzierten Tumorzellen aufweist [209]. Identifizierte tumorsuppressive microRNAs sind beispielsweise die *hsa-miR-125a-5p* und die microRNA *hsa-miR-148a*. Die *hsa-miR-125a-5p* ist ein bekannter Tumorsuppressor in diversen Krebserkrankungen, wie beispielsweise dem Mamma- oder Ovarialkarzinom [182]. Es gibt Anzeichen, dass die *hsa-miR-125a-5p* in SK-N-BE Zellen des Neuroblastom unter Behandlung mit Retinsäure hochreguliert ist und die Zellproliferation kontrolliert [121]. Weiterhin ist die microRNA ein bedeutender prognostischer Marker bei der Behandlung des Magenkarzinoms und spielt eine Rolle bei der Invasion und Metastasierung im Bronchialkarzinom. Eine geringe Expression kann mit der Tumorgroße, Tumordinvasion, Lebermetastasen und einer schlechten Prognose assoziiert werden [98, 182].

	log2FC	FDR	biol. Assoziation / Vorkommen	Referenz
mRNA				
ASPM	3,52	1,11E-8	Zellteilung/div. Krebserkrankungen	[82, 89, 134]
CCT2	4,98	2,22E-8	Peptidfaltung/Mamma- und Adenokarzinom	[77, 290]
KIF20A	3,50	8,13E-8	Motorprotein/div. Krebserkrankungen	[138, 207, 233]
ZNF695	4,29	1,38E-7	Ovarialkarzinom	[101]
FANCA	2,20	1,63E-7	DNA-Raplikation/div. Krebserkrankungen	[249, 257]
SKA3	3,59	3,82E-7	Mikrotubuli, Onkogen/Prostatakarzinom	[80, 126, 202]
FANCI	2,59	4,12E-7	DNA-Raplikation/div. Krebserkrankungen	[32, 45, 249]
KIF14	3,08	4,14E-7	Motorprotein/div. Krebserkrankungen	[94, 138, 207]
RAB3B	-3,65	4,00E-9	Exocytose, Differenzierung/Pankreaskarzinom/ megakaryoblastische Leukämie	[123]
PVRL1	-2,54	8,13E-8	Herpes-simplex-Virus	[54, 250]
PLXNA4	-3,52	8,14E-8	Tumorentwicklung/Alzheimer, Parkinson	[102, 107, 223]
PLEKHG5	-2,45	3,95E-7	NFκB Pathway	[151]
ZHX2	-2,10	4,14E-7	Differenzierung/Hepatozelluläres Karzinom, Hodgkin-Lymphom	[90, 174], [278, 280]
ATP2B4	-2,63	2,32E-6	Calcium Dysregulation/div. Erkrankungen	[5, 72, 129]
microRNA				
miR-18b-5p	2,35	4,20E-8	Biomarker/hepatozelluläres Karzinom	[172]
miR-345-5p	1,61	3,40E-7	Biomarker/kolorektales Karzinom	[243]
miR-3144-3p	5,31	1,10E-6	div. Krebserkrankungen	[83, 152, 270]
miR-9-5p	2,72	1,60E-6	Differenzierung, Tumorwachstum/Schizophrenie	[33, 251]
miR-3651	6,11	1,60E-6	mögliche Krebsassoziation	[191, 209, 219]
miR-505-3p	1,80	1,60E-6	Biomarker/Primär biliäre Cholangitis, Synovial-Sarkom	[68, 181]
miR-181d-5p	1,73	3,48E-6	mögliche Assoziation mit Basalzellenkarzinom, Phäochromozytom und Neuroblastom	[218, 240]
miR-125a-5p	-1,89	2,88E-7	Tumorsuppressor, Proliferation/div. Tumore	[98, 121, 182]
miR-148a(-3p/5p)	-2,91	1,24E-6	Metastasierung, Biomarker/ hep. Karzinom, Pankreaskarzinom	[125, 160]
miR-30a-3p	-1,48	1,15E-5	Biomarker, Migration, EMT/ Bronchialkarzinom, Mammakarzinom	[37, 41], [194, 244]
miR-30c-1-3p	-1,66	2,35E-5	Biomarker, Resensibilisierung, Invasivität/ Mammakarzinom, kolorektales Karzinom	[19, 56], [213, 284]
miR-328-3p	-1,60	4,83E-5	Biomarker, Metastasierung/Bronchialkarzinom	[238, 252, 276]
miR-487a-5p	-2,47	8,92E-5	EMT, Invasivität, Resensibilisierung/ Mammakarzinom	[144, 145]
miR-30e-3p	-1,32	14,84E-5	Biomarker, Invasivität/ div. Krebserkrankungen	[17, 118, 136]

Tabelle 3.3.: Putativ neue Transkripte, basierend auf der differentiellen Expression der mRNAs Protein-codierender Gene und microRNAs, die in einem direkten Zusammenhang mit dem hochrisikoreichen Neuroblastom stehen könnten, aber deren Funktion im Neuroblastom aktuell noch nicht publiziert ist. Gezeigt sind die Transkripte mit ihren Fold Changes, den zugehörigen FDR-Werten und der bisher bekannten biologischen Assoziationen und/oder deren Vorkommen. Alle hier gelisteten Transkripte sind als putativ neu (blau) in Abb. 3.13 und 3.14 markiert und wurden aus den 20 signifikantesten hoch- bzw. runterregulierten DE-mRNAs und DE-microRNAs ermittelt. Um die Dimension der Tabelle gering zu halten, wurden die Namen der microRNAs in verkürzter Schreibweise, ohne des Präfix „hsa-“, dargestellt.

Die Runterregulation der microRNA *hsa-miR-148a-3p* und *-5p* ist mit der Metastasierung im hepatozellulären Karzinom und dem Pankreaskarzinom assoziiert. Sie kennzeichnet eine schlechte Prognose und ist aufgrund der suppressiven Wirkung als prognostischer Marker für Krebspatienten geeignet [125, 160].

Wie erwähnt, würde eine ebenso detaillierte Betrachtung der weiteren Vergleiche den Rahmen dieser Arbeit überschreiten. Das risikoreiche und aggressive Neuroblastom, mit der darin auftretenden *MYCN* Amplifikation, nimmt aus klinischer Sicht für Diagnose und Therapie eine besondere Stellung ein. Die Überlebenszeit fällt gering und die Prognose äußerst schlecht für die Patienten aus. Trotz zahlreicher wissenschaftlicher Befunde können dank moderner Techniken, wie z. B. Hochdurchsatzsequenzierungen, immer noch neue Erkenntnisse gewonnen werden. Die Daten der differentiellen Expression der mRNA- als auch der microRNA-Sequenzierung für den hier exemplarisch gewählten Vergleich 1 – *MYCN* zeigen valide Ergebnisse für bereits bekannte Transkripte der Literatur. Darüber hinaus können anhand der Daten mögliche noch unbekannte Transkripte identifiziert werden, deren Untersuchung zur Wirkungsweise und Relevanz im risikoreichen und aggressiven Neuroblastom Ansätze für weiterführende Analysen bietet. Zusammengenommen bilden die Daten der differentiellen Expression eine valide Grundlage für die nachfolgenden Untersuchungen in dieser Arbeit. Die ermittelten Veränderungen der mRNAs und microRNAs sind hierbei unter anderem für die Gewinnung der spezifischen Gensignaturen zwischen den Stadien und der putativ neuen microRNA-Target Beziehungen von entscheidender Bedeutung.

3.2.3. Gensignaturen des Neuroblastoms

Die klinische Einteilung der Tumorstadien des Neuroblastoms erfolgt nach dem *International Neuroblastoma Staging System* (INSS), welches weitestgehend auf pathologischen Befunden, wie der Lokalisation des Tumors, beruht und neben wenigen prognostischen Markern, wie beispielsweise der *MYCN* Amplifikation, kaum molekularbiologischen Befunde mit einbezieht. Die in dieser Arbeit identifizierten Signaturen bieten die Möglichkeit, die pathologischen Befunde um molekularbiologische Aspekte, in Form von Stadien repräsentierenden Mengen an Genen, zu erweitern (vgl. Abschnitt 2.2.4). Auf der Grundlage der differentiell exprimierten mRNAs der Neuroblastomstadien wurden mithilfe der entwickelten Methodik neun Gensignaturen zwischen den Vergleichen der klinisch Stadien, bestehend aus 16 Gen-Sets, identifiziert. Die Gen-Sets der Signaturen unterteilen sich dabei in *Up-* und *Down-Sets*, welche

Name der Gensignatur	Vergleich	#Gene $Up\text{-Set}$	#Gene $Down\text{-Set}$
NB.metas	1/2/3 – 4/4S	27	15
NB.MYCN	1 – MYCN	1188	466
NB.s4	1 – 4	1103	182
NB.s4M ⁺	1 – 4 ^{+M}	979	526
NB.s4M ⁻	1 – 4 ^{-M}	1961	120
NB.s4S	1 – 4S	20	(1)
NB.s4S_4	4 – 4S	88	13
NB.s4S_4M ⁺	4 ^{+M} – 4S	522	153
NB.s4S_4M ⁻	4 ^{-M} – 4S	25	(4)

Tabelle 3.4.: Übersicht über die identifizierten Gensignaturen und deren Namen. Die Up- und Down-Sets der Gensignaturen enthalten verschiedene Anzahlen hoch- und runterregulierte Gene, die eine signifikante Veränderung der Expression aus den Vergleichen zwischen den klinischen Stadien aufweisen. Die zwei eingeklammerten Down-Sets wurden aufgrund der geringen Genanzahl verworfen.

durch die signifikanten hochregulierten bzw. herunterregulierten Genen der Signaturen der Stadienvergleiche repräsentiert werden ($|\log\text{FC}| \geq 1$, $\text{FDR} \leq 0,05$, vgl. 2.2.4). Die Anzahl an Genen der jeweiligen Up- und Down-Sets sind in Tab. 3.4 dargestellt. Aufgrund ihrer geringen Anzahl an Genen, die unterhalb des gesetzten Grenze von mindestens zehn Genen pro Gen-Set liegen (vgl. 2.2.4), wurden die beiden Down-Sets der Signaturen NB.s4S und NB.s4S_4M⁻ verworfen, so dass diese Signaturen einzig aus einem Up-Set bestehen. Anhand der neun Vergleiche zwischen den Stadien (vgl. 2.2.3) ergeben sich die bereits erwähnten 16 Gen-Sets. Die Namen der Gensignaturen tragen das Präfix NB, stellvertretend für Neuroblastom und einer Benennung, die aus den Stadienvergleichen resultiert. Eine detaillierte Beschreibung der hier erzeugten Signaturen kann aus Tab. 3.5 entnommen werden. Einen Überblick darüber, wie ähnlich sich die in den Gen-Sets enthaltenen Gene hinsichtlich ihres Vorkommens in den weiteren Gen-Sets sind, ist in Abbildung 3.15 dargestellt. Die Abbildung fasst die prozentualen Anteile der Übereinstimmung der Gene innerhalb der einzelnen Gen-Sets der Signaturen zusammen. Die gezeigte Heatmap gibt für jede Zeile an, zu wie viel Prozent die Signaturgene des Up- oder Down-Sets in den Gen-Sets der Spalte vorkommen. Die Farbskala von Blau (0 %) über Weiß (50 %) nach Rot (100 %) kennzeichnet den Anteil der Übereinstimmung bezogen auf die Gesamtzahl der Signaturgene im Up- oder Down-Set der Zeile. Die Darstellung der Übereinstimmungen erfolgt hierbei anhand eines asymmetrischen Maßes, um die prozentualen Anteile der übereinstimmenden Gene bezogen auf die Anzahl der Genen im betrachteten Gen-Set herauszustellen. Da die Gen-Sets der Signaturen in ihrer Genanzahl stark variieren (vgl. Tab. 3.4), würde ein symmetrisches Maß die Deutung der Übereinstimmung zwischen

Name der Gensignatur	Beschreibung
NB metas	Signatur für stark metastasierende Stadien
NB_MYCN	Signatur für <i>MYCN</i> Amplifikation gegen Stadium 1
NB_s4	Signatur für Stadium 4 gegen 1
NB_s4M ⁺	Signatur für Stadium 4 mit <i>MYCN</i> Ampl. gegen Stadium 1
NB_s4M ⁻	Signatur für Stadium 4 ohne <i>MYCN</i> Ampl. Stadium 1
NB_s4S	Signatur für Stadium 4S gegen Stadium 1
NB_s4S_4	Signatur für Stadium 4S gegen Stadium 4
NB_s4S_4M ^{+/-}	Signatur für Stadium 4S gegen Stadium 4 mit/ohne <i>MYCN</i> Ampl.

Tabelle 3.5.: Beschreibung der zugrunde liegenden Vergleiche in den hier erstellten Gensignaturen. Die Namen der Gensignaturen entsprechen dem Vergleich zwischen den klinischen Stadien, wobei vorwiegend der Vergleich zum Stadium 1 erfolgte. Weitere Vergleiche sind in den Namen berücksichtigt und ergeben sich aus der Beschreibung der Signatur.

unterschiedlich großen Gen-Sets erschweren und nicht eindeutig abbilden. Die anhand des hier gewählten Maßes maximal ermittelte Übereinstimmung beträgt 95 %, wobei die 100 %-ige Überlappung der Gen-Sets mit sich selbst unberücksichtigt bleibt. Eine Übereinstimmung von mehr als 90 % ist anhand der mit einem Kreuz markierten Gen-Sets der Signaturen gekennzeichneten. Hohe Übereinstimmung sind zum einen darauf zurückzuführen, dass einige Proben in mehreren der Vergleiche der klinischen Stadien vorkommen. So sind beispielsweise acht Proben aus dem Vergleich des Up-Sets der NB_s4M⁺ Signatur in den zwölf Proben des Up-Sets der NB_MYCN Signatur enthalten. Beide Vergleiche sind dennoch relevant, um die Veränderungen zum Stadium 1 mit und ohne Abhängigkeit einer *MYCN* Amplifikation zu betrachten. Beide Up-Sets der Signaturen NB_s4M⁺ und NB_MYCN weisen mit 979 respektive 1188 enthaltenen Genen eine vergleichbar große Anzahl an Genen auf. Weitere Abgleiche, die zur Betrachtung der Abhängigkeit einer *MYCN* Amplifikation dienen, sind z. B. NB_s4S_4_DN mit NB_s4S_s4M⁺_DN, NB_s4S_4_UP mit NB_s4S_s4M⁺_UP, NB_s4M⁻_DN mit NB_s4_DN, NB_s4M⁺_UP mit NB_s4_UP oder NB_s4_UP mit NB_MYCN_UP.

Zum anderen können hohe Übereinstimmung auf eine geringe Anzahl an Genen in einem der betrachteten Sets zurückgeführt werden, was anhand des verwendeten asymmetrischen Vergleiches für das Set mit deutlich weniger Genen gezeigt werden kann. Dies betrifft die Vergleiche der 13 Gene des Down-Sets NB_s4S_4_DN mit den Up-Sets NB_MYCN_UP, NB_s4_UP und NB_s4M⁺_UP, die 1188, 1103 und 979 Gene enthalten und in einer Übereinstimmung von 92% der 13 Gene resultieren. Ähnliches gilt für die Übereinstimmungen zwischen dem Set NB_s4S_UP (20 Gene) mit NB_s4S_4_UP (88) und NB_s4S_4M⁺_UP (522), was zu einer

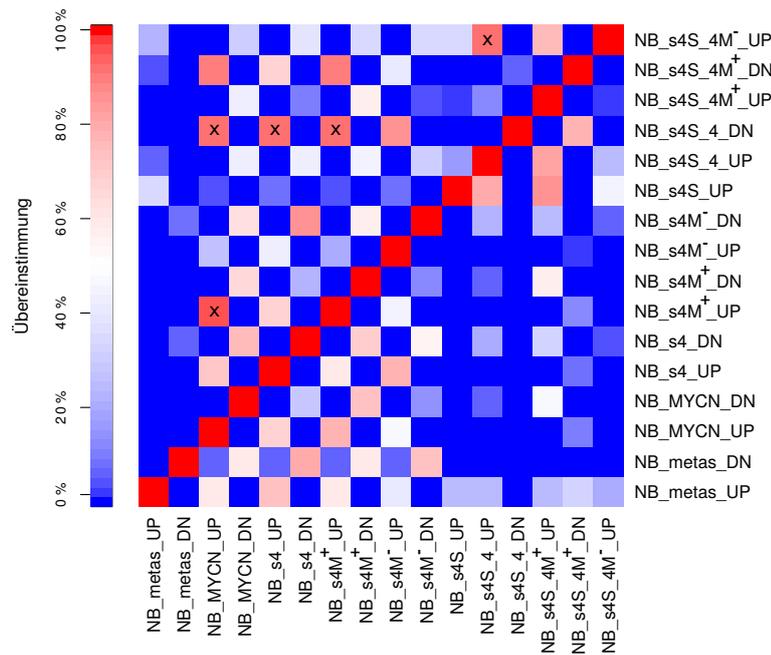


Abbildung 3.15.: Abgleich der Signaturgene in den identifizierten Signaturen des Neuroblastoms. Dargestellt ist der prozentuale Anteil an Genen die innerhalb eines Up- oder Down-Sets einer Signatur (Zeilen) mit den Genen eines anderen Sets übereinstimmen (Spalten). Die mit einem Kreuz markierten Sets weisen eine Übereinstimmung von mehr als 90 % auf.

Übereinstimmung von 80% bzw. 85% führt. Ein weiterer Vergleich zwischen den beiden Up-Sets NB_s4S_4M⁻_UP und NB_s4S_4_UP zeigt hingegen eine Kombination aus biologischer Ähnlichkeit der Proben aus Stadium 4 und einer kleinen Anzahl an Genen, was dazu führt, dass dieser Vergleich mit 25 bzw. 88 Genen in einer Übereinstimmung von 92% der 25 Gene mit dem Set NB_s4S_4_UP resultiert.

Übereinstimmungen, die sich nicht allein aus der biologischen Ähnlichkeit der Proben oder aus den Unterschieden in der Genanzahl ergeben, können Hinweise auf molekularbiologische Unterschiede zwischen den untersuchten Vergleichen der pathologischen Stadien des Tumors geben. Hierzu zeigt beispielsweise der Vergleich zwischen den Up-Sets NB_s4M⁺_UP und NB_s4M⁻_UP, dass beim Abgleich der Gene aus Vergleichen mit und ohne *MYCN* Amplifikation 45% der Gene übereinstimmen. Dies zeigt, dass beim Vergleich zwischen Stadium 1 und 4 diverse Gene unabhängig von einer Amplifikation des *MYCN* Gens in ihrer Expression verändert sind. Diese Gene stellen somit charakteristisch Kandidaten im Vergleich zwischen diesen beiden Stadien dar und können aus molekularbiologischer Sicht als Marker für Stadium 4 Tumore dienen. Die Betrachtung der Signatur NB_metas des Vergleiches der stark mit den weniger stark metastasierenden Stadien zeigt Übereinstimmungen zur Signatur NB_s4 aus dem Vergleich des

Stadium 4 zum Stadium 1. Auch wenn die Sets der Signatur NB metas nur wenige Gene enthalten (27 im Up- und 15 im Down-Set), weist dies auf Veränderungen von Genen im Stadium 4 hin, die mit der Metastasierung in Verbindung stehen, wie es für das stärker metastasierende Stadium 4 im Vergleich zum Stadium 1 zu erwarten wäre. Zudem scheinen nicht alle dieser Gene MYCN abhängig zu sein, da die Up- und Down-Sets ohne Betrachtung einer MYCN Amplifikation eine höhere Übereinstimmungen erzielen (Up 74% bzw. Down 80%), verglichen zu den Up- und Down-Sets der Signaturen NB_s4M⁺ (59%, 60%) und NB_s4M⁻ (41%, 73%), welche die MYCN Amplifikation getrennt betrachten. Eine hohe Übereinstimmung enthaltener Gene zeigt sich des Weiteren im Vergleich des Sets NB_s4S_4M⁺_DN mit den Sets NB_MYCN_UP (90%) und NB_s4M⁺_UP (89%). Hier wird bei Betrachtung von MYCN amplifizierten Tumoren deutlich, dass im Vergleich zum Stadium 1 und 4S ähnliche Gene einer Runterregulierung unterliegen, wie bereits anhand der differentiellen Expressionsanalyse deutlich wurde, deren Ergebnisse nur geringe Veränderungen im Vergleich 1 – 4S der Stadien zeigte (siehe 3.2.2).

Größere Übereinstimmungen, wie sie anhand der Heatmap in Abb. 3.15 deutlich werden, resultieren somit vorwiegend aus der biologischen Ähnlichkeiten der Proben untereinander oder aus Unterschieden in der Genanzahl, zu Ungunsten einer großen Übereinstimmung der wenigen Gene des kleineren Gen-Sets. Jedoch können auch Gen-Sets mit einer geringen Anzahl an Genen auf biologisch relevante Unterschiede zwischen den untersuchten Stadien hinweisen, wie es beispielsweise anhand des beschriebenen Vergleiches der Signatur NB metas und NB_s4 deutlich wird. Der überwiegende Teil der Gen-Sets der in dieser Arbeit identifizierten Signaturen zeigt jedoch eine geringe Überlappung der Gene untereinander und die Gen-Sets weisen daher molekulare Charakteristiken auf, um die klinischen Stadien des Neuroblastoms anhand von differentiell exprimierten Genen zu vergleichen und die Stadien zu identifizieren. Eine funktionale Analyse der enthaltenen Gene hinsichtlich ihrer Assoziationen zu bekannten biologischen Prozessen erfolgt im nachfolgenden Abschnitt der Gensignaturanalyse. Eine Übersicht über die einzelnen Gene aus den ermittelten Signatur kann aus dem Verweis im Anhang D.1 entnommen werden.

Gensignaturanalyse

Untersuchte Tumorproben und Zelllinien des Neuroblastoms werden in aktuellen wissenschaftlichen Publikationen (siehe z. B. PubMed⁴⁷) überwiegend anhand ihrer Risikogruppen und bezüglich molekularbiologischer Veränderungen in ihrer Expression oder anhand von Risikofaktoren eingeteilt und analysiert, wie beispielsweise *MYCN* Amplifikation, Mutationen oder Chromosomenabweichung in den Tumorproben (vgl. Abschnitt 1.2). Eine Einteilung einzig nach dem *International Neuroblastoma Staging System* (INSS) findet in der Analyse der Tumorerkrankung im Hinblick auf molekularbiologische Veränderungen keine Berücksichtigung und wird, so vorhanden, nur als Teil in neuere Einteilungssysteme mit einbezogen (vgl. 1.2). Das INSS beruht auf pathologischen Befunden, wie z. B. dem Patientenalter, der Lokalisation und Verbreitung der Tumore in Lymphknoten und anderen Organen (Lunge, Leber, Nieren etc.) und chirurgischen Befunden (z. B. Größe und Ausbreitung des Tumors), was eine schnelle Diagnose ermöglicht, jedoch in der Vergangenheit zu nicht eindeutigen Befunden oder Fehldiagnosen führte [169]. Eine Einteilung nach dem INSS fand nach seiner Einführung in den 90er Jahren oftmals Anwendung und wurde weltweit übernommen, auch heute findet trotz neuer und erweiterter Klassifizierungssysteme Einteilungen nach diesem System statt (vgl. 1.2). Für diese Art der Einteilung ist eine Erweiterung dank neuer Techniken wie Sequenzierungen, PCR-Methoden etc. auch über die pathologischen Befunde hinaus möglich. Eine Betrachtung der hier gewonnen Gensignaturen aus den Vergleichen der klinischen Stadien im Hinblick auf bekannte und eventuell unbekannte charakteristische Unterschiede im Transkriptom der analysierten Tumorproben kann interessante neue Erkenntnisse liefern. Dazu gehören z. B. Assoziationen zu anderen Erkrankungen oder Veränderungen in bekannten Signalwege.

Für das Neuroblastom und damit assoziierter biologischer Prozesse gibt es aktuell keine vorhandenen öffentlichen Gen-Sets. Ziel der Gensignaturanalyse ist es, die enthaltenen Gene der Signaturen hinsichtlich biologischer Assoziationen aus anderen Krebserkrankungen zu testen und mittels der öffentlich verfügbaren Gen-Sets dieser Erkrankungen unter Verwendung der GSEA Methode zu analysieren. Mithilfe der Gensignaturanalyse wurden die in den Signaturen enthaltenen Gene untereinander bezüglich allgemeiner biologischer Zusammenhänge, wie zelluläre Prozesse, Signal- bzw. Stoffwechselwege, Prozesse in Zellkompartimenten oder

⁴⁷<https://www.ncbi.nlm.nih.gov/pubmed/?term=neuroblastoma+classification>,
PubMed - Datenbank der *National Library of Medicine (USA)*

hinsichtlich spezieller Prozesse in unterschiedlichen Krebserkrankungen, wie beispielsweise Differenzierung, Metastasierung, Invasivität, Überlebenswahrscheinlichkeit etc. untersucht. Die Analyse der Signaturen erfolgte dafür mithilfe der GSEA Methode unter Verwendung öffentlich verfügbaren Gen-Sets der Kategorien C2, C5 und C6, die aus wissenschaftlichen Veröffentlichungen, Experimenten oder computergestützten Vorhersagen stammen (vgl. Abschnitt 2.2.4). Für die in dieser Arbeit ermittelten Gensignaturen des Neuroblastoms wurden dazu jeweils unabhängige Analysen der in den Signaturen enthaltenen Gene durchgeführt. Hierzu wurden die Gene des Up- und Down-Sets einer Signatur zu einer Rangliste, der nach ihrem logFC geordneten Werten aus der differentiellen Expressionsanalyse, zusammengeführt. Die Ranglisten der hier identifizierten Signaturen wurden anschließend mit den öffentlich verfügbaren Gen-Sets der drei Kategorien (C2, C5, C6) analysiert (vgl. 2.2.4). Eine vollständige Übersicht der Ergebnisse der Analyse kann aus dem Verweis im Anhang D.2 entnommen werden. Um Unterschiede zu verschiedenen biologischen Zusammenhängen zwischen den einzelnen Signaturen zu benennen, zeigen die Auffistungen in den Tabellen 3.6 und 3.7 bis zu zehn der am signifikantesten ($FDR \leq 0,05$) überrepräsentierten Gen-Sets aus der Analyse. Die erste Tabelle gibt die Gen-Sets an, die durch die hochregulierten Gene (Up-Set) der vergleichsbasierten Signatur zwischen den klinischen Stadien überrepräsentiert werden, während die zweite Tabelle die Gen-Sets darstellt, die anhand der herunterregulierten Gene (Down-Set) der Signatur überrepräsentiert werden (vgl. hierzu 2.2.4). Für die identifizierten Signaturen wurden hierfür die resultierenden Namen der öffentlich verfügbaren Gen-Sets, die Anzahl der übereinstimmenden Gene der Ranglisten mit dem Gen-Sets, die ES-, NES- und FDR-Werte extrahiert (vgl. 2.2.4) sowie eine fortlaufende Nummer vergeben. Der aufsummierte Enrichment Score (ES) entlang der Rangliste der Signaturgene gibt an, zu welchem Grad eines der öffentlichen Gen-Sets am Anfang oder am Ende der Rangliste überrepräsentiert ist und die Signaturgene der Rangliste der biologische Annotation des betrachteten Gen-Sets entsprechen. Der normalisierte Enrichment Score (NES) bezieht die Größe des Gen-Sets mit ein und lässt somit einen Vergleich der Gen-Sets untereinander zu. Der angegebene FDR-Wert (False Discovery Rate) gibt das korrigierte Signifikanzniveau wieder. Die öffentlich verfügbaren Gen-Sets der Ergebnistabelle wurden dabei nach ihren ES-Werten geordnet, um generelle Zusammenhänge anhand der am stärksten überrepräsentierten Gen-Sets zu betrachten, unabhängig von der Anzahl der repräsentierenden Gene. Für die Signaturen NB_metas, NB_s4S und NB_s4S_4M⁻ konnten keine oder nur wenige Gen-Sets ermittelt werden. Grund hierfür ist

die geringe Anzahl an Signaturgenen (vgl. Tab. 3.4), die keine ausreichende Übereinstimmung mit den derzeit öffentlich verfügbaren Gen-Sets ermöglichen. Durch Herabsetzen der minimalen Größe von zehn auf drei Gene konnte für die Signatur NB_s4S dennoch keine signifikanten Gen-Sets ermittelt werden. Für die Signatur NB_s4S_4M⁻ konnte nach Herabsetzen der Größe keine Gen-Set ermittelt werden, das anhand der herunterregulierten Gene der Signatur überrepräsentiert wird.

Das bereits erwähnte Ziel der Gensignaturanalyse ist es, die enthaltenen Gene der Signaturen hinsichtlich biologischer Assoziationen in Verbindung mit anderen Erkrankungen zu analysieren, da es keine Annotation öffentlich verfügbarer Gen-Sets gibt, die relevante biologische Prozesse im Neuroblastom beschreiben und damit assoziierte Gene zusammenfassen. Die signifikanten Gen-Sets der Tabelle zeigen Assoziationen zu diversen biologischen Zusammenhängen, die aus den Beschreibungen und Publikationen dieser öffentlichen Gen-Sets hervorgehen, beispielsweise zu unterschiedlichen Krebserkrankungen, wie dem Lungen-, Mamma- und Schilddrüsenkarzinom als auch zu Krebs im Kindesalter, oder verschiedenen Prozessen, wie der Transkription, Proliferation, Differenzierung, Apoptose oder Invasivität und Metastasierung. Zählt man die Assoziationen der gelisteten Gen-Sets der einzelnen Signaturen in den Tabellen zeigt sich bei dieser Betrachtung, dass die veränderten Gene aus den Signaturen der Vergleiche aggressiver und risikoreicher Tumore, NB_MYCN, NB_s4 und NB_s4M^{+/-}, deutlich mehr Zusammenhänge zur Entstehung und Entwicklung von Krebserkrankungen aufweisen, während die Gene der Signaturen der Vergleiche des risikoärmeren Stadiums 4S, wie NB_s4S, NB_s4S_4 und NB_s4S_4M^{+/-}, hingegen weniger dieser Zusammenhänge als auch eine an sich geringere Anzahl an signifikanten Gen-Sets zeigen. Die Signaturen aus den Vergleichen der aggressiven Tumorstadien mit weniger aggressiven und risikoärmeren Stadien 1 oder 4S zeigen anhand der Gensignaturanalyse eine größere Anzahl von signifikant überrepräsentierten Gen-Sets, sowohl hoch- als auch herunterregulierter Gene der Signaturen, die mit der Erhöhung der Metastasierung/Invasivität (z. B. Set-Nr. 80, 81, 82, 84, 116), der Transkription (12, 15, 100), des Zellzyklus (13, 17, 18, 36) oder des Zellwachstums (88, 93, 96) assoziiert sind, die Überlebenswahrscheinlichkeit (14, 78) negativ beeinflussen oder bereits im Zusammenhang mit dem Neuroblastom (78) und anderen Tumorerkrankungen (3, 5, 14, 26, 28, 45, 89) stehen. Die gelisteten Gen-Sets, die durch die Signaturgene aus dem Vergleich der aggressiven zu den risikoärmeren Stadien überrepräsentiert werden, weisen somit zahlreiche Assoziationen zu verschiedenen Krebserkrankungen auf und zu Prozessen, die die Tumorgenese sowie die

Gen-Sets überrepräsentiert durch hochregulierte Gene						
Name Gensignatur	Set-Nr.	Name öffentliches Gen-Set	#Gene	ES	NES	FDR
NB_metas	1	GOZGIT ESRI TARGETS UP	3	0,97	2,11	1,65E-002
	2	JAEGER METASTASIS DN	7	0,94	2,95	0,00E+000
	3	INAMURA LUNG CANCER SCC UP	3	0,92	2,07	1,66E-002
	4	TURASHVILI BREAST DUCTAL CARCINOMA VS DUCTAL NORMAL DN	3	0,92	1,99	2,60E-002
	5	LIU PROSTATE CANCER DN	3	0,92	2	2,68E-002
	6	ORGAN DEVELOPMENT	4	0,89	2,13	1,59E-002
	7	EPIDERMIS DEVELOPMENT	4	0,89	2,14	1,75E-002
	8	SYSTEM DEVELOPMENT	4	0,89	2,16	1,79E-002
	9	MULTICELLULAR ORGANISMAL DEVELOPMENT	4	0,89	2,18	1,93E-002
	10	ECTODERM DEVELOPMENT	4	0,89	2,14	1,96E-002
NB_MYCN	11	REACTOME AMYLOIDS	46	0,77	6,03	0,00E+000
	12	REACTOME RNA POL I PROMOTER OPENING	46	0,77	5,97	0,00E+000
	13	REACTOME PACKAGING OF TELOMERE ENDS	34	0,74	5,12	0,00E+000
	14	MONTERO THYROID CANCER POOR SURVIVAL UP	10	0,68	2,6	2,96E-004
	15	REACTOME RNA POL I TRANSCRIPTION	51	0,67	5,64	0,00E+000
	16	KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	61	0,67	6,13	0,00E+000
	17	REACTOME MEIOTIC SYNAPSIS	37	0,66	4,76	0,00E+000
	18	REACTOME MEIOTIC RECOMBINATION	55	0,65	5,58	0,00E+000
	19	FOURNIER ACINAR DEVELOPMENT LATE DN	11	0,64	2,57	3,59E-004
	20	FARMER BREAST CANCER CLUSTER 2	22	0,64	3,58	0,00E+000
NB_s4	21	REACTOME AMYLOIDS	48	0,62	4,97	0,00E+000
	22	REACTOME RNA POL I PROMOTER OPENING	47	0,62	4,88	0,00E+000
	23	MIKKELSEN MEF HCP WITH H3 UNMETHYLATED	11	0,61	2,48	1,16E-003
	24	REACTOME PACKAGING OF TELOMERE ENDS	34	0,61	4,14	0,00E+000
	25	KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	58	0,59	5,27	0,00E+000
	26	ACEVEDO LIVER CANCER WITH H3K27ME3 DN	32	0,57	3,81	0,00E+000
	27	SMID BREAST CANCER LUMINAL A DN	10	0,56	2,17	6,38E-003
	28	FARMER BREAST CANCER CLUSTER 2	21	0,56	3,15	1,75E-005
	29	NAKAYAMA SOFT TISSUE TUMORS PCA2 UP	33	0,55	3,62	0,00E+000
	30	MONTERO THYROID CANCER POOR SURVIVAL UP	10	0,55	2,08	9,82E-003
NB_s4M ⁺	31	REACTOME RNA POL I PROMOTER OPENING	45	0,75	5,89	0,00E+000
	32	REACTOME AMYLOIDS	45	0,75	5,82	0,00E+000
	33	REACTOME PACKAGING OF TELOMERE ENDS	33	0,74	5,04	0,00E+000
	34	REACTOME MEIOTIC RECOMBINATION	50	0,69	5,61	0,00E+000
	35	REACTOME MEIOTIC SYNAPSIS	36	0,66	4,68	0,00E+000
	36	REACTOME MEIOSIS	52	0,65	5,47	0,00E+000
	37	KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	60	0,65	5,82	0,00E+000
	38	LEE EARLY T LYMPHOCYTE UP	51	0,64	5,46	0,00E+000
	39	ZHAN MULTIPLE MYELOMA PR UP	31	0,63	4,22	0,00E+000
	40	REACTOME RNA POL I TRANSCRIPTION	51	0,63	5,31	0,00E+000
NB_s4M ⁻	41	KEGG OLFACTORY TRANSDUCTION	33	0,62	4,21	0,00E+000
	42	REACTOME OLFACTORY SIGNALING PATHWAY	31	0,58	3,81	0,00E+000
	43	KUNINGER IGF1 VS PDGFB TARGETS UP	11	0,53	2,09	3,11E-002
	44	SMID BREAST CANCER RELAPSE IN BONE UP	14	0,5	2,26	1,43E-002
	45	RICKMAN HEAD AND NECK CANCER C	14	0,49	2,19	2,02E-002
	46	LIEN BREAST CARCINOMA METAPLASTIC VS DUCTAL DN	13	0,47	2,02	4,23E-002
	47	MIKKELSEN IPS WITH HCP H3K27ME3	13	0,46	2,02	4,32E-002
	48	KONDO PROSTATE CANCER WITH H3K27ME3	18	0,41	2,14	2,71E-002
	49	SUBSTRATE SPECIFIC CHANNEL ACTIVITY	18	0,41	2,08	3,19E-002
	50	ION CHANNEL ACTIVITY	17	0,40	2,01	4,25E-002
NB_s4S	-	-	-	-	-	-
NB_s4S_4	51	JAEGER METASTASIS DN	18	0,71	3,18	0,00E+000
	52	SMID BREAST CANCER LUMINAL B DN	14	0,64	2,71	0,00E+000
	53	SMID BREAST CANCER BASAL UP	13	0,59	2,47	8,32E-004
	54	ONDER CDH1 TARGETS 2 DN	14	0,53	2,25	5,61E-003
	55	CHARAFE BREAST CANCER LUMINAL VS BASAL DN	11	0,53	2,03	2,00E-002
	56	SWEET LUNG CANCER KRAS DN	12	0,51	2,04	2,31E-002
NB_s4S_4M ⁺	57	HUPER BREAST BASAL VS LUMINAL UP	12	0,81	3,4	0,00E+000
	58	TURASHVILI BREAST LOBULAR CARCINOMA VS LOBULAR NORMAL UP	10	0,79	3,04	0,00E+000
	59	ECTODERM DEVELOPMENT	17	0,78	3,7	0,00E+000
	60	SENGUPTA NASOPHARYNGEAL CARCINOMA WITH LMP1 DN	10	0,76	2,91	3,18E-004
	61	EPIDERMIS DEVELOPMENT	15	0,75	3,45	0,00E+000
	62	TISSUE DEVELOPMENT	18	0,74	3,77	0,00E+000
	63	RICKMAN TUMOR DIFFERENTIATED WELL VS MODERATELY DN	12	0,73	3,1	0,00E+000
	64	TURASHVILI BREAST LOBULAR CARCINOMA VS DUCTAL NORMAL DN	10	0,72	2,84	5,22E-004
	65	STRUCTURAL CONSTITUENT OF CYTOSKELETON	10	0,67	2,55	2,72E-003
	66	PID A6B1 A6B4 INTEGRIN PATHWAY	10	0,57	2,25	1,49E-002
NB_s4S_4M ⁻	67	MURAKAMI UV RESPONSE 1HR UP	4	0,92	2,21	5,10E-003
	68	MURAKAMI UV RESPONSE 24HR	5	0,92	2,42	6,30E-004
	69	TISSUE DEVELOPMENT	5	0,92	2,36	1,57E-003
	70	ORGAN DEVELOPMENT	5	0,92	2,37	1,69E-003
	71	EPIDERMIS DEVELOPMENT	5	0,92	2,32	1,98E-003
	72	ECTODERM DEVELOPMENT	5	0,92	2,31	2,49E-003
	73	JAEGER METASTASIS DN	10	0,9	3,01	0,00E+000
	74	ONDER CDH1 TARGETS 2 DN	5	0,83	2,13	8,27E-003
	75	MULTICELLULAR ORGANISMAL DEVELOPMENT	7	0,73	2,17	6,16E-003
	76	SYSTEM DEVELOPMENT	7	0,73	2,15	6,85E-003

Tabelle 3.6.: Überrepräsentierte öffentliche Gen-Sets durch hochregulierte Gene der identifizierten Gensignaturen. Aufgelistet sind bis zu zehn der am stärksten überrepräsentierten öffentlich verfügbaren Gen-Sets der Kategorien C2, C5 und C6. Die enthaltenen Gene der öffentlichen Gen-Sets werden durch hochregulierte Gene der Signatur, welche am Anfang der Rangliste stehen, überrepräsentiert.

Gen-Sets überrepräsentiert durch herunterregulierte Gene						
Name Gensignatur	Set-Nr.	Name öffentliches Gen-Set	#Gene	ES	NES	FDR
NB_metas	77	REGULATION OF BIOLOGICAL QUALITY	4	-0,92	-2,25	1,80E-002
NB_MYCN	78	ASGHARZADEH NEUROBLASTOMA POOR SURVIVAL DN	10	-0,81	-3,17	0,00E+000
	79	LIU VAV3 PROSTATE CARCINOGENESIS UP	11	-0,81	-3,34	0,00E+000
	80	ONDER CDH1 SIGNALING VIA CTNNB1	12	-0,79	-3,25	0,00E+000
	81	KEGG CELL ADHESION MOLECULES CAMS	12	-0,76	-3,12	0,00E+000
	82	RICKMAN METASTASIS DN	14	-0,74	-3,3	0,00E+000
	83	KIM GLIS2 TARGETS UP	10	-0,74	-2,86	9,53E-005
	84	RODRIGUES NTN1 TARGETS DN	11	-0,74	-2,95	5,05E-005
	85	BASSO CD40 SIGNALING UP	10	-0,74	-2,84	1,01E-004
	86	GARGALOVIC RESPONSE TO OXIDIZED PHOSPHOLIPIDS TURQUOISE UP	11	-0,73	-2,9	8,66E-005
	87	DEMAGALHAES AGING UP	12	-0,72	-2,96	5,19E-005
NB_s4	88	AMIT EGF RESPONSE 60 HELA	13	-0,92	-3,99	0,00E+000
	89	VERHAACK GLIOBLASTOMA MESENCHYMAL	12	-0,9	-3,77	0,00E+000
	90	LEE NEURAL CREST STEM CELL UP	10	-0,89	-3,3	0,00E+000
	91	BOQUEST STEM CELL UP	14	-0,86	-3,85	0,00E+000
	92	PASINI SUZ12 TARGETS DN	16	-0,84	-3,96	0,00E+000
	93	NAGASHIMA EGF SIGNALING UP	12	-0,83	-3,43	0,00E+000
	94	SERVITJA ISLET HNF1A TARGETS UP	11	-0,81	-3,27	0,00E+000
	95	BURTON ADIPOGENESIS PEAK AT 2HR	12	-0,81	-3,43	0,00E+000
	96	TGFB UP,V1 UP	13	-0,81	-3,53	0,00E+000
	97	MANALO HYPOXIA UP	12	-0,8	-3,32	0,00E+000
NB_s4M⁺	98	SEKI INFLAMMATORY RESPONSE LPS UP	11	-0,78	-3,15	0,00E+000
	99	KEGG CELL ADHESION MOLECULES CAMS	11	-0,76	-3,06	1,26E-005
	100	LIU SMARCA4 TARGETS	11	-0,73	-2,92	2,08E-005
	101	KEGG HYPERTROPHIC CARDIOMYOPATHY HCM	11	-0,71	-2,87	2,94E-005
	102	NABA SECRETED FACTORS	15	-0,71	-3,29	0,00E+000
	103	TAKEDA TARGETS OF NUP98 HOXA9 FUSION 10D DN	18	-0,71	-3,57	0,00E+000
	104	LIU VAV3 PROSTATE CARCINOGENESIS UP	14	-0,71	-3,17	0,00E+000
	105	NADLER OBESITY UP	10	-0,7	-2,68	2,58E-004
	106	NAKAYAMA SOFT TISSUE TUMORS PCA1 UP	10	-0,68	-2,63	3,71E-004
	107	LIAN LIPA TARGETS 6M	10	-0,68	-2,61	4,14E-004
NB_s4M⁻	108	AMIT EGF RESPONSE 60 HELA	10	-0,96	-3,72	0,00E+000
	109	REN ALVEOLAR RHABDOMYOSARCOMA DN	13	-0,78	-3,35	0,00E+000
	110	BURTON ADIPOGENESIS PEAK AT 2HR	10	-0,77	-2,99	4,06E-005
	111	BURTON ADIPOGENESIS 1	10	-0,76	-2,98	5,68E-005
	112	CHEN LVAD SUPPORT OF FAILING HEART UP	12	-0,71	-3,02	2,15E-005
	113	NAKAMURA ADIPOGENESIS EARLY UP	10	-0,7	-2,75	1,62E-004
	114	ZHU CMV ALL DN	11	-0,7	-2,94	7,01E-005
	115	BOQUEST STEM CELL UP	15	-0,69	-3,16	0,00E+000
	116	SCHUETZ BREAST CANCER DUCTAL INVASIVE UP	17	-0,66	-3,28	0,00E+000
	117	PHONG TNF RESPONSE VIA P38 PARTIscheinbarAL	13	-0,65	-2,82	1,24E-004
NB_s4S	-	-	-	-	-	-
NB_s4S_4	118	MARSON BOUND BY E2F4 UNSTIMULATED	11	-0,88	-3,37	0,00E+000
	119	DUTERTRE ESTRADIOL RESPONSE 24HR UP	10	-0,86	-3,21	0,00E+000
	120	KINSEY TARGETS OF EWSR1 FLII FUSION UP	11	-0,84	-3,24	0,00E+000
	121	GOBERT OLIGODENDROCYTE DIFFERENTIATION UP	11	-0,77	-3,01	0,00E+000
	122	VECCHI GASTRIC CANCER EARLY UP	10	-0,75	-2,75	0,00E+000
	123	LASTOWSKA NEUROBLASTOMA COPY NUMBER UP	10	-0,61	-2,29	1,52E-003
	124	NUYTEN EZH2 TARGETS DN	12	-0,61	-2,42	8,80E-004
	125	DODD NASOPHARYNGEAL CARCINOMA DN	17	-0,55	-2,51	3,76E-004
NB_s4S_4M⁺	126	RHODES UNDIFFERENTIATED CANCER	10	-0,87	-3,41	0,00E+000
	127	GRAHAM NORMAL QUIESCENT VS NORMAL DIVIDING DN	11	-0,86	-3,43	0,00E+000
	128	KOBAYASHI EGFR SIGNALING 24HR DN	36	-0,86	-6,04	0,00E+000
	129	WINNEPENINCKX MELANOMA METASTASIS UP	15	-0,85	-3,89	0,00E+000
	130	FERREIRA EWINGS SARCOMA UNSTABLE VS STABLE UP	13	-0,85	-3,72	0,00E+000
	131	LEE EARLY T LYMPHOCYTE UP	15	-0,85	-4,01	0,00E+000
	132	ZHAN MULTIPLE MYELOMA PR UP	10	-0,85	-3,28	0,00E+000
	133	ZHOU CELL CYCLE GENES IN IR RESPONSE 6HR	12	-0,84	-3,48	0,00E+000
	134	CELL CYCLE PROCESS	15	-0,84	-3,9	0,00E+000
	135	WHITEFORD PEDIATRIC CANCER MARKERS	21	-0,84	-4,48	0,00E+000
NB_s4S_4M⁻	-	-	-	-	-	-

Tabelle 3.7.: Überrepräsentierte öffentliche Gen-Sets durch herunterregulierte Gene der identifizierten Gensignaturen. Aufgelistet sind bis zu zehn der am stärksten überrepräsentierten öffentlich verfügbaren Gen-Sets der Kategorien C2, C5 und C6. Die enthaltenen Gene der öffentlichen Gen-Sets werden durch herunterregulierte Gene der Signatur, welche am Ende der Rangliste stehen, überrepräsentiert.

Verbreitung des Tumors begünstigen und somit die Überlebenschancen verringern und die Prognose verschlechtern (vgl. [84]).

Betrachtet man hingegen die Signaturen aus den Vergleichen des weniger aggressiven Stadiums 4S mit den aggressiven Tumoren, NB_s4S_4, NB_s4S_4M⁺, NB_s4S_4M⁻, so werden vermehrt Gen-Sets überrepräsentiert, deren annotierte Prozesse eine Verminderung der Aggressivität erkennen lassen oder sogar die Tumorentwicklung hemmen oder daran beteiligte Prozesse umkehren bzw. abschwächen können. Dazu gehören Gen-Sets, welche herunterregulierte Gene während der Metastasierung/Invasivität beinhalten (z. B. Set-Nr. 51, 54) und die durch eine Hochregulierung dieser Gene im Vergleich 4 – 4S auf eine Verringerung der Metastasierung bzw. Invasivität im Stadium 4S schließen lassen. Dem entgegen zeigen ebenso herunterregulierte Gene der Signatur NB_s4S_4M⁺ Assoziationen zum Gen-Set, dessen enthaltenen Gene als hochreguliert im Zusammenhang mit der Metastasierung annotiert sind (129). Beides weist somit auf eine verminderte Metastasierung bzw. Invasivität im Stadium 4S hin und kann zur bekannten Verringerung der Aggressivität in diesem Stadium führen. Zu weiteren ermittelten Gen-Sets im Vergleich des Stadium 4S zu aggressiveren Stadien des Neuroblastoms zählen beispielsweise Prozesse der Transkription und des Zellzyklus (118, 120, 128, 133). Aber auch Prozesse, die mit Chromosomenabweichung im Neuroblastom (123), Mutationen (56) oder mit Krebs im Kindesalter (135) einhergehen. Des Weiteren werden diverse Prozesse, die mit der Entwicklung von Geweben oder dem Zytoskelett, der Zell-Zell-Kommunikation oder Zelladhäsion assoziiert sind durch hochregulierte Gene im Stadium 4S überrepräsentiert (59, 61, 62, 65, 66, 70). Diese genannten Zusammenhänge deuten eine Abnahme der Aggressivität im Stadium 4S an, indem typische Prozesse der Tumorgenese sowie der Verbreitung gehemmt werden, während Prozesse wie z. B. die Zell-Zell-Kommunikation oder Adhäsion durch eine Hochregulierung beteiligter Gene induziert werden. Besonders erwähnenswert an dieser Stelle ist die Überrepräsentierung von Gen-Sets im weniger aggressiven Stadium 4S, die im Zusammenhang mit der Differenzierung stehen (63, 121, 126) und auf einer Erhöhung des Differenzierungspotentials in den weniger aggressiven Tumoren hindeuten könnten. Dies könnte Ansätze für verbesserte Diagnosen und Therapien offen legen. Bekanntlich nimmt die Ausdifferenzierung von Tumorzellen im Neuroblastom eine wesentliche Stellung ein und ist Ziel diverser Therapieansätze, was nachweislich zu einer besseren Prognose und erhöhten Überlebenschancen führt [38,266]. So könnten die hier ermittelten annotierte Prozesse Hinweise auf Transkripte geben, die mit der Differenzierung im Zusammenhang stehen und

deren Analyse in weiterführenden Experimenten neue Möglichkeiten zur Therapie aggressiver Tumorstadien geben könnten.

Neben bekannten zu erwartenden Prozessen repräsentieren die identifizierten Signaturgene weitere annotierte Gen-Sets, die auf noch unbekannte Mechanismen, Signalwege oder Prozesse hindeuten könnten und die womöglich mit dem Neuroblastom assoziiert sind, wie beispielsweise Kanalaktivitäten (49, 50) oder Entzündungsreaktionen (98). Bestimmte Assoziationen sind dabei besonders hervorzuheben, da ihr Bezug zur Erkrankung nicht direkt ersichtlich wird. Das spezielle Neuroblastom 4S zeigt dabei innerhalb der Signatur NB_s4S_4M⁻ eine Hochregulierung von Genen, die annotierten Sets mit Verbindungen zur Reaktion bzgl. UV Bestrahlung (67, 68) überrepräsentieren. Dabei ist bekannt, dass UV-Strahlung Hautkrebs und Immunstörungen begünstigt und es konnte im Zusammenhang mit dem Neuroblastom anhand einer Zelllinie gezeigt werden, dass differenziertere Zellen eine höhere Sensitivität für UV-Strahlung mit Abnahme von DNA-Reparaturprozessen erkennen lassen. Wohingegen weniger differenzierte Zellen eine Zunahme an DNA-Schädigungen aufweisen [206]. Weiterhin konnte in der Literatur kein aktueller Zusammenhang zwischen dem Neuroblastom und der UV-Reaktion ermittelt werden. Die ermittelten Gen-Sets zeigen mit vier bzw. fünf übereinstimmenden Genen lediglich eine sehr geringe Anzahl, dennoch könnten diese Sets als Ansatz für weiterführende Experimente, hinsichtlich der Relevanz von UV-Reaktionen im Neuroblastom, dienen.

Für Signaturen aus den Vergleichen des aggressiven Tumorstadiums 4 lassen sich anhand der Gensignaturanalyse vermehrt Zusammenhänge zu Gen-Sets der Methylierung (23, 26, 47, 48) sowie auffällige Assoziationen zu Sets der Adipositas und Diabetes (94, 95, 105, 111, 113) ermitteln. Die Hochregulierung von Genen der Methylierung, die im wesentlichen mit einer Modifikation des Histons 3 assoziiert sind oder gar im Speziellen durch eine Tri-Methylierung gekennzeichnet sind, weisen auf einen Einfluss der Methylierung im Neuroblastom hin. Dabei handelt es sich um eine Form der Histonmodifikation, welche mit der Regulation der Zelldifferenzierung und der Genexpression verknüpft ist. Störungen oder Veränderungen der Methylierung führen dabei zur Beeinflussung der Genexpression durch Einflussnahme auf die Transkription und stellen einen wesentlichen Grund für die Heterogenität, Entstehung und Entwicklung von Tumoren dar [42, 63]. Die aus den Gen-Sets ermittelte Form der Methylierung (me) bezieht sich auf die Aminosäure Lysin (K) des Histon 3 (H3), wobei es sich hier um den Grad einer Tri-Methylierung handelt (me₃), welche als H3K27me₃ bezeichnete wird [42]. Durch Methylierung bzw. Demethylierung kann sich die Struktur des Chromatins in einen

aktiven oder inaktiven Zustand verändern, was zur genannten Beeinflussung der Transkription führen kann. Im Speziellen ist die H3K27me3 mit der Inaktivierung des Chromatins und der Inhibierung der Genexpression betroffener Gene assoziiert [42, 71]. Diese und weitere Modifikationen können genutzt werden, um Aussagen über die Prognose verschiedener Krebserkrankungen zu ermöglichen. Allgemeinen tritt dabei in Krebserkrankungen eine Veränderung der Histonmethylierungsmuster auf [42]. Aus der Gensignaturanalyse gehen hierbei unterschiedliche Überrepräsentierungen durch hoch- oder herunterregulierte Gene von methylierten bzw. nicht methylierten Gen-Sets hervor, die keinen eindeutigen Trend bezüglich einer hoch- oder herunterregulierten Expression von (nicht) methylierten Genen erkennen lässt. Deutlich wird jedoch, dass Signaturen der aggressiven zu weniger aggressiven Stadienvergleiche vermehrt Bezüge zur Methylierung und insbesondere der H3K27me3 aufzeigen. Dabei scheint dieser Befund unabhängig vom Vorhandensein einer *MYCN* Amplifikation aufzutreten, da derartige Gen-Sets sowohl in der gemischten Signatur NB_s4 als auch in der Signatur NB_s4M⁻ ohne Amplifikation enthalten sind. Aus der Literatur ist dazu bekannt, dass Methylierungen die Expression der Wnt und Dkk Genfamilien induziert, welche einen Einfluss auf die Differenzierung von Zellen und Geweben haben [186, 260, 262]. Was wiederum einen möglichen Einfluss auf das bekannte reduzierte Differenzierungspotential in den aggressiven Stadien des Neuroblastoms haben könnte.

Neben diesen - nicht gänzlich unbekanntem - Befunden war auffällig, dass in den durch herunterregulierten Genen überrepräsentierten Gen-Sets der Stadium 4 Signaturen mehrfach solche ermittelt werden konnten, die mit der Adipogenese oder der Diabetes assoziiert sind. Dieser Bezug zum Neuroblastom scheint zunächst ungewöhnlich, jedoch haben weiterführende Recherchen der Literatur ergeben, dass Methylierungen über den genannten Wnt-Signalweg einen direkten Einfluss auf die Differenzierung von Präadipozyten zu Adipozyten hat und somit die Generierung von Fettgewebe, der sogenannten Adipogenese, reguliert [71]. Dabei stellen die damit verbundenen Krankheiten, wie Adipositas (Fettsucht) und die Lipodystrophie (Fettverteilungsstörung), wichtige Risikofaktoren für den Typ-2-Diabetes dar [71]. Aktuell lassen sich in der Literatur hierzu vermehrt Zusammenhänge zur Regulierung der Adipogenese erkennen, welche auf der Modifikation von Histonen beruht und mithilfe der Methylierung den Wnt-Signalweg beeinflussen. Zudem ist beschrieben, dass Gene der Wnt Familie die Adipogenese hemmen können, indem eine verstärkte Expression zur Aktivierung des Signalweges führt [71, 186, 262]. Durch die Veränderungen der Methylierung im Stadium 4 könnte es hier zu

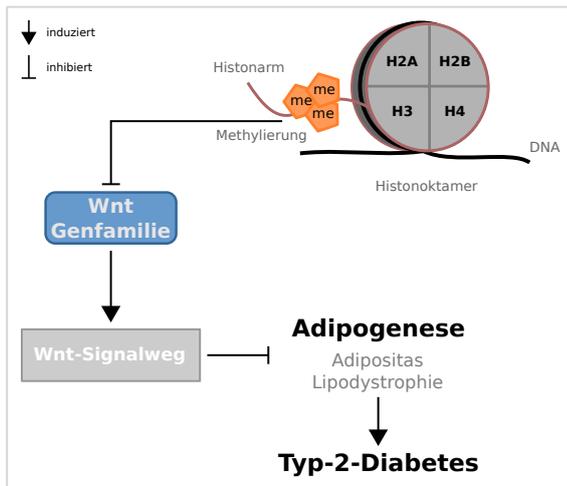


Abbildung 3.16.: Regulation des Wnt-Signalweges durch Histonmodifikationen, mit Auswirkungen auf die Adipogenese und damit einhergehender Regulierung der Typ-2-Diabetes. Befunde der Gensignaturanalyse deuten Zusammenhänge des aggressiven Neuroblastoms mit der Regulierung der Adipogenese, durch eine Trimethylierung des Lysins an Position 27 des Histon 3, an. Die damit verbundenen Krankheiten, Adipositas und Lipodystrophie, weisen einen nachgewiesenen starken Assoziation zur Diabetes auf.

einer Regulierung von Genen der Wnt Familie kommen, deren Produkte den Wnt-Signalweg aktivieren oder deaktivieren könnten. Eine Aktivierung des Wnt-Signalweges durch eine erhöhte Expression der Genfamilie könnte folglich zu einer Inhibierung der Adipogenese führen, bedingt durch eine - Wnt-Signalweg vermittelte - Störung der Präadipozytendifferenzierung. Abbildung 3.16 zeigt dazu die schematische Darstellung der beschriebenen Regulierung mit Einfluss auf den Typ-2-Diabetes. An dieser Stelle soll kein direkter Bezug zwischen dem Neuroblastom und der Adipogenese und/oder der Diabetes postuliert werden. Vielmehr soll der Fokus darauf gerichtet sein, dass sich anhand der Gen-Sets der Methylierung und Adipogenese Hinweise auf regulierte Gene andeuten, welche am Prozess der Differenzierung beteiligt sein könnten, da wie bereits erwähnt eine Behandlung der unterdrückten Differenzierung im aggressiven und risikoreichen Neuroblastom in zahlreichen Therapien zielführend ist.

Aus den drei ermittelten Gen-Sets der Gensignaturanalyse, die direkt mit der Adipogenese assoziiert sind (vgl. Set-Nr. 95, 111, 113 in Tab. 3.7), konnten verschiedene Gene ermittelt werden, die in mindestens zwei der am stärksten überrepräsentierten Sets vorhanden sind. Zu diesen gehören beispielsweise die Gene *ZFP36*⁴⁸, *CEBPD*⁴⁹ sowie Mitglieder der Nuklearrezeptorfamilie *NR4A*⁵⁰. Die hier im Zusammenhang mit der Adipogenese genannten Gene sind nachweislich auch an der Differenzierung diverser Zelltypen beteiligt, wie beispielsweise neuronale Zellen oder Keratinozyten [6, 18, 47, 48, 198, 204, 208], darüber hinaus ist für einige dieser Gene bereits bekannt, dass diese durch Methylierungen [6, 113, 231] und/oder den Wnt-Signalweg [10] reguliert werden. Für die CEBP Proteine sind bereits Beziehungen zur

⁴⁸ *ZFP36 ring finger protein*, auch als TTP bezeichnet.

⁴⁹ *CCAAT/enhancer binding protein delta*.

⁵⁰ *NR4A1* und *NR4A2* (*nuclear receptor subfamily 4 group A member*).

Differenzierung im Neuroblastom durch weitere Mitglieder der Familie nachgewiesen [47]. Die aus der Gensignaturanalyse resultierenden Zusammenhänge, der Methylierung und Adipogenese in den aggressiven Tumoren, lassen sich somit anhand verschiedener Publikationen bestätigen. Die ermittelten Gene der Adipogenese deuten dabei auf neue Zusammenhänge bezüglich der Methylierung und Präadipozytendifferenzierung hin, vermittelt durch den Wnt-Signalweg. Damit verbunden könnten dies neue Möglichkeiten für die Differenzierung im aggressiven Neuroblastom aufzeigen. Zudem lässt sich hier ein wiederkehrender Bezug zum beschriebenen Wnt-Signalweg erkennen, der bereits in der differentiellen Expression hinsichtlich hochregulierter Wnt Genen auffällig war (vgl. Abschnitt 3.2.2) und ebenso in der Analyse der microRNA-Targets sichtbar wird (vgl. Abschnitt 3.2.4). Alle Befunde bezüglich des Wnt-Signalweges in dieser Arbeit deuten auf eine erhöhte Expression der Wnt Genfamilie und somit auf eine verstärkte Aktivität des Signalweges in den aggressiven und risikoreicheren Stadien des Neuroblastoms hin. Ein bisher nicht vollends untersuchter Zusammenhang des Signalweges mit dem Neuroblastom deutet sich dabei ebenfalls in einigen Publikationen an, wie beispielsweise [65, 137]. Neben den hier bereits genannten Genen, könnten weitere der identifizierten Gen-Sets neue Transkripte in Erscheinung bringen, die als therapeutische Targets im Neuroblastom von Belang sein könnten. Gerade im Bezug auf die Differenzierung aggressiver Tumore könnten die Befunde zur Regulationen durch den Wnt-Signalweg und/oder der Methylierung vielversprechende Ansätze für weitere Analysen bieten. Genaue Untersuchungen, zur biologische Relevanz dieser ermittelten Regulierungen aus den Sets der Methylierung und Adipogenese, sollten anhand von weiterführenden Experimenten und Analysen stattfinden. Erkenntnisse aus diesen Zusammenhängen könnten womöglich neue Targets zur Regulierung von Differenzierungseffekte identifizieren und bieten damit neue Möglichkeiten für die Diagnose, den Verlauf oder die Therapie der aggressiven Stadien.

Insgesamt betrachtet zeigt sich anhand der Ergebnisse der Gensignaturanalyse, dass eine Kombination der pathologischen und molekularbiologischen Merkmale durch die ermittelten Signaturen sinnvoll ist und darüber hinaus der Vergleich mit öffentlichen annotierten Gen-Sets neue Hinweise, beispielsweise zu anderen Erkrankungen oder auf veränderte Signalwege, anhand transkriptomischer Unterschiede hervorbringen kann. Die in dieser Arbeit entwickelte Methode sollte sich dabei auf jedweder Art von Tumorerkrankung übertragen lassen, um daraus neue Erkenntnisse für die Diagnose und/oder Therapie der Erkrankung zu gewinnen, so sich in irgendeiner Weise die differentielle Expression von Transkripten zwischen den Tumorentitäten bestimmen lässt.

Validierung der identifizierten Gensignaturen

Um die in dieser Arbeit identifizierten stadienbasierten Gensignaturen des Neuroblastoms zu validieren, wurden öffentlich verfügbaren Daten primärer Neuroblastomproben verwendet. Da aktuell keine umfangreichen stadienbasierten Signaturen existieren und somit kein direkter Vergleich vorgenommen werden kann, ist das Ziel der Validierung die ermittelten Signaturen mithilfe von Ranglisten aus den öffentlichen Daten zu analysieren. Die Analyse erfolgt dabei mittels GSEA Methode (vgl. Abschnitt 2.2.4) auf den Ergebnissen der differentiellen Expression der öffentlichen Neuroblastomdaten. Es soll somit überprüft werden, ob die hier ermittelten Signaturen anhand der entwickelten Identifizierungsmethode der stadienbasierten Vergleichen durch die differentielle Expression identischer Vergleiche aus öffentlicher Daten repräsentiert werden können und wie verlässlich die Veränderungen in den öffentlichen Daten durch die Signaturen charakterisiert werden.

Bei den zur Validierung verwendeten Daten handelt es sich um zwei öffentliche Datensätze, welche über die *R2 Genomics Analysis and Visualization Platform*⁵¹ des *Academic Medical Center* (AMC) in Amsterdam zugänglich sind. Einer der Datensätze umfasst 498 primären Tumorproben [236] (GSE62564)⁵² und basiert auf einer RNA-Sequenzierung unter Verwendung eines Illumina HiSeq 2000 Geräts. Dieser Datensatz wurde somit mittels einer vergleichbaren Technik zu den hier analysierten Daten generiert und wird im Folgenden als „R2-Seq.“ bezeichnet. Die Daten liegen als RPM (*Reads Per Million*) normalisierte Counts vor. Um die Daten mittels der entwickelten Methode - identisch zu den in dieser Arbeit analysierten Daten - auswerten zu können, wurden diese denormalisiert, anschließend TMM normalisiert und hinsichtlich differentieller Expression analysiert (vgl. hierzu 2.2.3). Die notwendigen Librarygrößen der Rohdaten für die Denormalisierung wurden dafür beim Autor von [236] erfragt. Der zweite Datensatz umfasst 88 primäre Tumorproben, deren Genexpression mithilfe eines Microarrays (Affymetrix Human Genome U133 Plus 2.0 Array, MAS5.0 Normalisierung) analysiert wurde [167] (GSE16476)⁵³. Dieser Datensatz wird in der vorliegenden Arbeit folgend als „R2-Array“ bezeichnet. Da sich die Methodik zur Bestimmung differentieller exprimierter Gene aus dieser Arbeit (vgl. 2.2.3) nicht für die Analyse von Microarraydaten eignet, erfolgte der Test auf differentielle Expression mithilfe der R2-Plattform unter Anwendung des einfaktoriellen

⁵¹<http://r2.amc.nl>

⁵²<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62564>

⁵³<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16476>

	s1	s2	s3	s4	s4S	MYCN
RNA-Seq.	14	5	8 (4)	25 (8)	9	12
R2-Seq.	121 (13)	78 (5)	63 (15)	183 (65)	53 (4)	92
R2-Array	8	15	13 (1)	40 (15)	12	16

Tabelle 3.8.: Verteilung der Probenanzahl auf die Tumorstadien. Übersicht über die 61 primären Neuroblastomproben des analysierten Datensatzes (RNA-Seq.) und die öffentlich verfügbaren Daten der „R2 Genomics Analysis and Visualization Platform“ mit 498 Proben, die mittels RNA-Sequenzierung (R2-Seq.) und 88 Proben, welche durch eine Microarray (R2-Array) analysiert wurden. Die eingeklammerten Werte geben die Verteilung der MYCN amplifizierten Proben auf die Stadien der Datensätze an. In der letzten Spalte ist die Gesamtzahl an Proben angegeben, die eine MYCN Amplifikation aufweisen.

Anova-Tests. Zur Übersicht stellt Tab. 3.8 die Verteilung der Probenanzahl, der jeweiligen Neuroblastomstadien der analysierten und der öffentlichen Daten dar. Eine Übersicht über die veränderten Expressionen der einzelnen Gene zwischen den Stadien kann aus dem Verweis im Anhang E entnommen werden.

Auf Grundlage dieser öffentlichen Datensätze wurden Ranglisten, basierend auf den logFC-Werten aller Gene aus der Analyse der differentiellen Expression dieser Daten, generiert (siehe Verweis zu den Daten im Anhang E). Diese Ranglisten der öffentlichen Daten wurden anschließend mittels GSEA Methode analysiert. Hierzu wurden sowohl öffentliche verfügbaren kuratierte Gen-Sets der Kategorie C2 verwendet (vgl. 2.2.4), die annotierte und experimentell validierte Mengen von Genen zu diversen biologischen Prozessen, wie Differenzierung, Wachstum, Migration, Invasivität etc. enthalten, als auch die Gen-Sets der in dieser Arbeit identifizierten Signaturen des RNA-Seq. Um sowohl Gen-Sets der Signaturen mit geringer als auch großer Anzahl an Genen zu erlauben, wurden für die Validierung die GSEA Parameter angepasst. Dazu wurden die Standardwerte für die minimale und die maximale Größe der Sets von 15/500 auf 10/2000 - entsprechend der Genanzahl in den Gen-Sets der identifizierten Signaturen - geändert (vgl. Tab. 3.4). Somit erfolgt für die Gen-Sets der hier identifizierten Signaturen keine Einschränkung der Analyse, in der Form, dass dabei Gen-Sets aufgrund einer sehr hohen Anzahl von übereinstimmenden Genen mit den Ranglisten herausgefiltert werden. Bei einer Unterschreitung von mindestens zehn übereinstimmenden Genen wurden das jeweilige Gen-Sets verworfen (vgl. hierzu 2.2.4). In den beiden Tabellen 3.9 und 3.10 sind die Ergebnisse der Validierung mittels GSEA Methode zusammengefasst. Diese Position der Gen-Sets der Signaturen aus den Ergebnissen der GSEA soll den Zusammenhang der identifizierten Signaturen zum Neuroblastom deutlich machen und zeigen, dass die Ranglisten

aus den neuroblastombezogenen öffentlichen Daten diese Signaturen überrepräsentieren. Dabei müssen die Gen-Sets der identifizierten Signaturen nicht an erster Stelle stehen, vielmehr ist es wichtig, dass diese vordere Positionen erreichen, da die analysierten Ranglisten aus den gleichen stadienbasierten Vergleichen des Neuroblastoms stammen und somit eine biologisch ähnliche Datengrundlage zu den Signaturen bilden. Die Analyse mittels GSEA soll den Zusammenhang der Gene der Rangliste zum Neuroblastom und zu den Vergleichen aus den Stadien erkennen lassen. Die Analyse soll zeigen, dass die hier ermittelten Signaturen in erster Linie durch Ranglisten der öffentlichen Neuroblastomdaten überrepräsentiert werden anstelle anderer Gen-Sets der Kategorie C2, die keinen direkten Zusammenhang zum Neuroblastom aufweisen. In den Tabellen sind die erzielten Positionen der Gen-Sets (Zeilen) aus der GSEA für die einzelnen Ranglisten (Spalten) der öffentlichen Daten gezeigt. Dabei sind die Positionen der Gen-Sets der hier identifizierten Signaturen, die durch hochregulierte Gene der Ranglisten überrepräsentiert werden in Tab. 3.9, Gen-Sets die durch herunterregulierte Gene der Ranglisten überrepräsentiert werden in Tab. 3.10 dargestellt. Die Positionen sind für beide Datensätze, den R2-Seq. und R2-Array, gemeinsam aufgelistet. Ein fehlender Eintrag in einem oder beiden Datensätzen kennzeichnet, dass das Gen-Sets nicht in den Ergebnissen der Analyse für die untersuchte Rangliste enthalten ist. Die Kennzeichnung „NA“ markiert die fehlenden Down-Sets der Signaturen NB_s4S und NB_s4S_4M⁻, die aufgrund ihrer geringen Genanzahl verworfen wurden (vgl. hierzu Tab. 3.4). Die vollständigen Ergebnisse der Analyse mittels GSEA können aus dem Verweis im Anhang E entnommen werden.

Bei unabhängiger Betrachtung der Regulierungsrichtung der Gene in den Ranglisten zeigen die Diagonalen beider Tabellen, dass die Gen-Sets der Signaturen NB metas, die Up-Sets der Signaturen NB_s4S und NB_s4S_4M⁻ und das Down-Set der Signatur NB_s4S_4 hintere Plätze bei den Positionen belegen. Die Gen-Sets dieser Signaturen werden weniger durch die Ranglisten der öffentlichen Daten überrepräsentiert. Das bedeutet, dass die Gene im Gen-Set nicht am Anfang oder am Ende Rangliste auftreten sondern sich entlang der Rangliste verteilen (vgl. hierzu 2.2.4) und daher keine Plätze unter den ersten Positionen in der Ergebnisliste der GSEA erreichen. Andere Gen-Sets, der in dieser Arbeit identifizierten Signaturen oder aus den öffentlichen Daten der Kategorie C2, erzielen somit besser Ergebnisse bei der Analyse dieser Ranglisten. Auffällig hierbei ist, dass die weniger gut repräsentierten Gen-Sets der genannten Signaturen die geringsten Mengen an Genen aus den Stadienvergleichen beinhalten (vgl. Tab. 3.4). Die hinteren Positionen aus den Ergebnissen der GSEA deuten darauf hin, dass

Signatur	Set	Position der Ranglisten R2-Seq. / R2-Array								
		metas	MYCN	s4	s4M ⁺	s4M ⁻	s4S	s4S ₋₄	s4S _{-4M⁺}	s4S _{-4M⁻}
NB_metas	UP	294/1524	340/795	334/706	305/756	624/864	2371/634			-/1337
	Down						1167/-	1275/438	847/839	
NB_MYCN	UP	1/1	1/1	1/1	1/1	3/1	2/2			
	Down						2/25	3/4	2/-	
NB_s4	UP	3/3	3/3	2/3	3/3	2/2	4/4			
	Down						5/157	5/90	4/290	
NB_s4M ⁺	UP	2/2	2/2	3/2	2/2	4/3	3/3			
	Down						1/108	1/10	3/-	
NB_s4M ⁻	UP	4/5	4/5	4/4	4/5	1/4	7/10			
	Down						11/419	51/-	9/349	
NB_s4S	UP	1036/1358	-/1196	-/901	-/1047	1212/678	920/467	1487/801	1636/822	1076/767
	Down									
NB_s4S ₋₄	UP							13/59	30/59	27/146
	Down	254/209	255/169	229/142	254/164	205/114	276/184			
NB_s4S _{-4M⁺}	UP							3/25	2/6	7/-
	Down	11/11	9/4	9/6	9/4	65/15	119/24			
NB_s4S _{-4M⁻}	UP						2181/-	157/251	335/253	145/254
	Down									

Tabelle 3.9.: Positionen der Gen-Sets der hier identifizierten Signaturen, die durch hochregulierte Gene der Ranglisten aus den öffentlichen Daten repräsentieren werden. Aufgelistet sind die erreichten Positionen aus den Ergebnissen der GSEA Methode für die Ranglisten aus dem R2-Seq. und dem R2-Array.

Signatur	Set	Position der Ranglisten R2-Seq. / R2-Array								
		metas	MYCN	s4	s4M ⁺	s4M ⁻	s4S	s4S ₋₄	s4S _{-4M⁺}	s4S _{-4M⁻}
NB_metas	UP	368/691	432/1856	282/2408	378/1953	354/933	480/115	269/2596	314/2695	551/-
	Down									
NB_MYCN	UP							1/1	1/1	3/34
	Down	1/1	3/2	1/2	3/2	1/1	6/4			-/1580
NB_s4	UP							2/3	3/3	2/11
	Down	4/6	5/12	4/11	4/11	4/11	27/30			
NB_s4M ⁺	UP							3/2	2/2	4/54
	Down	2/3	2/3	2/3	2/3	2/2	9/2			-/142
NB_s4M ⁻	UP							4/7	5/20	1/6
	Down	5/23	10/127	5/69	8/119	5/42	31/77		-/783	
NB_s4S	UP									
	Down		2485/-	2102/-	2493/-		NA/NA			
NB_s4S ₋₄	UP	7/12	15/20	7/40	17/23	7/74	314/641			
	Down							247/835	280/1041	307/612
NB_s4S _{-4M⁺}	UP	3/2	1/1	3/1	1/1	3/3	12/6			-/440
	Down							12/39	11/23	90/536
NB_s4S _{-4M⁻}	UP	67/103	272/755	97/672	387/523	73/785	-/3125			
	Down									NA/NA

Tabelle 3.10.: Positionen der Gen-Sets der hier identifizierten Signaturen, die durch herunterregulierte Gene der Ranglisten aus den öffentlichen Daten repräsentieren werden. Aufgelistet sind die erreichten Positionen aus den Ergebnissen der GSEA Methode für die Ranglisten aus dem R2-Seq. und dem R2-Array.

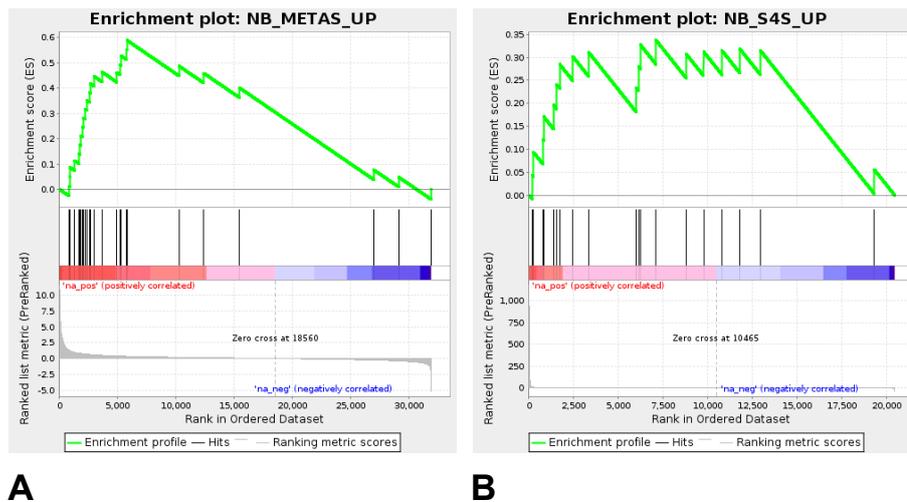


Abbildung 3.17.: Enrichment Plots aus der Analyse der Ranglisten der öffentlichen Daten. (A) Plot des Up-Sets der Signatur NB_metas aus der Analyse der R2-Seq. Rangliste aus dem Vergleich der stark zu den gering metastasierenden Neuroblastomproben. (B) Plot des Up-Sets der Signatur NB_s4S aus der Analyse der R2-Array Rangliste aus dem Vergleich des Stadiums 4S zu Stadium 1.

die hier identifizierten Gene dieser Signaturen nicht ausreichend oder charakteristisch genug für einen Vergleich dieser Stadien sind und die Veränderungen zwischen den Stadien nicht deutlich genug abbilden. Die Signaturen NB_metas und NB_s4S erzielen insgesamt die schlechtesten Ergebnisse. Abbildung 3.17 zeigt beispielhaft das Ergebnis der GSEA als Enrichment Plots für die Up-Sets der Signaturen NB_metas und NB_s4S aus der Analyse des R2-Seq. bzw. des R2-Array Datensatzes. Die Abbildung zeigt die geringe Übereinstimmung der Ranglistengene mit den hochregulierten Signaturgenen im Up-Set und macht deutlich, dass die wenigen übereinstimmenden Gene über die Rangliste verteilt sind und sich nicht alle am Anfang der Liste befinden. Zudem zeigen 3 bzw. 4 der Gene eine Herunterregulierung und treten somit am Ende der Rangliste auf. Die Gene der Signatur sind somit nicht eindeutig bzw. charakteristisch genug, um auch anhand anderer Daten die Gen-Sets der Signaturen so überrepräsentieren zu können, dass diese eine der vorderen Positionen in den Ergebnissen einnehmen und somit klar auf biologischen Zusammenhänge starker Metastasierung und Zugehörigkeit zum Stadium 4S des Neuroblastoms hindeuten. Eine mögliche Erklärung, warum diese Signaturen nicht charakteristisch genug sind, könnte in der Heterogenität des Vergleiches der stark metastasierenden zu den schwach bis nicht metastasierenden Stadien liegen. Bei den hier vorliegenden Daten beinhalten die Proben des Stadium 4 zum einen eine große Anzahl MYCN amplifizierter Proben. Zum anderen zeigt die Signatur NB_s4S des Vergleiches zwischen

Stadium 1 und 4S, dass diese sich basierend auf der differentiell Expressionsanalyse kaum in ihrer Genexpression unterscheiden (vgl. 3.2.2). Der Vergleich „vermischt“ somit in den stark metastasierenden Proben schwache und starke Veränderungen in der Genexpression, die nicht allein durch den Prozess der Metastasierung beeinflusst werden.

Für die spaltenweise Betrachtung der Ranglisten MYCN, s4, s4M⁺, s4M⁻ und s4S_4M⁺ befinden sich die jeweiligen fünf Signaturen NB_MYCN, NB_s4, NB_s4M⁺, NB_s4M⁻ und NB_s4S_4M⁺ auf den vorderen Positionen in den Ergebnissen der GSEA. Die vorderen Plätze zeigen deutlich den Zusammenhang der Ranglisten der öffentlichen Daten zum Neuroblastom sowie Abhängigkeiten zum Stadium 4 und einer MYCN Amplifikation auf Basis der Stadienvergleiche der Signaturen NB_MYCN, NB_s4 und NB_s4M⁺. Da sich die zugrunde liegenden Probenmengen zur Identifizierung der Signaturen überschneiden, erzielen diese Signaturen ähnlich gute Positionen. Bei Betrachtung der Übereinstimmung der Signaturgene war bereits zu erkennen, dass ein geringer Anteil der enthaltenen Gene in mehreren Signaturen vorkommt (vgl. hierzu Abb. 3.15).

Für Ranglisten, die im Zusammenhang mit einer MYCN Amplifikation im Vergleich zum Stadium 1 stehen, erzielen auch die Signatur NB_s4M⁻ ohne MYCN Amplifikation und die Signatur NB_s4M⁺ vordere Plätze bei den Positionen. Für die Signatur NB_s4M⁻ bedeutet dies, dass die Signatur neben den MYCN abhängigen Gene auch Gene enthält, die für das Stadium 4 charakteristisch sind, unabhängig von einer vorliegenden MYCN Amplifikation. Dies zeigt sich auch daran, dass, bei zeilenweiser Betrachtung, die Signatur NB_s4M⁻ für die zugehörige Rangliste s4M⁻ die vorderste Position erzielt. Die Signatur somit charakteristisch für Stadium 4 Tumore ohne eine MYCN Amplifikation ist. Für die Rangliste s4S_4M⁺ erzielt die entsprechende Signatur NB_4S_4M⁺ eine vordere Position bei den durch hochregulierten Genen überrepräsentierten Signaturen (vgl. Tab. 3.9). Die Signaturen NB_s4M⁺ und NB_MYCN des R2-Seq. bzw. des R2-Array erreichen knapp bessere Positionen. Hieraus wird deutlich, dass ähnliche Veränderungen der Gene im Vergleich des Stadiums 4 mit MYCN Amplifikation zum Stadium 4S bzw. 1 vorliegen, somit Überschneidungen in den exprimierten Genen zwischen Stadium 4S und Stadium 1 Tumoren vorhanden sind (vgl. Abb. 3.15). Dies lässt sich auch anhand der geringen differentiellen Expressionsveränderungen zwischen diesen beiden Stadien belegen (vgl. Abschnitt 3.2.2). Die Signatur NB_4S_4M⁺ beinhaltet entsprechend dieser Ergebnisse relevant hochregulierte Gene aus dem Vergleich des Stadium 4S zum Stadium 4 mit MYCN Amplifikation. Für die herunterregulierten Gene zeigt sich anhand der Rangliste

in Tabelle 3.10, dass die Signatur NB_4S_4M⁺ schlechter abschneidet als beispielsweise die Signaturen NB_MYCN oder NB_s4M⁺, da diese bessere Positionen in den Ergebnissen der GSEA erzielen. Die Signaturgene des Down-Sets der NB_4S_4M⁺ Signatur scheinen somit nicht hinreichend charakteristisch genug für den Vergleich zwischen den herunterregulierten Genen des Stadium 4S zum Stadium 4 mit einer *MYCN* Amplifikation zu sein. Anhand der Positionen des Down-Sets der Signatur aus den Ergebnissen wird deutlich, dass die Signatur jedoch einen Zusammenhang zum Neuroblastom aufzeigt und für Analysen mittels GSEA genutzt werden kann, um differentielle Veränderungen zwischen den Neuroblastomstadien zu charakterisieren. Anhand der hier durchgeführten Validierung mittels Ranglisten aus öffentlichen Expressionsdaten zeigen die in dieser Arbeit identifizierten fünf Signaturen NB_MYCN, NB_s4, NB_s4M⁺, NB_s4M⁻ und NB_s4S_4M⁺ Charakteristiken, um anhand veränderter Gene zwischen den Stadienvergleichen Zusammenhänge zur Erkrankung des Neuroblastoms und auf die zugrunde liegenden Neuroblastomstadien erkennen zu lassen. Neben den Vergleichen der Stadien kann zudem eine *MYCN* Amplifikation in die Betrachtung einbezogen werden. Für die vier Signaturen NB metas, NB_s4S und NB_s4S_4M⁻ und NB_s4S_4 zeigt sich im Vergleich zu den anderen Signaturen dieser Arbeit, aber auch zu anderen Gen-Sets der Kategorie C2, dass die enthaltenen Gene dieser Signaturen eine ungenügende Charakteristik für die Vergleiche darstellen und ein klarer und eindeutiger Bezug zum Neuroblastom und der biologischen Zusammenhänge der Vergleiche fehlt. Hier sollte in weiterführenden Arbeiten eine Optimierung im Hinblick auf eine Größere Menge an Genen als auch deren Eindeutigkeit und Relevanz im Bezug auf die Stadienvergleiche erfolgen. Dennoch macht die Validierung der identifizierten Signaturen dieser Arbeit deutlich, dass anhand molekularbiologischer Aspekte charakteristische Gene ergänzend zu der INSS basierten Stadieneinteilung genutzt werden können, um die Stadien genauer zu definieren. Dank neuester Methoden der Genexpressionsanalyse können differentielle Veränderungen in den Stadien heutzutage schnell und kostengünstig erfasst werden und zur eindeutigeren Einteilung und damit verbunden zur Verbesserung der Prognose und Einschätzung von Therapiemöglichkeiten verwendet werden.

3.2.4. microRNA-Targets im Neuroblastom

MicroRNAs spielen eine wesentliche Rolle als posttranskriptionelle Regulatoren der Genexpression. Durch die Bindung an den 3'-UTR der mRNAs ihrer Zielgene (microRNA-Targets) beeinflussen sie die Target-Expression, was zur translationellen Repression oder zum

Abbau bzw. einer Destabilisierung der Target mRNA führen kann [85]. Die microRNAs können als Tumorsuppressoren oder Onkogene wirken, indem sie mRNAs onkogen oder tumorsuppressiv agierender Gene binden und deren Expression positiv oder negativ regulieren. Dabei binden onkogene microRNAs bevorzugt tumorsuppressive mRNAs und umgekehrt, wodurch in Krebserkrankungen suppressive Prozesse inaktiviert oder unterdrückt werden können. MicroRNAs mit onkogenem Potential wirken sich somit begünstigend auf die Tumorentwicklung aus, indem sie z. B. Gene der Differenzierung, Apoptose und/oder tumorsuppressiv Gene negativ regulieren. Tumorsuppressive microRNAs hingegen inhibieren die Entwicklung des Tumors, beispielsweise durch eine negative Regulation von Onkogenen und/oder positiven Regulation von Genen, die die Zelldifferenzierung oder Apoptose steuern. Für Erkrankungen, darunter auch Krebserkrankungen, sind neben den regulierten Targets die Prozesse der Biogenese zur Entstehung der microRNAs und der microRNA vermittelte mRNA-Abbau von Interesse. Eine Deregulierung der Biogenese hat zur Folge, dass die Produktion, Stabilität und Reifung der microRNAs beeinflusst wird, was deren Expression und die Regulation ihrer Targets beeinträchtigen kann. Veränderungen in der microRNA Biogenese und damit verbundener Auswirkungen auf die posttranskriptionelle Regulation der Targets bildet aktuell die Hypothese zur Entstehung von Krebserkrankungen [81, 214, 268].

Dank des gewonnenen Materials aus ein und derselben Tumorprobe für die RNA- und microRNA-Sequenzierung, war es in der hier vorliegenden Arbeit möglich die transkriptomischen Veränderung der Expression zwischen den klinischen Stadien auf mRNA- und microRNA-Ebene miteinander zu korrelieren (vgl. Abschnitt 2.2.5). Dies ermöglicht - *in silico* - die Untersuchung der bevorzugten Bindung von onkogenen microRNAs an tumorsuppressive mRNAs und umgekehrt sowie die Identifizierung putativ neuer microRNAs, Targets und microRNA-Target Beziehungen im Neuroblastom. Durch die Materialgrundlage konnten zudem die Veränderungen in der microRNA Biogenese und Aktivität anhand der Expressionsdaten verschiedener Faktoren, die an der der microRNA Entstehung und dem microRNA vermittelten Abbau von mRNAs beteiligt sind, innerhalb der vorliegenden Arbeit untersucht werden. Die Analyse der veränderten microRNAs und ihrer möglichen Targets in dieser Arbeit (vgl. 2.2.5) stützt sich dabei auf die differentielle Expression der microRNAs und mRNAs gemäß der Vergleiche zwischen den klinischen Stadien (vgl. Abschnitt 2.2.3). Die Vorhersage der microRNA-Targets erfolgte anhand der selektierten DE-microRNAs der differentiellen Expression (vgl. 2.2.3) bei einem $|\logFC| \geq 1$

und einem mittleren CPM-Wert⁵⁴ von 100 (vgl. 2.2.5). Diese Werte wurden gewählt, um zum einen die Vorhersagen nur auf die microRNAs zu beziehen, die eine signifikante Veränderung zwischen den Stadien zeigen und ein gewisses Mindestmaß an Veränderung bzw. Expression zwischen und innerhalb der Proben aufweisen. Im Anschluss an die Selektion der microRNAs und die Vorhersage deren putativ direkten Targets wurden die microRNA-Target Beziehungen unter Einbeziehung der Expressionsdaten, der Ergebnisse aus den Vorhersagedatenbanken, den Untersuchungen aus den Überlebenszeitanalysen und bindungsspezifischer Informationen anhand eines Scores bewertet (vgl. Abschnitt 2.2.5). Die Bewertung der microRNA-Target Beziehungen dient der Beurteilung, wie wahrscheinlich die hier in dieser Arbeit ermittelten putativ direkten Bindungen aus den Vergleichen der Expressionsdaten zwischen den klinischen Stadien ist. Zudem soll eine Bewertung der ermittelten Beziehungen untereinander ermöglicht werden, um die wahrscheinlichsten Verbindungen sowie mögliche noch unbekannte Beziehungen zu identifizieren. Dabei liegt ein hoher Score der Bewertung vor, wenn alle aus der Targetanalyse stammenden Daten auf eine putativ direkte Bindung zwischen microRNA und Target hinweisen und eine Regulierung des Targets durch die microRNA wahrscheinlich ist (vgl. 2.2.5).

microRNA Expression

Bereits bei der Betrachtung der microRNA Expressionswerte werden Unterschiede in der Expression innerhalb und zwischen den Stadien deutlich, die auf ein onkogenes bzw. tumorsuppressives Potential diverser microRNAs in den unterschiedlich stark aggressiven und risikoreichen Stadien hindeuten. Zur Darstellung dieser Unterschiede zeigt die Heatmap in Abb. 3.18 die Medianwerte der Expression für bekannte onkogene und tumorsuppressive microRNAs bzw. microRNA Familien innerhalb der einzelnen Neuroblastomstadien. Die Mediane der Expression basieren auf den CPM-Werten der annotierten microRNA Reads der Proben der jeweiligen Stadien. Die gezeigten microRNAs sind als onkogen oder tumorsuppressiv wirkend im Zusammenhang mit dem Neuroblastom [155, 224] oder in diversen anderen Tumorerkrankungen, wie beispielsweise Hirntumoren, dem Mamma-, Pankreas- oder Schilddrüsenkarzinom, beschrieben [226, 282] und wurden hier für die Darstellung der Expressionsunterschiede aus der genannten Literatur zusammengetragen.

Onkogene microRNAs weisen in Tumoren mit schlechterer Prognose und stärkerer Aggressivität

⁵⁴Counts per Million (CPM) gibt die Anzahl der annotierten microRNA Reads, normalisiert auf eine Million Reads, an.

eine erhöhte Expression auf, um die Entwicklung des Tumors voranzutreiben. Die Abbildung zeigt, dass vor allem das aggressive und risikoreiche Neuroblastom, Stadium 4 mit *MYCN* Amplifikation (4^{+M}), eine stark erhöhte Expression bekannter onkogener microRNAs aufweist (vgl. Abb. 3.18). Als Beispiele seien hier einige der bekanntesten onkogen wirkenden microRNAs genannt. Dazu zählt die in den vorliegenden Daten am stärksten exprimierte miR-17 Familie des miR-17~92 Clusters (Chromosom 13) und seine paralogen Clustern miR-106a~363 (Chromosom X) und miR-106b~25 (Chromosom 7). Alle drei Cluster besitzen ein hohes onkogenes Potential und sind mit diversen Krankheiten, darunter auch dem Neuroblastom, assoziiert. Der *MYCN* induzierte miR-17~92 Cluster reguliert dabei eine Vielzahl diverser Targetgene, die wesentliche krebsbezogene Prozesse wie Zellproliferation, -adhäsion oder verschiedene Signalwege regulieren [46,159,165]. Eine weitere zu nennende onkogene microRNA ist die *hsa-miR-9*, die mit diversen Kreberkrankungen assoziiert ist, darunter mit dem Glioblastom und dem Neuroblastom. Die Darstellung der hier analysierten Daten zeigt, dass sie in den Stadien 2, 3 und 4^{+M} eine erhöhte Expression aufweist. Es ist beschrieben, dass die microRNA in verschiedenen Krebszellen unterschiedlich stark exprimiert wird und eine Aktivierung durch *MYCN* stattfindet. Dabei ist für das Neuroblastom bekannt, dass die Expression mit der von *MYCN* korreliert und sich ähnlich zur Expression der microRNAs des onkogenen 17~92 Clusters verhält [143,175,281]. Die hier vorliegenden Expressionsdaten bestätigen dies und zeigen eine vergleichbare Expression zum Cluster in den anderen Stadien der Abb. 3.18. Aufgrund ihrer stark erhöhten Expression in den Stadien $4^{+M/-M}$ sind die microRNAs *hsa-miR-21* und *hsa-miR-181* im Vergleich zu den anderen Stadien zu nennen. Beide microRNAs besitzen ein onkogenes Potential und zeigen in weniger aggressiven Stadien eine verringerte Expression. Innerhalb der Expressionsdaten weisen sie, wie in der Literatur beschrieben, in *MYCN* amplifizierten Tumoren sowie Tumoren mit ungünstigem Verlauf eine hochregulierte Expression auf [38,155].

Im Gegensatz zu den genannten onkogen wirkenden microRNAs zeigen bekannte tumorsuppressive microRNAs eine oftmals verringerte Expression in den aggressiven Stadien. In Stadien, die sich hingegen durch eine geringe Aggressivität und bessere Prognose auszeichnen, wie beispielsweise Stadium 1, 2 und 4S (vgl. Abschnitt 1.2), weisen tumorsuppressive microRNAs eine erhöhte Expression auf, um gezielt Wege der Zelldifferenzierung und Apoptose zu regulieren und die Weiterentwicklung des Tumors zu hemmen. Ein Beispiel für eine bekannte Familie von tumorsuppressiven microRNAs ist die *let-7* Familie. Diese ist nachweislich an Prozessen des Zellwachstums, der Differenzierung und Tumorsuppression beteiligt und nimmt in diversen

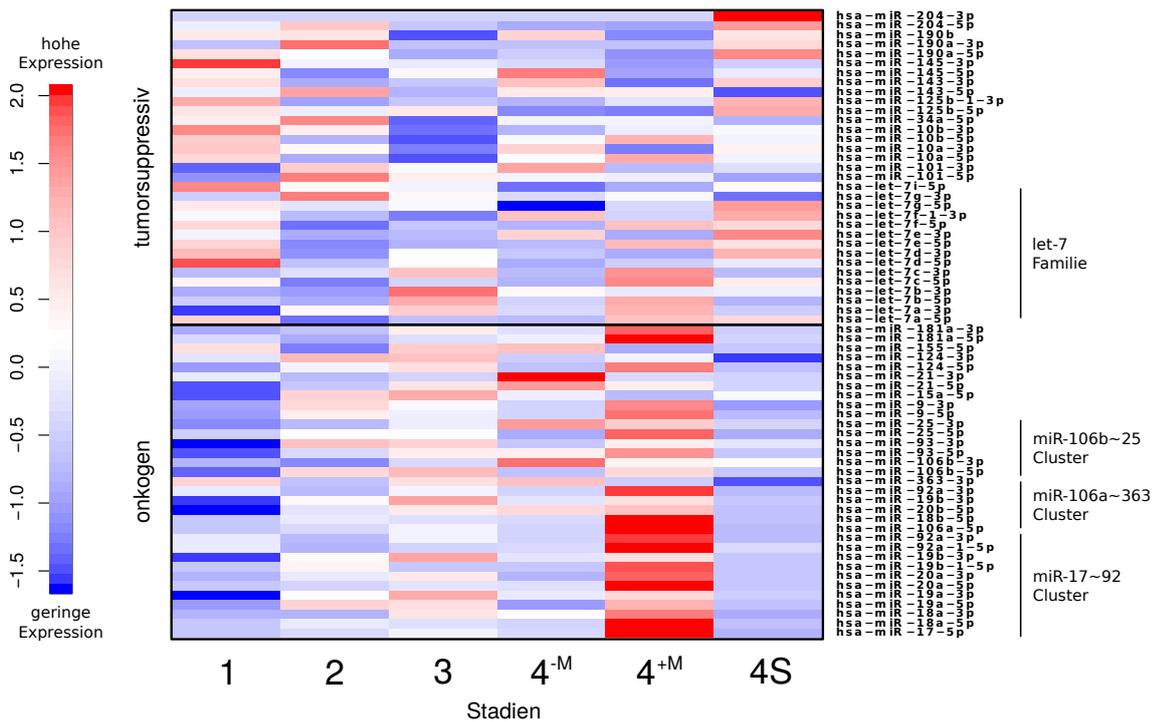


Abbildung 3.18.: Heatmap der Medianwerte der Expression bekannter onkogener und tumorsuppressiver microRNAs im Vergleich der Stadien des Neuroblastoms aus den hier analysierten Sequenzierungsdaten. Für die Darstellung der Expression fand eine zeilenweise Skalierung der Mediane statt, so dass für alle microRNAs Mittelwerte von 0 und Standardabweichungen von 1 erreicht wurde.

Krebserkrankungen eine wichtige Rolle in der Diagnose und Therapie ein [21, 197]. Darüber hinaus kommt der microRNA Familie eine besondere Bedeutung zu, da sie über zahlreiche microRNA Mitglieder verfügt, die über verschiedene Spezies hinweg konserviert sind und diverse Funktionen innerhalb sowie zwischen den Spezies erfüllen. Trotz des beschriebenen vornehmlich tumorsuppressiven Potentials der microRNA Familie, weisen die einzelnen Mitglieder unterschiedlich starke - sogar konträre - Expressionen in verschiedenen Zelltypen oder Geweben auf. In Säugetieren ist bekannt, dass eine erhöhte Expression der microRNAs in der Embryogenese, dem sich entwickelnden Gehirn und während der Differenzierung vorliegt. Im Gegensatz dazu wurde eine oftmals reduzierte Expression in diversen Krebserkrankungen festgestellt, wobei Teile der Familie sich dennoch durch eine Hochregulierung auszeichnen und die allgemeine tumorsuppressive Funktion der Familie nicht für alle Mitglieder zutreffen ist [21]. Auch die in dieser Arbeit analysierten Expressionsdaten der Neuroblastomproben lassen diese Gegensätze erkennen. Deutlich wird dies vor allem in den Stadien 1 und 4S, in denen ein Teil der Familie hochreguliert und der andere Teil runterreguliert wird, z. B. *hsa-let-7e-3p* und *hsa-let-7g-5p* respektive *hsa-let-7b-5p* und *hsa-let-7c-5p*. Im Stadium 4^{+M}

verhält sich die Expression dieser microRNAs hingegen vollständig umgekehrt. Betrachtet man zudem die Expression der microRNA Familie zwischen den Stadien 4^{-M} und 4^{+M} , so wird deutlich, dass sich die Expression aller Mitglieder, in ihrer Stärke und/oder Richtung, voneinander unterscheidet. Dabei sind die grundlegenden Funktionen und Mechanismen, durch welche sich die Mitglieder in ihrer Expressionsstärke und -richtung unterscheiden, weitestgehend noch ungeklärt. Neueste Publikationen deuten an, dass eine Deregulierung der let-7 microRNAs auf verschiedenen Wegen erfolgen kann, wie beispielsweise der Unterdrückung der microRNA Biogenese durch das RNA-bindende Protein LIN28B oder die Anwesenheit des onkogenen MYCN Proteins [21,197]. Für das Neuroblastom ist beschrieben, dass die Expression von *LIN28B* in aggressiveren Tumoren erhöht und die Abundanz im Neuroblastom *MYCN* unabhängig ist [166]. Durch die Expression von *LIN28B* wird zudem die let-7 Familie in nicht *MYCN* amplifizierten Tumoren unterdrückt, was somit die Expression von *MYCN* vorantreibt, während *LIN28B* einen geringen Einfluss in *MYCN* amplifizierten Tumoren aufweist [197]. Entsprechend der Literatur zeigen auch die Befunde aus den hier analysierten Daten, dass die let-7 microRNA Expression in den Stadien 4 mit *MYCN* Amplifikation die höchste Abundanz aufweisen. In nicht *MYCN* amplifizierten Stadium 4 zeigt die microRNA eine wesentlich geringere Expression und wird womöglich wie in der Literatur beschrieben von *LIN28B* unterdrückt. Bestätigend zur beschriebenen negativen Regulation der let-7 microRNA, durch *LIN28B* [166,197], weist *LIN28B* in den Stadien 4 (mit oder ohne *MYCN* Amplifikation) die höchste Expression im Vergleich zu den anderen Stadien auf (vgl. hierzu Tab. 3.11). Im Stadium 4^{+M} wird der Einfluss von *LIN28B* jedoch möglicherweise durch *MYCN* oder andere Faktoren verringert. Dies kann anhand der vorliegenden Expressiondaten in dieser Arbeit gezeigt werden, da im Stadium 4^{+M} im Vergleich zum Stadium 4^{-M} eine starke Zunahme der let-7 Expression unter erhöhter *MYCN* Expression und bei nahezu gleichbleibender *LIN28B* Expression vorliegt (vgl. Tab. 3.11). Darüber hinaus ist *MYCN* selbst ein Target der let-7 Familie [21,26], was zu einer negativen Regulation von *MYCN* durch microRNAs der let-7 Familie führt. Neueste Erkenntnisse in der Literatur zeigen jedoch, dass es bei einer Amplifizierung von *MYCN* zur Bindung eines Großteils der microRNAs der let-7 Familie an die *MYCN* mRNA kommen kann, wodurch die microRNAs stark in ihrer Expression unterdrückt werden können [197]. All diese genannten Mechanismen sind weitestgehend noch ungeklärt und bedürfen weiterer Untersuchungen und Nachweise.

Weiterhin fallen bei den tumorsuppressiven microRNAs in Abb. 3.18 die *hsa-miR-145* und *hsa-miR-204* anhand ihrer Expressionsunterschiede in bestimmten Stadien auf. Die

	s1	s2	s3	s4 ^{-M}	s4 ^{+M}	s4S
MYCN	48,30	78,67	619,75	64,36	1721,30	64,91
LIN28B	8,14	3,89	8,46	19,03	18,87	7,01
hsa-let-7a-5p	68987,56	32835,65	43292,77	42553,24	75093,09	67576,99

Tabelle 3.11.: Mediane der CPM-Werte aus den hier analysierten Proben der einzelnen Neuroblastomstadien für die Gene *MYCN* und *LIN28B* sowie der microRNA *hsa-let-7a-5p*, exemplarisch für die *let-7* Familie.

hsa-miR-145 zeigt eine sehr starke Expression im Stadium 1 (*hsa-miR-145-3p*) und Stadium 4^{-M} (*hsa-miR-145-5p*), während sie in den anderen Stadien wenig und im Stadium 4^{+M} kaum exprimiert wird. Die microRNA ist als tumorsuppressiv wirkend in verschiedenen Krebsarten beschrieben und hemmt die Tumorgenese durch die Bindung einer Vielzahl von Zielgenen, welche an der Regulation von Prozessen wie Proliferation, Invasion, Migration, Apoptose oder Differenzierung beteiligt sind. Eine reduzierte Expression der microRNA ist bekanntermaßen überwiegend mit aggressiven Tumoren verschiedener Erkrankungen, erhöhter Tumorgröße, Metastasen und einer schlechteren Prognose gekoppelt [49]. In den analysierten Daten bestätigt sich die geringe Expression der *hsa-miR-145* vor allem im aggressiven und risikoreichen Stadium 4^{+M}. Für die zweite genannte microRNA *hsa-miR-204* liegt im Stadium 4S der hier analysierten Daten, gegenüber allen weiteren Stadien, die höchste Expression vor. Im Median weist z. B. die *hsa-miR-204-5p* einen CPM-Wert von rund 385 auf, während beispielsweise für die Stadien 1 und 4^{+M} Werte von 160 und 117 vorliegen. Laut Literatur ist eine geringe Expression der microRNA mit einer schlechteren Prognose assoziiert und lässt der tumorsuppressiven microRNA als Biomarker bei der Diagnose eine wesentlichen Rolle zukommen, die sowohl im Neuroblastom als auch in verschiedener Krebserkrankungen angewendet wird, wie beispielsweise dem Mammakarzinom oder Gliom [28, 130, 272]. Im Neuroblastom dient die microRNA unter anderem zur unabhängigen Vorhersage der Überlebenswahrscheinlichkeit und kommt bei der Chemotherapieresistenz zum Einsatz. Eine erhöhte Expression der microRNA führt dabei zu einer gesteigerten Sensitivität auf Chemotherapeutika, was folglich zu einer verringerten Rückfallrisiko und einer erhöhten Überlebenswahrscheinlichkeit führt [216]. Auch die hier analysierten Daten zeigen eine generell niedrigere Expression der microRNA (z. B. im Stadium 3 oder 4^{+M}), wobei insbesondere die *hsa-miR-204-5p* verringerte Expressionswerte im Vergleich der aggressiven Stadien (4^{+M/-M}) zu den weiteren Stadien (1, 2 und 3) aufweist und das bekannte tumorsuppressive Potential der microRNA bestätigen [216]. Die erhöhte Expression

im Stadium 4S könnte auf eine Reduzierung der Aggressivität in den vorliegenden Proben dieses Stadiums hindeuten. Womöglich deutet sich hier aufgrund der erhöhten Expression zudem ein Ansatz für weiterführende Untersuchungen in Bezug auf eine gesteigerte Wirkungsfähigkeit von Chemotherapeutika speziell im Stadium 4S des Neuroblastoms an (vgl. hierzu [216]).

Zusammengenommen kann somit gezeigt werden, dass sich diverse bekannte microRNAs mit nachgewiesenem onkogenem bzw. tumorsuppressivem Potential anhand der Expressionsdaten aus den Analysen dieser Arbeit bestätigen lassen. Dabei können klare Unterschiede zwischen der Aggressivität und dem Risiko der einzelnen Tumorstadien dargestellt werden. Eine Betrachtung der differentiellen Expression zwischen den Stadien erlaubt jedoch die Identifizierung signifikant veränderter microRNAs, wie sie im nachfolgenden Abschnitt der Targetanalyse Anwendung findet.

Targetanalyse differentieller microRNAs

Die durchgeführte Targetanalyse der innerhalb dieser Arbeit als differentiell identifizierten microRNAs dient der Ermittlung möglicher direkter Bindungsbeziehungen zwischen den veränderten microRNAs der Stadien und den durch diese microRNAs putativ direkt regulierten Target mRNAs. Anhand der ermittelten Beziehungen sollen in diesem Abschnitt zunächst die Ergebnisse der Targetanalyse dargestellt und im Folgenden hinsichtlich publizierter sowie putativ neuer microRNA-Target Beziehungen im Neuroblastom untersucht werden. Auf Basis der differentiellen Expression (vgl. 2.2.3) erfolgte innerhalb der Targetanalyse die Selektion und Vorhersage möglicher Targets der veränderten microRNAs sowie die Korrelation und Bewertung der Target mRNAs bzw. der Beziehungen zwischen den selektierten microRNAs und deren möglichen Targets (vgl. Abschnitt 2.2.5). Die sich daraus ergebene Anzahl an ermittelten microRNAs und ihrer putativ direkten Targets sind, bezüglich der differentiellen Vergleiche zwischen den Neuroblastomstadien, in der nachfolgenden Abb. 3.19 zusammengefasst. Dabei ist die Anzahl der hoch- und runterregulierten microRNAs sowie Targets, die für eine direkte Beziehung eine entgegengesetzte differentielle Expression aufweisen sollten (vgl. 2.2.5), zusammen mit der Veränderungsrichtung aus der differentiellen Expression dargestellt. Insgesamt konnten mithilfe der Methode (vgl. 2.2.5) der hier entwickelten Targetanalyse 69 verschiedene microRNAs und 2808 verschiedene potentielle Target mRNAs identifiziert werden, woraus mehr als 30000 mögliche Bindungen resultieren. Ergänzend hierzu ist in Abb. 3.20 die Veränderungen der Expression aller ermittelten differentiellen microRNAs auf

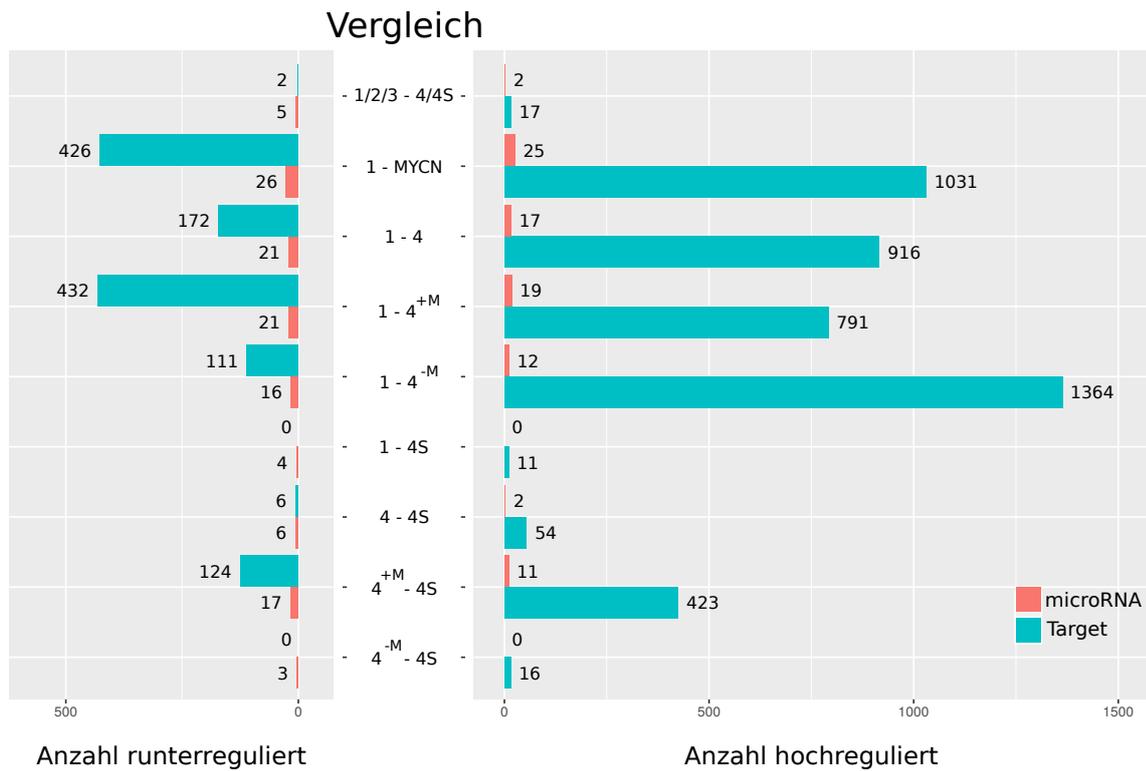


Abbildung 3.19.: Anzahl der durch die microRNA-Targetanalyse ermittelten differentiellen microRNAs und ihrer putativ direkten Targets, die eine entgegengesetzte differentielle Expression zwischen den Stadien aufweisen (entgegengesetzte Balken im Diagramm). Die orangefarbenen Balken kennzeichnen die Anzahl der microRNAs, während die türkisfarbenen Balken im Diagramm die Anzahl der ermittelten potentiellen Targets darstellen.

Basis der logFC-Werte in einer Heatmap dargestellt. Weiterhin sind in der Abbildung bereits in der Literatur beschriebene onkogene und tumorsuppressive microRNAs gekennzeichnet (vgl. vorheriger Abschnitt). Beide Abbildungen lassen erkennen, dass Vergleiche, die risikoreiche und aggressive Stadien mit risikoärmeren und weniger aggressiven Stadien umfassen, die größte Anzahl an ermittelten microRNAs und Targets aufweisen. Genauer betrachtet finden starke Veränderungen der Expression z. B. in den Stadien 1 – MYCN und 1 – 4 verglichen mit den risikoarmen bzw. weniger aggressiven Stadium 1 statt, was entsprechende Befunde aus der differentiellen Expressionsanalyse in dieser Arbeit widerspiegelt und demnach auch für die Targetanalyse eine erhöhte Anzahl an identifizierten microRNAs bzw. mRNAs zur Folge hat (vgl. Abschnitt 3.2.2). Aus den stadienbasierten Vergleichen 1/2/3 – 4/4S , 1 – 4S , 4 – 4S und 4^{-M} – 4S hingegen ließen sich keine oder nur wenige microRNAs und Targets identifizieren. Insbesondere konnten beim Vergleich 4^{-M} – 4S keine differentiell hochregulierten microRNAs entsprechend der Methode selektiert und somit keine Targets vorhergesagt werden (vgl. 2.2.5).

Beim Vergleich 1 – 4S konnten zwei hochregulierten DE-microRNAs selektiert werden, jedoch konnten keine Targets ermittelt werden, die eine entgegengesetzte Expression aufwiesen und/oder von mindestens zwei der Datenbanken vorhergesagt wurden (vgl. 2.2.5). Aus diesem Grund wurde für die hochregulierten DE-microRNAs des Vergleiches keine microRNA-Target Beziehung erstellt. Trotz einer geringen Anzahl an differentiell exprimierten microRNAs aus einigen Vergleichen, wie z. B. 1/2/3 – 4/4S oder 1 – 4S (vgl. Abb. 3.19), kann gezeigt werden, dass zwischen den Stadien der untersuchten Tumorproben bekannte onkogene microRNAs hoch- und tumorsuppressive herunterreguliert werden (vgl. Abb. 3.20). Vornehmlich die Vergleiche der aggressiveren und risikoreicheren Stadien 1 – MYCN und 1 – 4^{+M} weisen dabei eine starke Veränderung in der microRNA Expression auf, wie es bereits aus der differentiellen Expression zu erwarten ist (vgl. 3.2.2). Wie zuvor erwähnt zeigen die letzten drei Vergleiche des weniger aggressiven Stadiums 4S mit den Proben des aggressiven Tumors eine geringe Anzahl an DE-microRNAs. Darüber hinaus wird hier anhand der differentiellen Expression deutlich, dass im Stadium 4S diverse microRNAs keine Veränderung aufweisen oder eine teilweise Umkehrung der microRNA Expression statt findet. Eine unveränderte Expression dieser microRNAs legt nahe, dass sie in diesem Tumorstadium gleichermaßen exprimiert werden, verglichen zu den aggressiveren Stadien. Die Umkehrung der Expression zeigt, dass einige microRNAs, vor allem onkogen wirkende, nach unten reguliert werden (z. B. *hsa-miR-17-5p*, *-92b-3p*), während andere microRNAs, darunter auch bekannte tumorsuppressiv wirkende, hochreguliert werden (z. B. *hsa-miR-204-5p*, *-744-5p*). Diese microRNAs können entscheidend an der besseren Prognose, der geringeren Aggressivität und der Möglichkeit zur spontanen Regression im Tumorstadium 4S beteiligt sein (vgl. hierzu Abschnitt 1.2).

Die aus der Targetanalyse erhaltenen Daten zeigen bekannte molekulare Mechanismen, wie die erwartete Präsenz von onkogenen microRNAs in aggressiven Tumoren und tumorsuppressiven microRNAs in weniger aggressiven Stadien und können dazu genutzt werden beschriebene microRNAs des Neuroblastoms aus verschiedenen Publikationen zu bestätigen (vgl. [155, 224]). Weiterhin lassen sich aus Abb. 3.20 diverse microRNAs entnehmen, die nicht allgemeingültig als onkogen oder tumorsuppressiv wirkend (im Neuroblastom) beschrieben sind, z. B. *hsa-miR-30e-3p*, *-92b-3p*, oder die nur zwischen bestimmten Stadien ermittelt werden, wie beispielsweise *hsa-miR-769-5p*, *-1291*, und somit auf eine mögliche, vielleicht neue, Funktion oder Regulation in einzelnen Stadien des Tumors hinweisen könnten. Um dies *in silico* genauer zu untersuchen, sind detailliertere Betrachtungen der ermittelten microRNA-Target Beziehungen

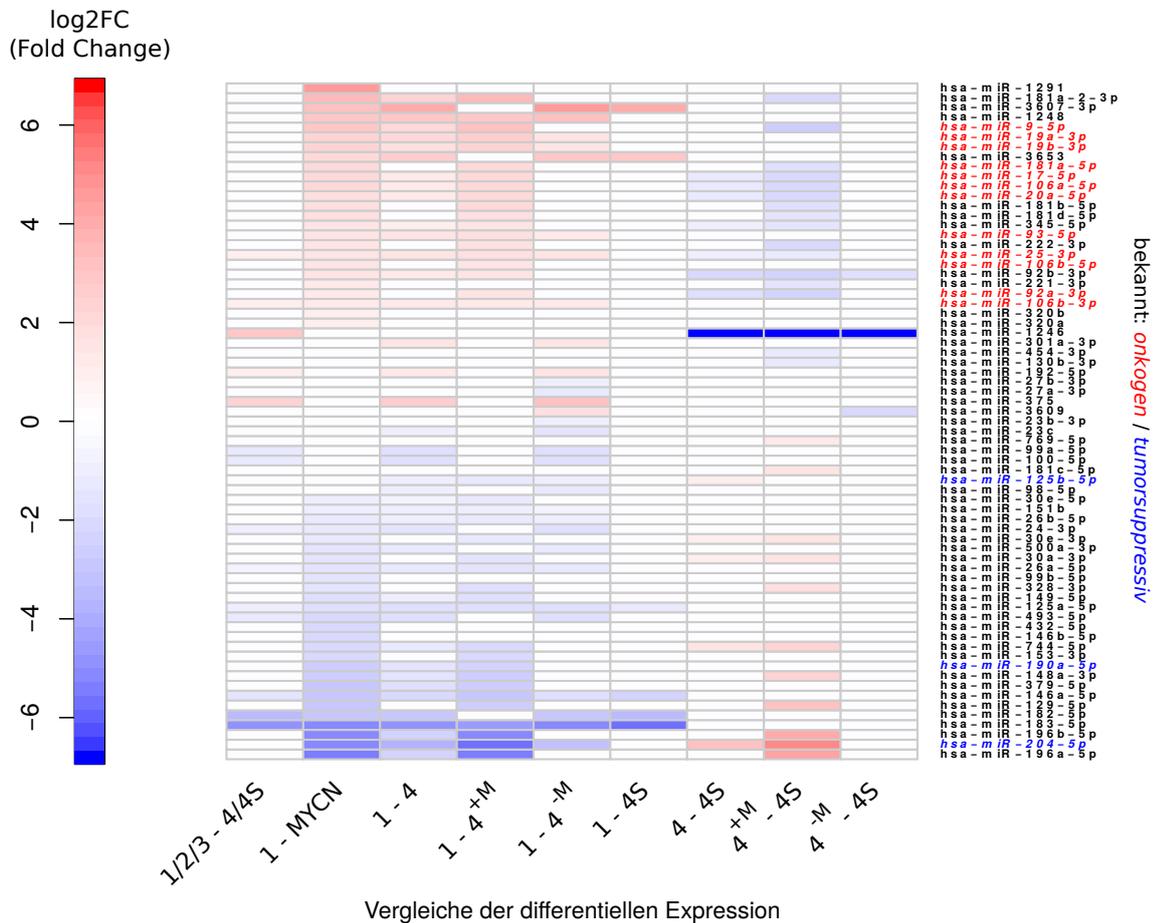


Abbildung 3.20.: Differentielle Expressionen der mittels Targetanalyse ermittelten microRNAs. Dargestellt sind die Veränderungen der 69 ermittelten microRNAs aus den differentiellen Vergleichen der Neuroblastomstadien durch deren logFC-Werte. Nicht signifikant und unveränderte microRNAs sind weiß, signifikant hochregulierte microRNAs eines Vergleiches sind rot, herunterregulierte blau dargestellt. Die Namen bereits bekannter onkogener und tumorsuppressiver microRNAs sind kursiv in rot respektive blau gekennzeichnet. Es fand keine Skalierung der logFC-Werte statt, um neben den Veränderungen zwischen den verglichenen Stadien auch die Veränderungen innerhalb eines Vergleiches zwischen den Expressionen der microRNAs deutlich zu machen. Die Sortierung der microRNAs erfolgte entsprechend der logFC-Werte aus dem Vergleich 1 – MYCN , welcher das aggressivste und risikoreichste Stadium mit MYCN Amplifikation gegen das wenig aggressive und risikoärmere Stadium 1 vergleicht.

notwendig. Als Grundlage hierzu und zur Klärung der Frage ob sich neben den bereits bekannten Targets auch putativ neue Targets oder Beziehungen bestätigen lassen, dienen die weiterführenden Analysen der nachfolgenden Abschnitte sowie die resultierenden Tabellen der microRNA-Targetanalyse, welche aus dem Verweis im Anhang F.2 entnommen werden können.

Bekannte Beziehungen

Um die Eignung der entwickelten Methode zur Ermittlung von microRNA-Target Beziehungen zu testen, kann - ohne experimentelle Nachweise - nur geprüft werden ob sich erwiesene onkogene und tumorsuppressive microRNAs des Neuroblastoms sowie ihrer beschriebenen Targets *in silico* bestätigen lassen. Dazu wurden mittels Literaturrecherchen beschriebene und wenn möglich nachgewiesene microRNA-Target Beziehungen im Neuroblastom und/oder weiteren Krebserkrankungen herausgesucht und bezüglich der in dieser Arbeit ermittelten Beziehungen untersucht. Anhand der Expression in Abb. 3.18 und der differentiellen Expression der ermittelten microRNAs in Abb. 3.20 lassen sich etwa die Veränderungen des stark onkogenen miR-17~92 Clusters sowie der paralogen Cluster miR-106a~363 und miR-106b~25 erkennen. Der *MYCN* induzierte Cluster weist eine hohe differentielle Expression in den aggressiven Tumoren auf, vor allem natürlich in Vergleichen mit *MYCN* amplifizierten Proben. In den weniger aggressiven Stadien zeigt sich eine unveränderte Expression der zugehörigen microRNAs. Bedingt durch die *MYCN*-Induzierung des Clusters, weist der Vergleich 1 – *MYCN* die stärkste (differentielle) Expression des Clusters auf (vgl. Abb. 3.18 und 3.20). Die microRNAs des Cluster regulieren eine Vielzahl diverser Targetgene unterschiedlicher Signalwege, was einen entscheidenden Einfluss auf die negative Prognose und die Überlebenswahrscheinlichkeit in diesem Tumorstadium zur Folge hat (vgl. [46,159,165]). Mittels microRNA-Targetanalyse lassen sich beim Vergleich 1 – *MYCN* der analysierten Neuroblastomdaten die beiden beschriebenen Targets *DKK3* und *TGFBR2* des 17~92 Clusters bestätigen. Beide Targets sollten durch den *MYCN* induzierten Cluster inhibiert werden und eine verringerte Expression aufweisen. Die zusammengefassten Ergebnisse in Tab. 3.12 stellen die extrahierten microRNAs des Clusters dar, für die eine Bindung mit den beiden Targets ermittelt wurde. Des Weiteren ist die Anzahl der Datenbanken, die diese Bindung vorhergesagt haben, sowie der berechnete $Score_{pdB}$ aus der Bewertung der ermittelten microRNA-Target Beziehung angegeben (vgl. hierzu Abschnitt 2.2.5). Dieser Score gibt unter Einbeziehung der Vorhersagedatenbanken, der Überlebenswahrscheinlichkeit und der Anzahl an microRNA-Bindestellen des Targets an,

microRNA	DKK3		TGFBR2	
	#DB	$Score_{pdB}$	#DB	$Score_{pdB}$
hsa-miR-17-5p	-	-	7	0.30
hsa-miR-19a-3p	6	0.23	8	0.33
hsa-miR-19b-3p	6	0.23	9	0.37
hsa-miR-20a-5p	-	-	9	0.37
hsa-miR-92a-3p	6	0.24	-	-
hsa-miR-92b-3p	6	0.24	-	-
logFC	-2,06		-1,71	
FDR	0,0004		0,033	

Tabelle 3.12.: Ergebnis der microRNA-Targetanalyse im Vergleich 1 – MYCN für die bekannten Targets *DKK3* und *TGFBR2* des onkogenen, *MYCN* induzierten, miR-17~92 Clusters des Neuroblastoms. Gelistet sind die ermittelten microRNAs, welche die Targets regulieren könnten, sowie die Anzahl der Vorhersagedatenbanken und der $Score_{pdB}$ für die ermittelte microRNA:mRNA Bindung. Des Weiteren sind die logFC- und die FDR-Werte der Targets für die differentiellen Expression zwischen den Stadien des Vergleiches angegeben.

wie wahrscheinlich eine Regulierung des Targets durch die microRNA ist, wobei ein maximalen Wert von 1.0 erreichen werden kann (vgl. 2.2.5). In den *MYCN* amplifizierte Tumoren findet eine signifikante Runterregulation beider Targets statt (negativer logFC-Wert), die dabei teilweise von verschiedenen hochregulierten microRNAs des Clusters beeinflusst werden. Für die Targets *DKK3* und *TGFBR2* ist beschrieben, dass der Cluster diese negativ reguliert, was sich hier anhand der analysierten Daten in dieser Arbeit bestätigen lässt.

Zur Veranschaulichung der Regulation der Targets sowie damit einhergehender Prozesse ist in Abb. 3.21 die schematische Wirkungsweise dargestellt. Für *DKK3* (*Dickkopf WNT Signaling Pathway Inhibitor 3*) ist beschrieben, dass es in Wechselwirkungen mit dem Wnt-Signalweg zur Reaktion der Zelle auf äußere Einflüsse steht und es indirekt durch den *MYCN* induzierten 17~92 Cluster reguliert wird, was mit einer schlechten Prognose des Krankheitsverlaufes korreliert [51, 114]. Speziell für die beiden in dieser Arbeit ermittelten microRNAs *hsa-miR-19b* und *hsa-miR-92a* wurde bereits in anderen Arbeiten eine direkte Bindung und Regulation der *DKK3* mRNA nachgewiesen [51]. Tabelle 3.12 zeigt, dass diese beiden microRNAs mithilfe der Targetanalyse ermittelt werden können. Daneben sind weitere microRNAs des Clusters vorhanden, die möglicherweise direkt *DKK3* regulieren könnten, was sich anhand der aktuellen Literatur jedoch nicht belegen lässt. Das mit dem TGF β -Signalweg assoziierte zweite Target *TGFBR2* (*Transforming Growth Factor Beta Receptor 2*)

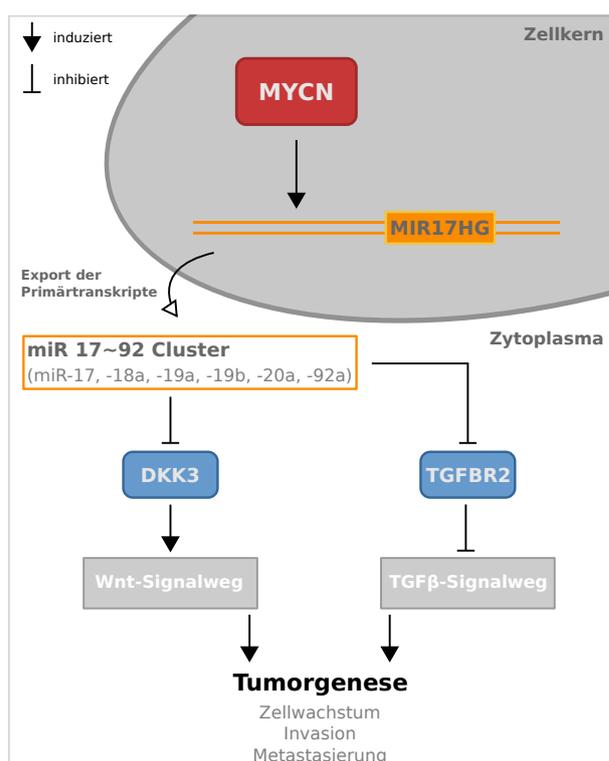


Abbildung 3.21.: Schematische Darstellung der Wirkungsweise des miR-17~92 Clusters. Gezeigt ist die Wirkung des MYCN induzierten Host Gene (*MIR17HG*) des onkogenen microRNA Clusters auf die in der in der Literatur beschriebenen Targets *DKK3* und *TGFBR2*. Die bekannte microRNA-Target Beziehung beider Targets lässt sich mithilfe der entwickelten Methode der Targetanalyse aus den analysierten Neuroblastomdaten bestätigen.

spielt eine wesentliche Rolle in der embryonalen Entwicklung und reguliert Prozesse der Proliferation, Differenzierung, Apoptose und Adhäsion von Zellen. Neben diversen anderen Krebserkrankungen ist *TGFBR2* mit dem Neuroblastom assoziiert, wobei eine erhöhte Expression zur Aktivierung oder Verstärkung des Signalweges mit folglich Reduzierung des Zellwachstums und der Abnahme der Tumorbildung einhergeht [159,177]. Für dieses Target des 17~92 Clusters ist nachgewiesen, dass die microRNAs *hsa-miR-17* und *hsa-miR-20a* direkt an der Regulation beteiligt sind [159]. Anhand der Targetanalyse ließ sich diese microRNA-Target Beziehungen in den Daten bestätigen (vgl. Tab. 3.12). Darüber hinaus wurden die microRNAs *hsa-miR-19* und *hsa-miR-92* identifiziert, welche in der angegebenen Literatur nicht als direkte Regulatoren bestätigt werden konnten, diese jedoch eine identische microRNA Seed-Sequenz aufweisen und somit als putative Bindungspartner in Betracht gezogen werden können [159]. Die Herunterregulierung beider Targets durch die microRNAs des 17~92 Clusters bewirken in den *MYCN* amplifizierten Stadien des Vergleiches 1 – *MYCN* somit eine nachgewiesene Verstärkung der Tumorgenese, was für die Aggressivität und schlechte Prognose der Tumore mit einer *MYCN* Amplifikation spricht. Die Auswirkungen der Unterdrückung der tumorrepressiven Targets durch den onkogenen Cluster im Neuroblastom zeigt sich folglich auch in der Überlebenszeitanalyse (vgl. hierzu 2.2.5), wie in Abb. 3.22 dargestellt ist. Beide Targets zeigen

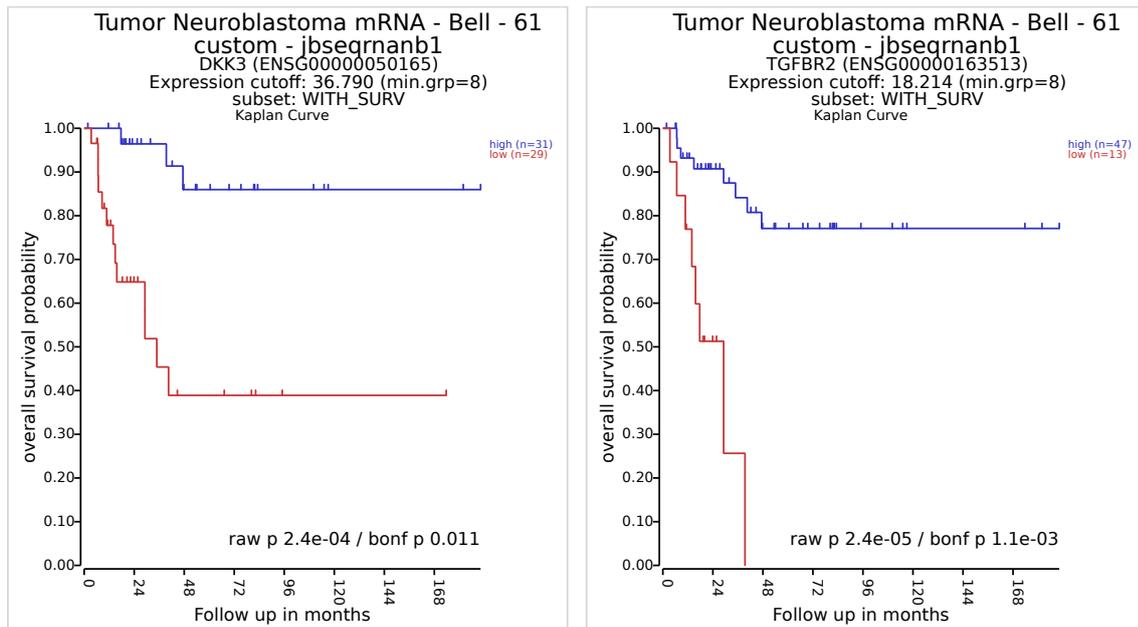


Abbildung 3.22.: Kaplan-Meier Kurven der Targets *DKK3* und *TGFBR2* des miR-17~92 Clusters. Anhand der Überlebenszeitanalyse wird deutlich, dass beide Targets bei geringer Expression (rote Kurve) im Neuroblastom zu einer signifikant verringerten Überlebenswahrscheinlichkeit führen.

anhand der analysierten Daten, dass eine geringe Expression eine signifikante Reduzierung der Überlebenswahrscheinlichkeit der Patienten zur Folge hat (rote Kurve). Im Detail betrachtet führt die Inhibierung von *DKK3* zu einer Verstärkung des Wnt-Signalweges, was eine höhere Expressions von Wachstumsfaktoren und eine erhöhte Transkription verschiedener Gene des Signalweges bewirkt. Die Herunterregulierung von *TGFBR2* hingegen führt zur Inhibierung des TGF β -Signalweges, was sich negativ auf Prozesse wie z. B. Apoptose auswirkt und die Tumorgenese fördert (vgl. Abb. 3.21). In nicht *MYCN* amplifizierten Tumoren, wie es z. B. durch den Vergleich 1 – 4^{-M} deutlich wird, ist zu erkennen, dass die microRNAs des Cluster nur noch teilweise differentiell exprimiert werden (vgl. Abb. 3.20). Auch lassen sich bei diesen Vergleichen die genannten Targetbeziehungen aus dem Vergleich 1 – *MYCN* nicht mehr mittels microRNA-Targetanalyse bestätigen. Des Weiteren zeigt sich erwartungsgemäß, dass beim Vergleiche von Stadium 4S mit aggressiven Tumoren eine Runterregulation des Clusters und eine Hochregulierung der Targets statt findet.

In einigen wenigen Krebserkrankungen, darunter auch dem Neuroblastom, ist die microRNA *hsa-miR-129-5p* beschrieben, der eine überwiegend tumorsuppressive Funktion beigemessen wird, wenngleich in einigen Fällen ein onkogenes Potential nachgewiesen wurde [173, 224, 279]. Eines der beschriebenen Haupttargets der microRNA ist *TWIST1* (*Twist Family Basic Helix-*

Vergleich	hsa-miR-129-5p		TWIST1	
	logFC	FDR	logFC	FDR
1 – MYCN	-2,86	2,20E-004	3,26	8,38E-004
1 – 4 ^{+M}	-2,54	0,03	3,38	2,52E-003
4 ^{+M} – 4S	3,00	0,01	-3,46	0,03

Tabelle 3.13.: Ergebnis der microRNA-Targetanalyse für die beschriebene Bindung zwischen der microRNA *hsa-miR-129-5p* und der *TWIST1* mRNA. Anhand der Analyse kann die Target-Beziehung beider RNAs bestätigt und das tumorsuppressive Potential der microRNA sowie die onkogene Wirkung des Targets im amplifizierten Neuroblastom verifiziert werden. Für die jeweiligen Vergleiche zwischen den Stadien, in denen die microRNA-Target Beziehung ermittelt wurde, sind die logFC- und die FDR-Werte der differentiellen Expression angegeben.

Loop-Helix Transcription Factor 1), das den Prozess der Epithelial-mesenchymalen Transition (EMT) fördert und in der Zellproliferation, der Invasivität sowie Metastasierung eine wesentliche Stellung einnimmt und welches mit *MYCN* co-exprimiert wird [199,279]. In den hier analysierten Daten weist die *hsa-miR-129-5p* eine Runterregulation im aggressiven und risikoreichen Stadien des Neuroblastoms mit *MYCN* Amplifikation auf. Wohingegen sich die microRNA in den anderen Vergleichen, mit Ausnahme von 4^{+M} – 4S, unverändert zeigt. Im Vergleich 4^{+M} – 4S wird die microRNA signifikant hochreguliert, was auf eine erhöhte Expression im weniger aggressiven Stadium 4S, verglichen mit *MYCN* amplifizierten Tumoren des Stadiums 4, hinweist (vgl. Tab. 3.13). Darüber hinaus kann die *TWIST1* mRNA als Target der *hsa-miR-129-5p* in den Vergleichen 1 – MYCN, 1 – 4^{+M} und 4^{+M} – 4S verifiziert werden. Dabei zeigt das Target eine signifikant starke Hochregulierung in den *MYCN* amplifizierten Stadien sowie eine ähnlich starke Runterregulation im Stadiums 4S (vgl. Tab. 3.13). Aus den in dieser Arbeit untersuchten Daten kann anhand der Targetanalyse somit die beschriebene microRNA-Target Beziehung als auch das onkogene Potential von *TWIST1* und die tumorsuppressive Wirkung der *hsa-miR-129-5p* für das Neuroblastom bestätigt werden. Das tumorsuppressive Potential der microRNA als auch die onkogene Wirkung des Targets können ebenfalls anhand der Überlebenszeitanalyse verifiziert werden. Dazu zeigt Abb. 3.23 die Kaplan-Meier Kurven, die eine signifikante Verringerung der Überlebenswahrscheinlichkeit bei verminderter Expression der microRNA (Abb. links) und gleichzeitig erhöhter Expression des Targets (Abb. rechts) bestätigen⁵⁵.

⁵⁵Die unterschiedliche Probenanzahl von 66 bzw. 61, dieser und der folgenden Kaplan-Meier Kurven, ergibt sich aus den genannten Unterschieden der Anzahl in der microRNA- und mRNA-Sequenzierung (vgl. Abschnitt 2.2.2).

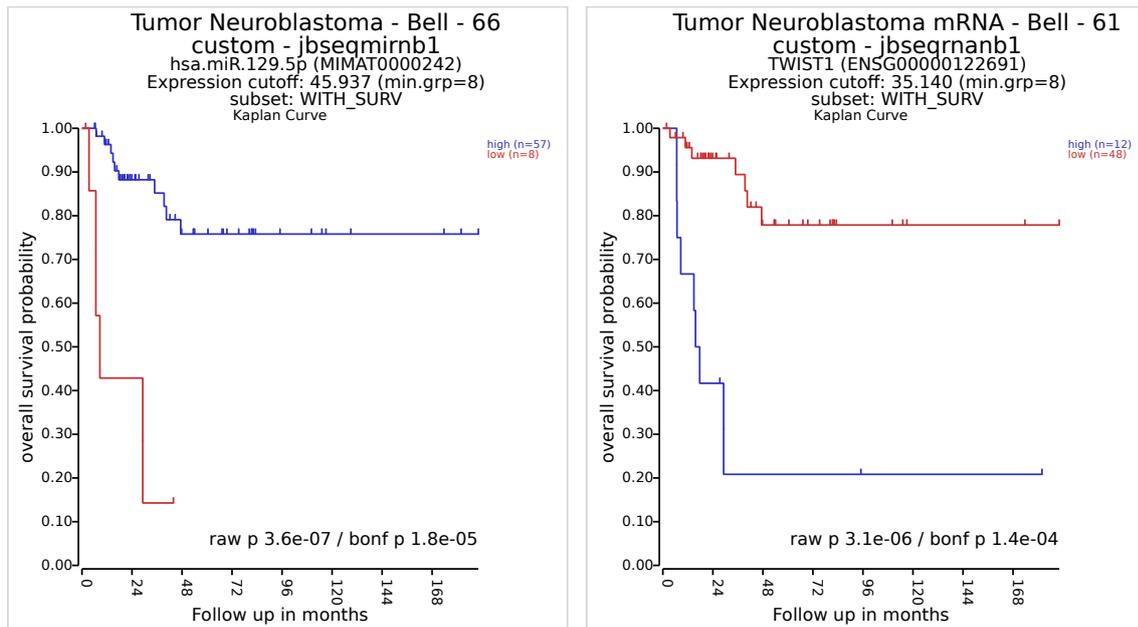


Abbildung 3.23.: Kaplan-Meier Kurven der microRNA *hsa-miR-129-5p* und des beschriebenen Targets *TWIST1*. Anhand der Überlebenszeitanalyse wird deutlich, dass eine verringerte Expression der tumorsuppressiven microRNA (rote Kurve, links) und die erhöhte Expression des onkogenen Targets (blaue Kurve, rechts) zu einer signifikant verringerten Überlebenswahrscheinlichkeit führen.

Ein weiteres Beispiel für eine tumorsuppressiv wirkende microRNA ist die *hsa-miR-204*. Sie wird als Biomarker zur Diagnose sowie Prognose des Krankheitsverlaufes verschiedener Krebserkrankungen genutzt, wie beispielsweise dem Neuroblastom, Mammakarzinom, Gliom etc. Eine Herunterregulation der microRNA Expression ist mit einer schlechteren Prognose der Krankheit assoziiert [28, 130, 155, 216, 272]. Für die microRNA sind in der genannten Literatur aktuell nur wenige Targets beschrieben, die darüber hinaus eine genauere Untersuchung erfordern. Auf den Zellwachstumshemmer Cisplatin wirkt die microRNA teilweise sensitivitätssteigernd durch die negative Regulation des anti-apoptischen Faktors *BCL2* (*B-Cell Lymphoma 2*). Des Weiteren ist für die microRNA die direkte Regulation des *NTRK2* (*Neurotrophic Receptor Tyrosine Kinase 2*) Onkogens nachgewiesen, welches eine wesentliche Rolle bei der Chemotherapie-resistenz des Neuroblastoms spielt. Das Target *BCL2* weist in keiner der Vergleiche eine differentielle Expression zwischen den Stadien auf womit sich die beschriebene microRNA-Target Beziehung anhand der analysierten Daten nicht identifizieren lässt. *NTRK2* hingegen weist in *MYCN* amplifizierten Stadien eine signifikant starke differentielle Expression auf und lässt sich mithilfe der Methode als Target der *hsa-miR-204* bestätigen. In Tabelle 3.14 sind die logFC- und FDR-Werte der microRNA und des

Vergleich	hsa-miR-204-5p		NTRK2	
	logFC	FDR	logFC	FDR
1 – MYCN	-5,14	1,60E-006	2.55	6,14E-004
1 – 4 ^{+M}	-5,64	1,40E-004	2.76	6,88E-004

Tabelle 3.14.: Ergebnis der microRNA-Targetanalyse für die beschriebene Bindung zwischen der microRNA *hsa-miR-204-5p* und der *NTRK2* mRNA. Anhand der Analyse kann für amplifizierte Tumoren des Neuroblastoms die Target-Beziehung zwischen microRNA und mRNA bestätigt werden. Die veränderte Expression bestätigt zudem das tumorsuppressive Potential der microRNA sowie die onkogene Wirkung des Targets. Für die jeweiligen Vergleiche, in denen die microRNA-Target Beziehung ermittelt wurde, sind die logFC- und die FDR-Werte der differentiellen Expression angegeben.

Targets in den differentiellen Vergleichen aufgelistet. Anhand der veränderten Expression ist eine Herunterregulierung der *hsa-miR-204* sowie eine Hochregulierung von *NTRK2* in *MYCN* amplifizierten Tumoren erkennbar, was das beschriebene tumorsuppressive Potential der microRNA und die onkogene Wirkung von *NTRK2* in aggressiven und risikoreichen Tumoren bestätigt. Das tumorsuppressive und onkogene Potential beider Transkripte lässt sich ebenso anhand der Überlebenszeitanalyse verifizieren (vgl. Abb. 3.24). Die Kaplan-Meier Kurven zeigen eine signifikante Erhöhung der Überlebenswahrscheinlichkeit durch eine verstärkte Expression der microRNA (Abb. 3.24 links) sowie eine signifikant verringerte Wahrscheinlichkeit bei erhöhter Expression des Targets (Abb. 3.24 rechts).

Anhand der ausgewählten Beispiele bekannter krebsbezogener microRNAs, welche auch für das Neuroblastom eine wesentliche Stellung im Bezug auf die Diagnose, den Verlauf sowie die Therapie der Erkrankung einnehmen, wird deutlich, dass sich mithilfe der entwickelten Methode der Targetanalyse microRNA-abhängige Regulationen nachgewiesener Targets bestätigen lassen. Zudem zeigen die Ergebnisse, dass neben der Bindung von microRNAs und putativen Targets auch Rückschlüsse über das onkogene bzw. tumorsuppressive Potential der microRNAs und Targets anhand der *in silico* Analysen aus der Expressionsveränderung zwischen den Stadien ermittelt werden können. Neben der hier bestätigten Auswahl an Transkripten ist zu erwähnen, dass zum einen für die selektierten microRNAs der Analyse zahlreiche Targets von den Datenbanken vorhergesagt werden, deren tatsächliche Bindung und Regulation durch die microRNAs unter Umständen noch nicht experimentell validiert wurden. Zum anderen lassen sich gegebenenfalls nicht alle in der Literatur beschriebenen Targets einer microRNA aus den in dieser Arbeit analysierten Daten bestätigen. Die hierfür vorliegenden Gründe können vielfältig sein, wie beispielsweise die fehlende signifikante Veränderung der microRNA oder

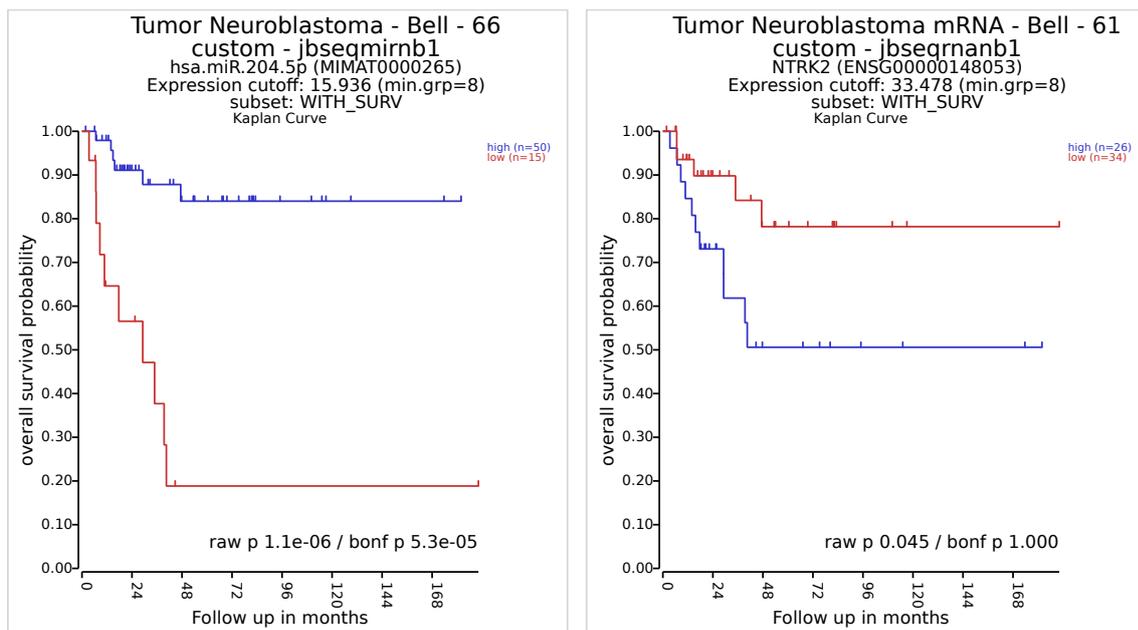


Abbildung 3.24.: Kaplan-Meier Kurven der microRNA *hsa-miR-204-5p* und des beschriebenen Targets *NTRK2*. Anhand der Überlebenszeitanalyse wird deutlich, dass eine verringerte Expression der tumorsuppressiven microRNA (rote Kurve, links) und eine Erhöhung der Expression des onkogenen Targets (blaue Kurve, rechts) zu einer verringerten Überlebenswahrscheinlichkeit führen. Dabei zeigt sich für das Target *NTRK2* zwar keine signifikante Verringerung der Überlebenswahrscheinlichkeit bei Betrachtung des korrigierten p-Wertes, dennoch ist der negative Trend bei einer erhöhten Expression erkennbar, der zu einer auf rund 50 % reduzierten Überlebenswahrscheinlichkeit führt.

des Targets zwischen den untersuchten Proben (siehe *BCL2* als Target der *hsa-miR-204*), der Targetnachweis der Literatur aus Zelllinien, Methoden zur Gewinnung der Expressionsdaten (Microarrays oder PCR-Analysen), Einfluss verschiedener RNA-bindender Proteine (RBPs), welche die microRNA Wirkung modulieren etc.

Putativ neue Beziehungen

Basierend auf den Ergebnissen der microRNA-Targetanalyse lassen sich neben bekannten auch putativ neue microRNAs, Targets als auch neue microRNA-Target Beziehungen im Neuroblastom identifizieren. Die Prüfung auf einen bestehenden Zusammenhang zum Neuroblastom erfolgte mittels Recherche über aktuell verfügbarer Literatur anhand der *PubMed*⁵⁶ Datenbank der *National Library of Medicine (USA)*. Dazu wurden Verbindungen der ermittelten microRNAs und ihrer möglichen Targets hinsichtlich bekannter Veröffentlichungen

⁵⁶<https://www.ncbi.nlm.nih.gov/pubmed/>

im Zusammenhang mit dem Neuroblastom gesucht und inhaltlich geprüft. Tabelle 3.15 zeigt hierzu die 67 am besten bewerteten microRNA:mRNA Bindungen die mittels Targetanalyse identifiziert wurden und von denen die microRNA und/oder das Target noch nicht direkt mit dem Neuroblastom assoziiert sind. Die linke Seite der Tabelle zeigt für jeden der stadienbasierten Vergleiche die differentiell hochregulierten microRNAs und deren herunterregulierten Targets, während die rechte Seite der Tabelle die runterregulierten microRNAs und hochregulierten Target mRNAs zeigt. Zudem ist der bereits erwähnte $Score_{pdB}$ zur Bewertung der gelisteten Beziehung angegeben, der es ermöglicht die Bindungen innerhalb und zwischen den Vergleichen der Stadien zu bewerten (vgl. 2.2.5). Die vollständige Bestenliste der bis zu zehn am besten bewerteten Bindungen kann in Tab. F.4 des Anhangs eingesehen werden, ein Verweis auf die Gesamtlisten aller ermittelten microRNA-Target Beziehungen befindet sich zudem im Anhang F.2.

Für die Betrachtung putativ neuer Beziehungen im Neuroblastom sollen die ermittelten Bindungen untersucht werden, von denen weder die microRNA noch das ermittelte Target sich aus aktuellen Publikationen mit dem Neuroblastom assoziieren lassen und die somit auf eine mögliche neue microRNA-Target Bindung zwischen den Stadien des Neuroblastoms hindeuten könnten. Von den 67 ermittelten, trifft dieses Kriterium auf sieben microRNA:mRNA Bindungen zu (vgl. Tab. 3.15, microRNA und Target sind hervorgehoben). Dabei konnten zwei Bindungen aus dem Vergleich 1/2/3 – 4/4S , jeweils eine Bindung aus 4 – 4S und 4^{+M} – 4S und drei aus dem Vergleich 4^{-M} – 4S ermittelt werden. Wie bereits erwähnt, kann auch hier letztendlich kein realer Nachweis für die direkte Bindung an sich oder die Funktion im Neuroblastom erbracht werden. Alle putativ neuen Bindung sollen jedoch anhand der verfügbaren Literatur untersucht und hinsichtlich einer möglichen Assoziation zum Neuroblastom betrachtet werden. Für alle im Folgenden betrachteten Beziehungen wären jedoch weiterführende Untersuchungen in Form von experimentellen Nachweisen notwendig, um genauen Aufschluss über die Bindung sowie mögliche Funktionen und Relevanz im Neuroblastom aufzuzeigen. Der erste hier betrachtete Vergleich 1/2/3 – 4/4S , welcher den Vergleich zwischen den lokalisierten bzw. geringfügig metastasierenden Tumoren mit den stark metastasierenden darstellt, resultiert in zwei putativ neue Bindungen, die anhand der hier durchgeführten Targetanalyse ermittelt wurden. Zum einen die Bindung *hsa-miR-1246:FAM124A* einer hochregulierten und zum anderen die Bindung *hsa-miR-493-5p:WDR43* einer runterregulierten microRNA aus dem Vergleich. Die *hsa-miR-1246* der ersten Bindung steht dabei nicht völlig außerhalb des Neuroblastom-Kontext. In der Literatur

wird die microRNA in einer Publikation im Zusammenhang mit der Infektion einer untersuchten

Vergleich	Target ↓	microRNA ↑	$Score_{pdB}$	Target ↑	microRNA ↓	$Score_{pdB}$
1/2/3 – 4/4S	FAM124A	miR-1246	0,149	LIN28B	miR-493-5p	0,198
	FAM124A	miR-375	0,078	WDR43	miR-493-5p	0,189
1 – MYCN				CREBZF	miR-26b-5p	0,378
				LBR	miR-148a-3p	0,374
				RPS6KA6	miR-26b-5p	0,367
1 – 4	ANKRD52	miR-93-5p	0,347	LBR	miR-148a-3p	0,374
	MINK1	miR-20a-5p	0,342	CLP1	miR-98-5p	0,368
	MINK1	miR-106a-5p	0,342	CCDC47	miR-196a-5p	0,358
	TOM1L2	miR-19b-3p	0,326	TIMM17B	miR-98-5p	0,349
	ANKRD52	miR-20a-5p	0,312	RAVER2	miR-99a-5p	0,343
1 – 4^{+M}				LBR	miR-148a-3p	0,374
				RPS6KA6	miR-26b-5p	0,367
1 – 4^{-M}	ANKRD52	miR-93-5p	0,347	FBXW7	miR-27a-3p	0,412
	TOM1L2	miR-19b-3p	0,326	CLP1	miR-98-5p	0,368
	SERPINE1	miR-301a-3p	0,301	TIMM17B	miR-98-5p	0,349
	CHD5	miR-301a-3p	0,294	YOD1	miR-98-5p	0,344
	ZEB2	miR-25-3p	0,287	RAVER2	miR-99a-5p	0,343
				MET	miR-23b-3p	0,341
			TAF5	miR-98-5p	0,336	
1 – 4S				KRT1	miR-146a-5p	0,238
				COL17A1	miR-183-5p	0,183
				PKP1	miR-125a-5p	0,181
				PKP1	miR-182-5p	0,181
				PKP1	miR-183-5p	0,181
				KRT10	miR-146a-5p	0,177
				KRT6B	miR-146a-5p	0,177
				DSC3	miR-125a-5p	0,145
				FLG	miR-125a-5p	0,142
				KRT6B	miR-183-5p	0,142
4 – 4S	CDT1	miR-204-5p	0,113	MINK1	miR-20a-5p	0,342
	PRR11	miR-204-5p	0,111	MINK1	miR-106a-5p	0,342
	HJURP	miR-744-5p	0,078	COL1A2	miR-92a-3p	0,318
				MINK1	miR-17-5p	0,307
			ZEB2	miR-92b-3p	0,287	
4^{+M} – 4S	HNRNPA1	miR-129-5p	0,313	NDEL1	miR-130b-3p	0,375
	ELOVL6	miR-204-5p	0,280	ZBTB4	miR-130b-3p	0,360
	FAM19A5	miR-196a-5p	0,259	ZFPM2	miR-130b-3p	0,360
	FAM19A5	miR-196b-5p	0,259	ZEB2	miR-221-3p	0,358
	FRS2	miR-129-5p	0,252	ZFPM2	miR-454-3p	0,353
4^{-M} – 4S				KRT1	miR-92b-3p	0,238
				MATN2	miR-92b-3p	0,203
				MATN2	miR-1246	0,203
				ABR	miR-92b-3p	0,195
				DSG1	miR-92b-3p	0,177
				CREB5	miR-1246	0,152
				CREB5	miR-3609	0,152
				MATN2	miR-92b-3p	0,132

Tabelle 3.15.: Auszug der ermittelten microRNA-Target Bindungen aus der finalen Extraktion der Targetanalyse. Gelistet sind die 67 putativ neuen microRNA:mRNA Bindungen der bis zu zehn am besten bewerteten Bindungen für die hoch- und runterregulierten microRNAs aus den stadienbasierten Vergleichen. Sind die microRNA und/oder das Target aktuell nicht direkt mit dem Neuroblastom assoziierte, so sind diese als fett gedruckt gekennzeichnet. Um die Dimension der Tabelle gering zu halten, wurden die Namen der microRNAs in verkürzter Schreibweise, ohne des Präfix „hsa-“, dargestellt.

Zelllinie des Neuroblastoms (SH-SY5Y) mit dem humanen Enterovirus 71 (HEV71) genannt [275]. Dabei wird jedoch nur eine Zelllinie betrachtet und keine Untersuchungen an primären Tumorproben durchgeführt. Zudem stützen sich die Untersuchungen auf die Virusinfektion der Zellen, wobei beeinflusste mRNAs ermittelt wurden, die mit dem Zellzyklus, der Proliferation und der Apoptose zusammenhängen. Eine Runterregulierung der microRNA hat dabei zu einer signifikanten Veränderung in der Expression von *DLG3* (*Disc-Large Homolog 3*) geführt, welches

mit neuronalen Störungen in Verbind gebracht wird und wodurch schlussfolgernd die *hsa-miR-1246* mit neuronalen Krankheitsentwicklungen assoziiert wird (vgl. [275]). Ergänzende Untersuchungen fanden jedoch nicht statt, so dass die biologische Relevanz der microRNA für neuronale Erkrankungen oder das Neuroblastom weiterer Nachforschungen bedürfen und hier dennoch als nicht direkt mit dem Neuroblastom assoziiert betrachtet wird. Aus weiteren Publikationen ist bekannt, dass die *hsa-miR-1246* regulatorischen Einfluss auf die Metastasierung, Migration und Invasivität verschiedener Tumorerkrankungen hat, wie beispielsweise im Hepatozellulären- oder Bronchialkarzinom [93, 239]. Für das identifizierte Target *FAM124A* konnten keinerlei Zusammenhänge zum Neuroblastom, Tumor- oder allgemeinen Erkrankungen aus der Literatur ermittelt werden. In der Analyse dieser Arbeit zeigen beide Partner eine signifikante Veränderung in ihrer Expression, zudem wird die Bindung *hsa-miR-1246:FAM124A* von vier der 14 Datenbanken vorhergesagt (vgl. 2.2.5). In der Überlebenszeitanalyse zeigt jedoch lediglich die microRNA eine signifikante Veränderung, wobei eine hohe Expression zu einer verringert Überlebenswahrscheinlichkeit und somit zu einer schlechten Prognose führt. An dieser Stelle ist es somit nicht möglich ohne weitere experimentelle Nachweise eine Aussagen über die Zuverlässigkeit der ermittelten Bindung oder die Assoziation zum Neuroblastom und einer damit verbundenen möglichen Funktion zu treffen. Tendenziell geben die hier gemachten Befunde jedoch einen Hinweis darauf, dass die *hsa-miR-1246* mit dem Neuroblastom assoziiert ist und womöglich an Prozessen der Metastasierung beteiligt sein könnte.

Eine zweite aus diesem Vergleich ermittelte Bindung besteht zwischen der microRNA *hsa-miR-493-5p* und der *WDR43* mRNA. Dabei konnten keine wesentlichen Information über *WDR43* im Zusammenhang mit Tumorerkrankungen oder dem Neuroblastom in Erfahrung gebracht werden. Bei der *hsa-miR-493-5p* hingegen zeichnet sich eine Regulation des Wachstums und der Invasivität bei Krebserkrankungen ab sowie Befunde zur Beeinflussung der Chemosensitivität bzw. -resistenz [79, 97, 287]. Diese Erkenntnisse resultieren jedoch nicht aus Untersuchungen im Neuroblastom, sondern basieren auf diversen anderen Tumorerkrankungen, wie z. B. dem Mamma- oder dem Magenkarzinom. Anhand der Analyseergebnisse aus der vorliegenden Arbeit sind sowohl die microRNA als auch das Target zwischen den Stadien des Vergleiches 1/2/3 – 4/4S signifikant in ihrer Expression und Richtung verändert. Die ermittelte Bindung wurde zudem von fünf Datenbanken vorhergesagt. Veränderungen in der Überlebenszeitanalyse weisen jedoch keine Signifikanz auf, deuten aber entsprechend der Literatur eine schlechte Prognose bei

erhöhter microRNA Expression an. Zusammen, mit den geringfügigen Informationen aus der Literatur, lassen diese Befunde eine mögliche Verbindung der *hsa-miR-493-5p* zur Inhibierung der Invasivität und Metastasierung erkennen. Dies wiederum unterstützt die Erkenntnisse aus der Targetanalyse, dass die microRNA in den stark metastasierenden Stadien des Vergleiches 1/2/3 – 4/4S runterreguliert ist und auch im Neuroblastom auf eine Regulation der Invasivität, Metastasierung und Migration hindeutet. Genauere Aussagen zur tatsächlichen Bindung des Targets *WDR43* und zur biologischen Signifikanz im Neuroblastom lassen sich jedoch auch hier ohne weiterführende Experimente nicht treffen.

Im untersuchten Vergleich 4 – 4S wurde die Bindung *hsa-miR-744-5p:HJURP* identifiziert. Für die Partner der Bindung konnte aus der aktuellen Literatur keine direkte Interaktion nachgewiesen werden, noch sind beide Partner direkt mit dem Neuroblastom assoziiert. Die *hsa-miR-744-5p* wird zwar in einer Publikation zum Neuroblastom aufgelistet, ohne jedoch weiterführende Betrachtungen zu ihrer Funktion oder Stellung in der Tumorerkrankung vorzunehmen [155]. In anderen Krebserkrankungen ist die microRNA im Zusammenhang mit dem Wnt/ β -Catenin Signalweg und einer damit verbundenen schlechten Prognose beschrieben [289]. Des Weiteren konnte ein Zusammenhang zum TGF β -Signalweg nachgewiesen werden [149] und es wurde gezeigt, dass die *hsa-miR-744* sich als diagnostischer Biomarker eignet, beispielsweise im Pankreaskarzinom oder im Hepatozellulären Karzinom, wobei sich die Expression in den Erkrankungen unterschiedlich verhält [164, 242]. Aus den wenigen Befunden der Literatur kann ein Einfluss der microRNA auf Prozesse wie Proliferation, Differenzierung, Apoptose und die Adhäsion von Zellen abgeleitet werden, welche in der Tumorgenese eine wesentliche Stellung einnehmen [84]. Ein möglicher Zusammenhang zum Neuroblastom kann an dieser Stelle angenommen werden, bedarf jedoch aufgrund fehlender Nachweise weiterführender experimenteller Untersuchungen. Das aus der Analyse ermittelte Target *HJURP* (*Holliday Junction Recognition Protein*) weist ebenso keinerlei direkte Assoziationen zum Neuroblastom auf. *HJURP* steht im Zusammenhang mit der Faltung und Stabilität von Histonen und Histonkomplexen und zeigt in diversen Krebserkrankungen eine erhöhte Expression sowie Verringerung der Überlebensrate [255, 258]. Weiterhin wurde ein Zusammenhang zu Mechanismen der DNA-Reparatur im Glioblastom postuliert, wobei *HJURP* in die Reparatur von Doppelstrangbrüchen involviert ist [255]. Ähnlich wie bei der *hsa-miR-744-5p* kann somit auch für das Target *HJURP* ein Zusammenhang zum Neuroblastom an dieser Stelle nur vermutet werden. Auch die putativ direkte Bindung der microRNA zum Target kann nicht gänzlich

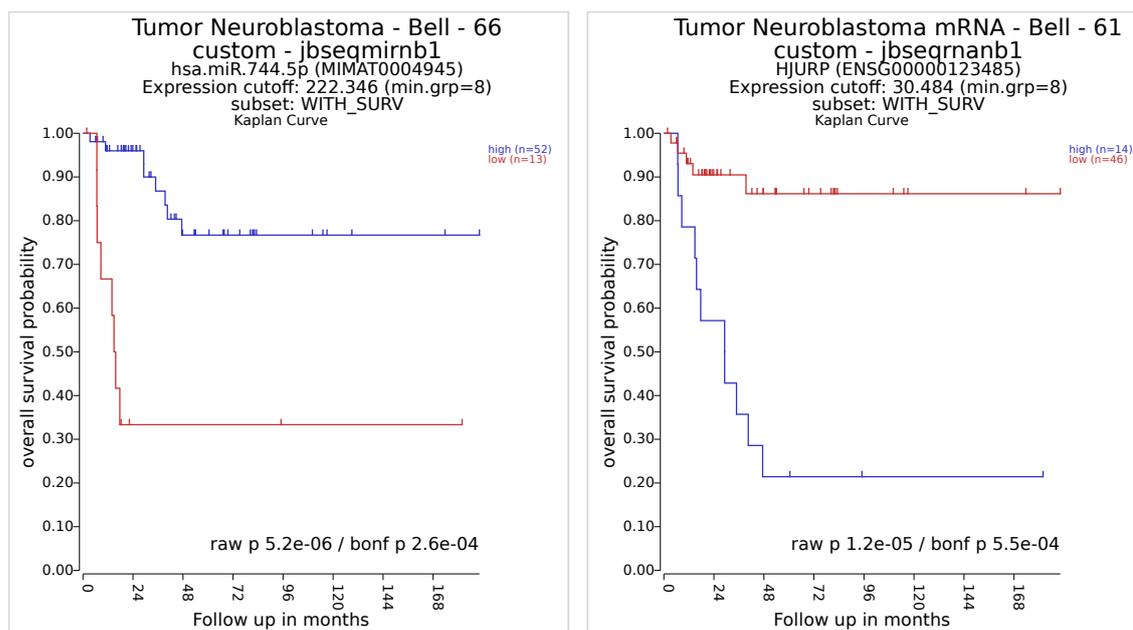


Abbildung 3.25.: Kaplan-Meier Kurven für die Partner der ermittelte Bindung *hsa-miR-744-5p:HJURP*, die eine putativ neue Beziehung im Zusammenhang mit dem Neuroblastom darstellen. Die Überlebenszeitanalyse lässt eine signifikant verringerte Überlebenswahrscheinlichkeit bei geringerer microRNA Expression (rote Kurve, links) und erhöhter *HJURP* Expression (blaue Kurve, rechts) erkennen.

nachgewiesen werden. Bei der Targetanalyse wird diese Bindung *hsa-miR-744-5p:HJURP* von nur zwei Datenbanken vorhergesagt und weist einen sehr geringen $Score_{pdB}$ von 0,078 auf. Dennoch zeigen beide Transkripte eine signifikant veränderte und entgegengesetzte Expression im Vergleich 4 – 4S sowie signifikante Veränderungen in der Überlebenszeitanalyse (vgl. Abb. 3.25). Dies weist zumindest auf eine putative Bindung beider Partner hin und unterstützt die beschriebenen Befunde, dass eine schlechte Prognose bei erhöhter *HJURP* Expression vorliegt. Für das Neuroblastom könnte dies eine tumorsuppressive Wirkung der *hsa-miR-744-5p* andeuten, welche im Stadium 4S hochreguliert ist und die Expression von *HJURP* negativ reguliert und somit zu einer verbesserten Prognose in diesem Stadium führt. Um letztendlich eine biologische Relevanz zum Neuroblastom oder zur Regulierung von *HJURP* durch die *hsa-miR-744-5p* zu klären sind weitere Untersuchungen auf Basis experimenteller Nachweise unumgänglich, da die vorliegenden Befunde und die aktuelle Literatur diese Fragen nicht ausführlich genug klären können.

Innerhalb des Vergleiches 4^{+M} – 4S konnte die Verbindung der *hsa-miR-130b-3p* mit der mRNA von *NDEL1* (*NudE Neurodevelopment Protein 1 Like 1*) als eine putativ neue Bindung bestimmt werden. Für die microRNA *hsa-miR-130b-3p* ist bekannt, dass sie an Prozessregulationen

wie Invasivität, Migration und dem Wachstum in Gliomzellen, dem Mammakarzinom sowie weiteren Krebserkrankungen beteiligt ist und sowohl onkogene als auch tumorsuppressive Funktionen einnimmt [78, 228]. Eine direkte Assoziation zum Neuroblastom konnte anhand aktueller Literatur nicht nachgewiesen werden. Mittels Targetanalyse wurde eine putativ direkte Bindung zwischen *hsa-miR-130b-3p:NDEL1* ermittelt, welche so weder für das Neuroblastom noch im Zusammenhang mit anderen Krebserkrankungen bekannt ist. Die Verbindung wurde durch zehn der 14 genutzten Datenbanken vorhergesagt, bei einem $Score_{pdB}$ von 0,375, was dem drittgrößten Score der Tabelle 3.15 entspricht und dem höchsten Score einer putativ neuen Verbindung. Zudem konnten signifikante Veränderungen in der Überlebenszeitanalyse beider RNAs ermittelt werden. Die zugehörigen Kaplan-Meier Kurven in Abb 3.26 zeigen die Ergebnisse der Überlebenszeitanalyse, die die Annahme einer putativ direkte Bindung ergänzend zu den *in silico* Vorhersagen und der entgegengesetzten differentiellen Expression unterstützen. Es kann gezeigt werden, dass eine erhöhte Expression der *hsa-miR-130b-3p* bei verminderter Expression der *NDEL1* mRNA zu einer signifikanten Verringerung der Überlebenswahrscheinlichkeit führen. Die Befunde der Targetanalyse und aus der Literatur legen nahe, dass eine direkte Bindung wahrscheinlich ist und diese in Prozesse der Tumorentwicklung eingreift. Im Mammakarzinom wurde für die microRNA nachgewiesen, dass sie den NOTCH-Signalweg durch die Regulierung beteiligter Liganden beeinflussen kann und dadurch die Invasivität und Migration des Tumors hemmt [228]. Der Signalweg dient der Regulierung des Wachstums und der Differenzierung neuronaler Vorläuferzellen und nimmt eine wesentliche Stellung bei der Zell-Zell-Kommunikation, der Migration, Invasivität und Metastasierung ein [228, 271, 273]. Eine Deregulierung des Signalweges ist mit verschiedenen Krebserkrankungen assoziiert und es ist bekannt, dass eine Aktivierung im Mammakarzinom zu einer Verschlechterung der Prognose und einer geringeren Überlebenswahrscheinlichkeit führt, gekoppelt an die Begünstigung aggressiver Tumorprozesse, wie beispielsweise die Epithelial-mesenchymale Transition (EMT) [228]. Der Signalweg wird wesentlich durch den Prozess der Zellteilung beeinflusst, wobei die neuronalen Vorläuferzellen mit der symmetrischen und der asymmetrischen Zellteilung zwei Arten von Teilungen durchlaufen. Die symmetrische Teilung führt zu einer Vermehrung der Zellen, die folglich durch Umschaltung zur asymmetrischen Zellteilung eine Differenzierung zur Ausbildung reifer Neuronen erfahren [271, 273]. Diese Arten der Zellteilung sind an die Orientierung der Spindelfasern gekoppelt, welche zum einen die Art der Teilung bestimmt und zum anderen den NOTCH-Signalweg beeinflussen. Eine fehlerhafte Orientierung kann

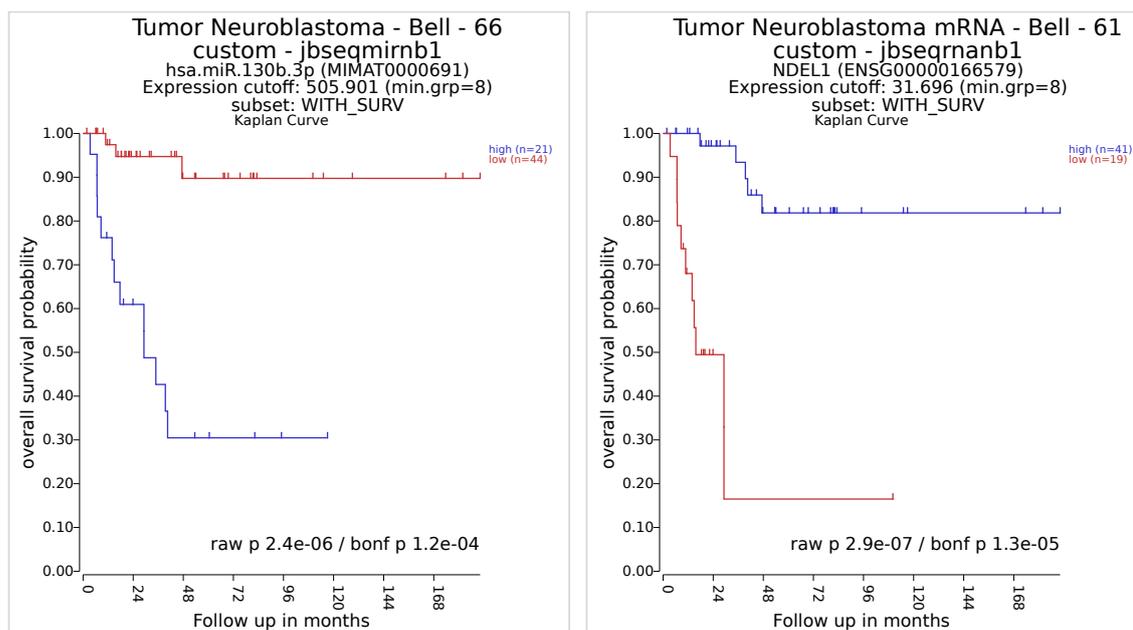


Abbildung 3.26.: Kaplan-Meier Kurven für die Partner der ermittelte Bindung *hsa-miR-130b-3p:NDEL1*, die eine putativ neue Beziehung im Zusammenhang mit dem Neuroblastom darstellen. Die Überlebenszeitanalyse lässt eine signifikant verringerte Überlebenswahrscheinlichkeit bei erhöhter microRNA (blaue Kurve, links) und verringerter *NDEL1* Expression (rote Kurve, rechts) erkennen.

darüber hinaus zum programmierten Zelltod führen [273]. Das aus der Targetanalyse ermittelte Target *NDEL1* steht im direkten Zusammenhang mit der Spindelorientierung, welche neben der bereits erwähnten *hsa-miR-130b* ebenfalls einen direkten Einfluss auf den NOTCH-Signalweg hat [271, 273]. Es ist nachgewiesen, dass *NDEL1* die Stabilität der Mikrotubuli beeinflusst, welche Grundbausteine der Spindelfasern darstellen. Zusammen mit *LIS1* (auch bekannt als *PFAFH1B1*) und dem zytoplasmatischen Motorprotein Dynein bildet *NDEL1* einen Komplex, der einen direkten Wirkung auf die Orientierung der Spindelfasern und Effekte im NOTCH-Signalweg hat [271, 273]. Unter Betrachtung der Literatur und der aus den Daten des Vergleiches $4^{+M} - 4S$ gezogenen Befunde lässt sich an dieser Stelle für das Neuroblastom eine putative Bindung der *hsa-miR-130b-3p* und des Targets *NDEL1* postulieren, die an Prozesse des Zellwachstums, der Differenzierung sowie der Migration, Invasivität und Metastasierung gekoppelt sind. Auf Basis der *hsa-miR-130b-3p* vermittelte Regulation der Spindelorientierung und des NOTCH-Signalweges kann hier mit hoher Wahrscheinlichkeit von einer naheliegenden Bindung der microRNA zum Target und zu einem Zusammenhang im Tumorkontext ausgegangen werden. Dazu ist in Abb. 3.27 eine schematische Zusammenfassung der hier geschlussfolgerten Interaktionen dargestellt. Die Regulation von *NDEL1* beeinflusst

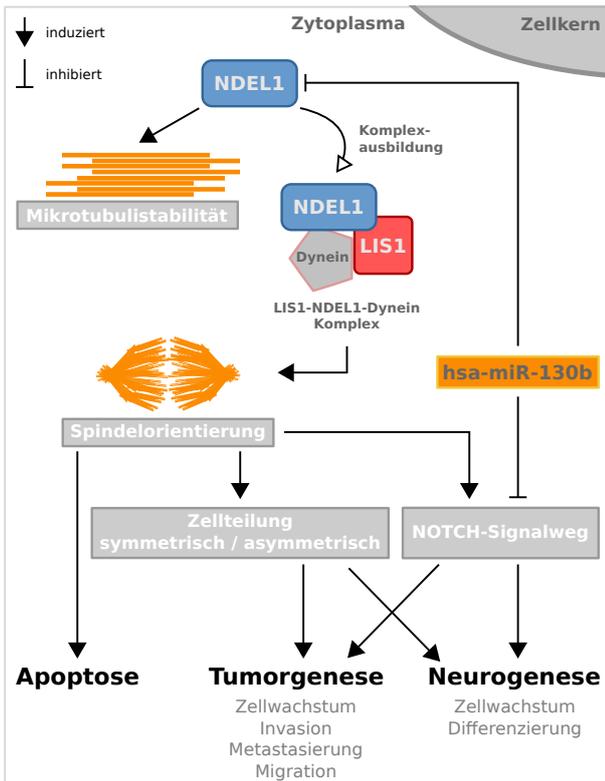


Abbildung 3.27.:

Schematische Darstellung zur möglichen Wirkungsweise der *hsa-miR-130b* im Neuroblastom auf Basis der ermittelten Bindung der *hsa-miR-130b-3p* und der *NDEL1* mRNA. Eine putative Bindung könnte im Neuroblastom zur Beeinflussung der Zellteilung und des NOTCH-Signalweges führen, was Auswirkungen auf die Tumor- und Neurogenese zur Folge hätte. Die microRNA könnte dabei in wesentliche Prozesse der Spindelorientierung sowie davon abhängiger Prozesse eingreifen und somit onkogene Mechanismen, wie Invasivität und Migration, positiv und tumorsuppressive Mechanismen, wie beispielsweise die Differenzierung, negativ regulieren. Dies macht die beiden Bindungspartner zu möglichen therapeutischen Targets für das Neuroblastom und sollte mithilfe von weiterführenden Experimenten genauer untersucht werden.

hierbei die Bildung des LIS1-NDEL1-Dynein Komplexes, was direkte Auswirkungen auf die Spindelfaserorientierung hat und somit die Zellteilung als auch den NOTCH-Signalweg beeinflusst. Durch die Regulation von Liganden des Signalweges kann zudem eine Beeinflussung weiterer Komponenten der Zell-Zell-Kommunikation stattfinden. Über beide Wege könnten im Neuroblastom somit anhand der *hsa-miR-130b-3p* vermittelten Regulation Prozesse des Wachstums, der Differenzierung sowie der Tumorgenese der neuronalen Zellen beeinflusst werden. Dabei sind für die Tumorgenese besonders die Migration, Invasivität und Metastasierung zu nennen. Unter Betrachtung der differentiellen Expression aus dem Vergleich $4^{+M} - 4S$, indem diese putativ neue Bindung ermittelt wurde, kann für das Stadium 4S eine verringerte Expression der *hsa-miR-130b-3p* und eine Erhöhung der *NDEL1* mRNA Expression gezeigt werden. Dies würde nach dem dargestellten Schema zu einer positiven Regulation des LIS1-NDEL1-Dynein Komplexes und damit zu keiner Beeinträchtigung der Spindelorientierung führen. Damit würde der NOTCH-Signalweg weder direkt noch indirekt negativ beeinflusst, so dass Prozesse der symmetrischen/asymmetrischen Zellteilung und Zell-Zell-Interaktion sich positiv auf Wachstum und Differenzierung auswirken und das unkontrollierte Wachstum undifferenzierter Zellen gehemmt wird. Was zu einer Verringerung der Invasivität und Metastasierung sowie zu einer

Verbesserung der Prognose und Überlebenswahrscheinlichkeit führen würde. Dies wiederum entspräche bekannten Charakteristiken des Stadiums 4S, wie beispielsweise die Abnahme der metastatischen Eigenschaften, eine geringere Aggressivität, Ausdifferenzierung von Zellen und eine verbesserte Prognose mit möglicher Regression des Tumors (vgl. hierzu auch 1.2). Die *hsa-miR-130b-3p* und das Target *NDEL1* könnten somit als therapeutische Targets im Neuroblastom dienen und sollten in weiterführenden Analysen hinsichtlich ihrer direkten Bindung und der Assoziation sowie Funktion im Neuroblastom auf experimentellem Wege tiefergehend untersucht werden.

Im sich anschließenden Vergleich $4^{-M} - 4S$ wurden drei putativ neue Bindungen ermittelt, welche auf die Regulation von *MATN2* durch die *hsa-miR-1246* sowie von *CREB5* durch die beiden microRNAs *hsa-miR-1246* und *hsa-miR-3609* hindeuten. Für die Untersuchung der Regulationen durch die *hsa-miR-1246*, wurde die microRNA bereits im Vergleich $1/2/3 - 4/4S$ genauer betrachtet und lässt demnach eine Assoziation zur Metastasierung im Neuroblastom erkennen. Das im Vergleich $4^{-M} - 4S$ ermittelte Target *MATN2* (*Matrilin 2*) ist in der Literatur kaum mit Tumorerkrankungen und nur indirekt mit dem Neuroblastom assoziiert. Es wird als positiv reguliertes Target von *GLI1* beschrieben, einem Transkriptionsfaktor, der im frühen Tumorstadium des Neuroblastoms präsent ist, eine tumorsuppressive Funktion ausübt und eine günstige Prognose bedingt [190]. Die beobachtete Runterregulation der *hsa-miR-1246* und die Hochregulation von *MATN2* im Stadium 4S könnten somit auf eine verbesserte Prognose hindeuten, verglichen mit dem aggressiveren Stadium 4. Eine positive Regulation von *MATN2* im Stadium 4S könnte, bezüglich der *GLI1* Interaktion, somit eine Hemmung der Tumorentwicklung und Metastasierung andeuten, was mit der geringeren Aggressivität und der Möglichkeit zur spontanen Regression in diesem Stadium einher ginge (vgl. 1.2). Aus funktioneller Sicht würde die hier ermittelte Bindung der *hsa-miR-1246* mit der *MATN2* mRNA mit den Befunden der Analyse und der Literatur konform gehen. Ergänzend dazu wird die Bindung von fünf unterschiedlichen Datenbanken vorhergesagt und die Expression beider Partner zeigen innerhalb der hier analysierten Daten eine verbesserte Überlebenswahrscheinlichkeit, auch wenn *MATN2* keine signifikante Veränderung aufweist (vgl. Abb. 3.28). Tendenziell lässt sich jedoch postulieren, dass die *hsa-miR-1246* eine regulatorische Wirkung auf *MATN2* im Neuroblastom ausüben könnte und die in dieser Arbeit ermittelte microRNA als auch das Target im Zusammenhang mit der Metastasierung, Migration und Invasivität im Neuroblastom stehen. Weiterführende Untersuchungen und Experimente hinsichtlich der Funktion und biologischen

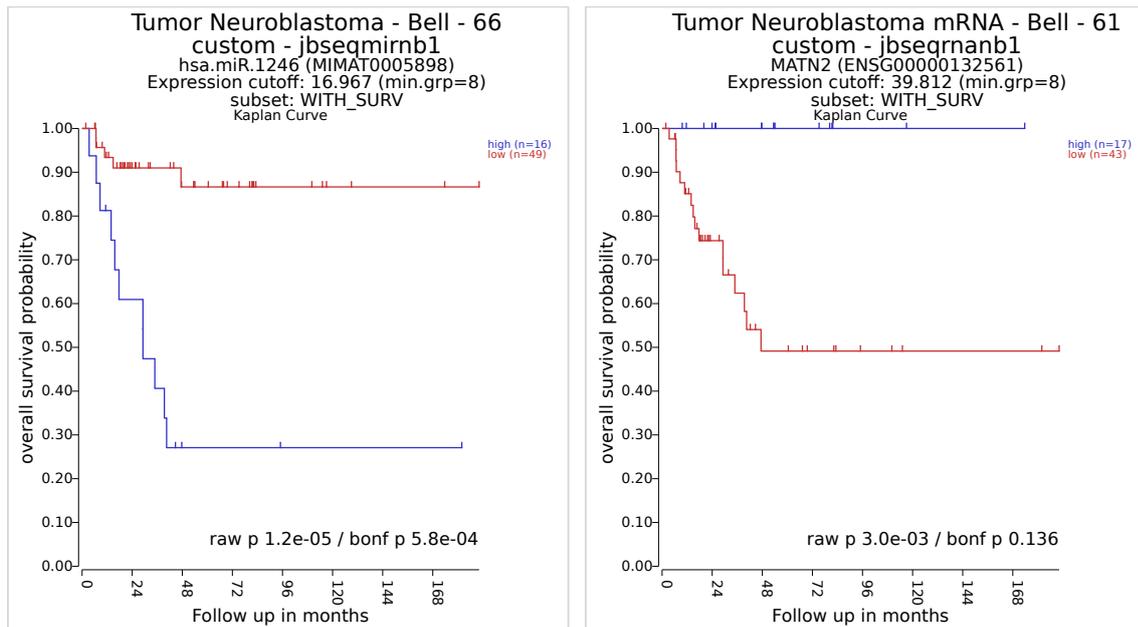


Abbildung 3.28.: Kaplan-Meier Kurven für die Partner der ermittelte Bindung *hsa-miR-1246:MATN2*, die eine putativ neue Beziehung im Zusammenhang mit dem Neuroblastom darstellen könnten. Die Überlebenszeitanalyse lässt eine verbesserte Überlebenswahrscheinlichkeit bei geringer microRNA (rote Kurve, links) und erhöhter *MATN2* Expression (blaue Kurve, rechts) erkennen.

Relevanz auf die Aggressivität im Neuroblastom und die Stellung im Stadium 4S sollten hier in Betracht gezogen werden.

Ein weiteres putativ neues Target aus dem Vergleich $4^{-M} - 4S$ stellt *CREB5* (*cAMP Responsive Element Binding Protein 5*) dar. Hierzu wurden zwei Bindungen zu den microRNAs *hsa-miR-1246* und *hsa-miR-3609* identifiziert. Bei Betrachtung des Vergleiches $1/2/3 - 4/4S$ wurde die *hsa-miR-1246* bereits im Zusammenhang mit [275] erwähnt, wobei in der Publikation auch *CREB5* als Target der *hsa-miR-1246* microRNA genannt wird. Die Vorhersage in diesem Papier erfolgte mittels TargetScan erfolge, eine der Datenbanken, die auch in der hier entwickelten Targetanalyse verwendet wurde. Über *CREB5* lassen sich darüber hinaus nur sehr wenige Nachweise über dessen Funktion in Tumorerkrankungen aus der Literatur entnehmen. Eine der Veröffentlichungen deutet jedoch auf eine wesentliche Stellung in der Metastasierung des kolorektalen Karzinoms hin [200]⁵⁷. Einhergehend mit den Befunden der *hsa-miR-1246* zeichnet sich auf funktionaler Ebene eine Regulierung von Prozessen der Metastasierung durch die microRNA und das Target *CREB5* ab. Die Ergebnisse der Targetanalyse deuten auf eine Beeinflussung der Metastasierung und Migration im Stadium 4S hin, wobei in den analysierten

⁵⁷Nur englische Zusammenfassung des Artikels verfügbar.

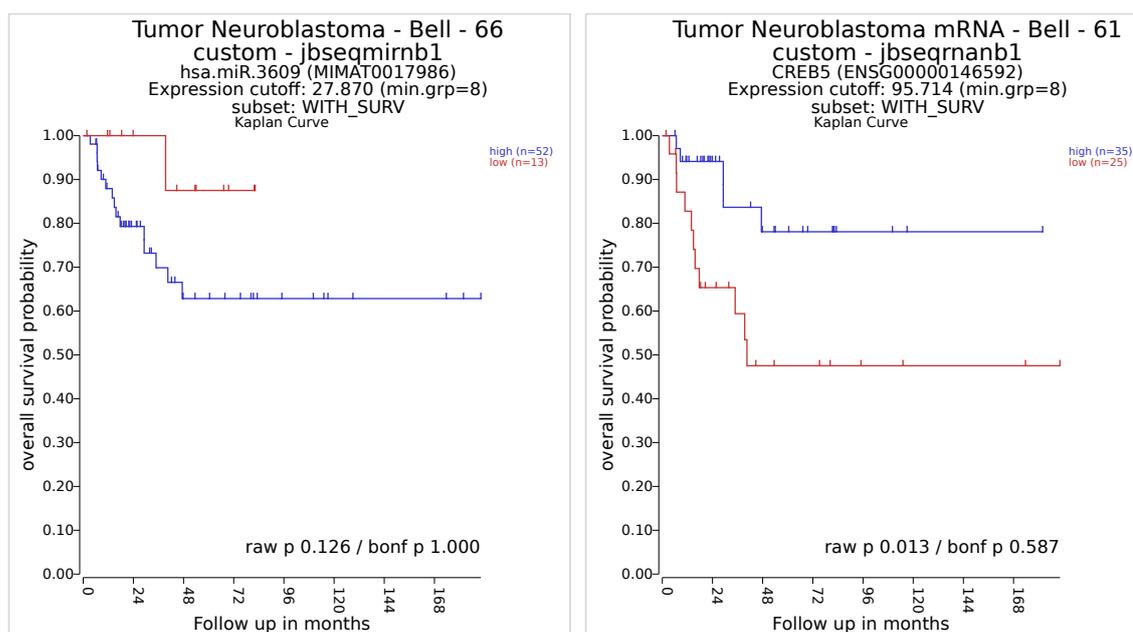


Abbildung 3.29.: Kaplan-Meier Kurven für die Partner der ermittelte Bindung *hsa-miR-3609:CREB5*, die eine putativ neue Beziehung im Zusammenhang mit dem Neuroblastom darstellen könnten. Die Überlebenszeitanalyse lässt eine verbesserte Überlebenswahrscheinlichkeit bei geringer microRNA (rote Kurve, links) und erhöhter *CREB5* Expression (blaue Kurve, rechts) erkennen.

Daten des Vergleiches $4^{-M} - 4S$ die microRNA runter- und *CREB5* hochreguliert werden. Die Überlebenszeitanalyse zeigt für *CREB5* keine signifikante Veränderung, dennoch deutet sich an, dass ein Rückgang der *hsa-miR-1246* Expression und eine Erhöhung der *CREB5* Expression zu einer erhöhten Überlebenswahrscheinlichkeit im Neuroblastom führen (vgl. Abb. 3.28 (links, für *hsa-miR-1246*) und 3.29 (rechts, für *CREB5*)). Somit könnten sowohl die *hsa-miR-1246* als auch *CREB5* eine Assoziation mit der Metastasierung im Neuroblastom aufweisen und zusätzlich zu der Publikation von [275] deutet sich die hier ermittelte Bindung als putativ neue und möglicherweise biologisch relevante Bindung im Neuroblastom an. Dazu sollten weiterführende Experimente folgen, die die Bindung und den Bezug zur Metastasierung im Neuroblastom, und den möglichen Zusammenhang zum Stadium 4S, genauer untersuchen. Auch für die microRNA *hsa-miR-3609* wurde *CREB5* von vier Datenbanken als möglicher Bindungspartner vorhergesagt. Die Überlebenszeitanalyse zeigt weder für die microRNA noch für das Target eine signifikante Veränderung. Dennoch deutet eine entgegengesetzte Expression einen möglichen direkten Zusammenhang zwischen den Transkripten an, wobei eine hohe Expression der microRNA bei gleichzeitig niedriger Expression des Targets zu einer Verschlechterung der Prognose durch eine reduzierte Überlebenswahrscheinlichkeit führt (vgl. Abb. 3.29). Die

hsa-miR-3609 ist in der Literatur nicht mit dem Neuroblastom assoziiert. Darüber hinaus existiert nur eine Publikation, in der die microRNA als putativ neu im Zusammenhang mit Gebärmutterhalskrebs aufgelistet wird [269]. Somit ist es an dieser Stelle äußerst schwierig schlussfolgernde Aussagen für die biologische Relevanz der microRNA im Neuroblastom oder für die Bindung und Regulierung des Targets *CREB5* zu treffen. Hierzu sind definitiv weitere Experimente von Nöten, die in einem ersten Schritt die mögliche Bedeutung der *hsa-miR-3609* im Kontext mit dem Neuroblastom klären sollten, um darauf basierend eventuell weitere Analysen durchzuführen.

Zusammenfassend betrachtet weisen drei der sieben ermittelten microRNA vermittelten Regulationen eine verstärkte biologische Relevanz für das Neuroblastom auf und sollten weiterführend, vor allem aber durch experimentelle Versuche gestützt, untersucht werden. Diesen Regulationen, von denen weder die microRNA noch das Target sich direkt mit dem Neuroblastom assoziieren lassen, sind gegeben durch die Bindungen *hsa-miR-744-5p:HHURP*, *hsa-miR-130b-3p:NDEL1* und *hsa-miR-1246:MATN2*. Anhand der Ergebnisse der Targetanalyse konnten diese Bindungen von mehreren Datenbanken vorhergesagt werden, weisen signifikante Veränderungen in der Überlebenszeitanalyse auf und deuten zudem durch Befunde der aktuellen Literatur bezüglich anderer Krebserkrankungen auf ihre Relevanz im Neuroblastom hin. Alle drei Verbindungen stammen aus Vergleichen des Stadiums 4S mit aggressiven und risikoreichen Stadien. Das Stadium 4S zeichnet sich durch eine geringere Aggressivität, eine bessere Prognose sowie durch die Möglichkeit zur spontanen Regression aus und ist bisher nur unzureichend untersucht (vgl. 1.2). Die Befunde der Targetanalyse zeigen des Weiteren - bereits innerhalb der Bestenliste in Tab. 3.15 - für dieses Stadium zahlreiche putativ neuen Transkripte und Verbindungen auf, von denen einer der Bindungspartner nicht direkt mit dem Neuroblastom assoziiert ist. Weitere Recherchen hierzu haben gezeigt, dass diverse Targets im Stadium 4S mit Keratinen (KRT1, KRT10) oder Desmosomen (DSC3, PKP1) assoziiert sind, Proteine der Motilität, Zellgröße, Zelladhäsion und des Zell-Zell-Kontaktes. Wodurch sich anhand der ermittelten Targets des Stadiums ein beschriebener Zusammenhang zur Epithelial-mesenchymalen Transition (EMT) andeutet [60, 105, 120]. Dieser Prozess beschreibt den Übergang von Zellen, aus Epithelien oder Karzinomen, zu Zellen mit mesenchymalen Eigenschaften und nimmt eine wesentliche Rolle bei der Entwicklung, Wundheilung und der Metastasierung von Krebszellen ein [86, 120]. Dabei werden genetische Prozesse wie Differenzierung und Motilität aktiviert und Prozesse wie Zell-Zell-Kontakt oder Zelladhäsion

runterreguliert. Mithilfe des EMT Prozesses ist eine metastasierende Krebszelle in der Lage vom Entstehungsort des Tumors in andere Gewebe oder Organe zu migrieren und neue Tumore zu formen. Um sich an diesen neuen Ort festzusetzen, setzt ein umgekehrter Prozess ein, die Mesenchymal-epitheliale Transition (MET), der die Aktivierung oder Inaktivierung der Prozesse umkehrt. Beide Prozesse sind eine Grundvoraussetzung für eine erfolgreiche Metastasierung des Tumors [86]. Durch die Hochregulierung der ermittelten Targets kann aufgrund der Umkehrung der Expression ein entgegengesetzter Effekt und somit ein Bezug zum MET angenommen werden [84, 86, 120]. Da diese Targets nicht nur bei Vergleichen zu den aggressiven und risikoreichen Tumoren auftreten (siehe Vergleich 1 – 4S), könnten diese Befunde einen relevanten Ausgangspunkt für weitere Untersuchungen im Bezug auf die Möglichkeit zur spontanen Regression im Stadium 4S liefern. Möglicherweise lassen sich die Targets als biologisch signifikant für das Stadium 4S bestätigen und können für therapeutische Maßnahmen in anderen Stadien zur Eindämmung der Matastasierung und Aggressivität und somit zu einer Verbesserung des Krankheitsverlaufes genutzt werden.

Wesentlich eingehender untersuchte Tumore wie die des Stadium 4 oder Tumore mit *MYCN* Amplifikation, die für eine sehr aggressive und risikoreiche Erkrankung stehen, liefern bei der Analyse in der vorliegenden Arbeit hingegen nur wenig neue microRNAs, Targets oder Bindungen. Grund dafür ist die stark veränderte Expression einer Vielzahl an Transkripten, die vor allem in amplifizierten Tumoren auftritt und bereits in zahlreichen Analysen primärer Proben oder Zelllinien nachgewiesen werden konnten (vgl. hierzu Abschnitt 3.2.2). Somit existieren für diese Stadien, mit sehr schlechte Prognose und geringen Therapiemöglichkeiten, bereits diverse Nachweise in der Literatur, die neue relevante microRNAs oder Targets im Neuroblastom identifiziert oder experimentell validiert haben. Dennoch können auch in diese Stadien mittels der entwickelten Targetanalyse putativ neue microRNAs oder Targets ermittelt werden, von denen einer der Bindungspartner nicht direkt mit dem Neuroblastom assoziiert ist und die eine mögliche Relevanz zur Erkrankung erkennen lassen. Dazu gehören beispielsweise die microRNAs *hsa-miR-23b-3p*, *-27a-3p* und *-301a-3p* des Vergleiches 1 – 4^{-M} sowie einige mRNAs der Vergleiche 1 – 4 , 1 – 4^{+M} und 1 – 4^{-M} , wobei hier nur die Bestenliste der Targetanalyse betrachtet sei (vgl. Tab. 3.15). Die Gesamtlisten aller ermittelten microRNA-Target Beziehungen lässt vermutlich weitere putativ neue Zusammenhänge und Erkenntnisse zu (vgl. Verweis im Anhang F.2), deren Analysen jedoch den Rahmen dieser Arbeit überschreiten würden. Die hier erzielten Befunde liefern nichtsdestoweniger interessante Ansätze für weiterführende

Untersuchungen von Transkripten, die für das Stadium 4S aber auch für bereits gut untersuchte Stadien des Neuroblastoms relevant sein könnten und die Funktionen oder Mechanismen regulieren, die für diagnostische bzw. therapeutische Maßnahmen von Belang sein könnten.

microRNA Biogenese und Aktivität

Veränderungen in der Biogenese und Aktivität von microRNAs haben einen verstärkten Einfluss auf die Regulation der zahlreichen microRNA-Targets mittels RNA-Stillegung (*RNA-Silencing*) und somit auf die daran gekoppelten Prozesse. In Krebserkrankungen ist dabei zu beobachten, dass eine globale Reduktion von microRNAs durch eine Deregulation der Biogenese-Faktoren stattfindet, was dem Prozess der Deregulation an sich eine große Bedeutung innerhalb der Krebsforschung zukommen lässt [135]. Die Biogenese vom Transkript zur reifen microRNA unterläuft dabei einen mehrstufigen Reifungsprozess, der nicht für alle microRNAs identisch, jedoch für die Entstehung einer reifen microRNA entscheidend ist [135, 214, 268]. Durch diverse Mechanismen oder auf posttranskriptionellem Wege ist es möglich diesen Prozess zu regulieren und zu beeinflussen, indem Stufen des Prozesses umgangen, aufgehoben oder ersetzt werden. Als Folge davon wird die Produktion, Stabilität und Reifung der microRNAs beeinflusst, was möglicherweise die veränderte Expression in diversen Krebserkrankungen, bei einer globalen Reduktion reifer microRNAs, erklärt. Die Deregulation der microRNA Biogenese legt aktuell vielmals die Grundlage für die Hypothese der Krebsentstehung durch Veränderungen in den Biogenese-Faktoren und des Reifungsprozesses, jedoch fehlen hierfür einschlägige experimentelle Nachweise [214, 268]. Durch aktuelle Forschungen ist jedoch bekannt, dass die Regulation auf verschiedenen Ebenen der Transkription, Prozessierung durch DROSHA und DICER1, durch Argonaut-Modifizierungen oder beim RNA-Abbau erfolgt. Dabei finden unterschiedlichste Strategien der Einflussnahme Anwendung, wie beispielsweise die Rekrutierung von Transkriptionsfaktoren, RNA-bindenden Proteinen oder Protein- und RNA-modifizierenden Enzymen [81]. In vielen bereits untersuchten Krebserkrankungen regulieren diverse Transkriptionsfaktor (TF) die microRNA Transkription. Nachgewiesenermaßen zeigen einige TF ein tumorsuppressives oder onkogenes Potential, welches die Transkription verschiedener microRNAs fördern oder hemmen kann. Als Beispiel sei hier MYC als TF genannt, der die Transkription onkogener microRNAs, wie die Mitglieder des 17~92 Clusters, fördert [214]. Darüber hinaus gibt es eine Vielzahl bekannter TF, die die microRNA Expression bereits in einem der ersten Reifungsschritte positiv oder negativ beeinflussen, wie

beispielsweise p53, ZEB1, ZEB2 oder MYOD1 [81]. Zudem können Änderungen in der RNA Sequenz und/oder Struktur die Reifung und Aktivität von microRNAs beeinflussen. Oftmals sind microRNAs aber auch an sogenannte *Feedback Loops* gekoppelt, bei denen die microRNA-Targets wiederum einen Einfluss auf die Produktion der regulierenden microRNAs haben und somit deren Regulation verstärken oder schwächen können. Des Weiteren können Faktoren der microRNA Biogenese auch in anderen RNA Stoffwechselwegen beteiligt sein und sich auf die Stabilität und den Abbau von mRNAs auswirken [81]. Genauere Untersuchung bezüglich der bekannten Faktoren der Biogenese als auch der Aktivität können somit Hinweise offenlegen, ob und in welchem Schritt des Reifungsprozesses oder der mRNA-Regulation Veränderungen auftreten. Das Verständnis über den Ablauf der microRNA Biogenese ist entscheidend, um die Rolle dieses Prozesses in der Tumorgenese zu verstehen und um zukünftige Anwendungen entwickeln zu können, die auf die Diagnose, Risikoeinschätzung und auf therapeutische Strategien der Krebsbehandlung abzielen [135, 214]. Die in dieser Arbeit analysierten Daten bieten dazu eine ideale Grundlage, da das Material der Sequenzierungen aus den selben Proben entnommen wurde und Veränderungen der Expression aus Basis der microRNA als auch mRNA Expression gleichermaßen analysiert werden konnten.

Betrachtet man die differentielle Expression und die Anzahl ermittelter microRNA-Target Beziehungen lässt sich erkennen, dass die aggressiven und risikoreichen Stadien aus den Vergleichen eine erhöhte Anzahl an veränderten microRNAs im Neuroblastom zeigen (vgl. hierzu Tab. 3.2 und Abb. 3.19). Im Detail weisen dabei die Vergleiche 1 – MYCN und 1 – 4^{+M} die höchste Anzahl an veränderten microRNAs auf, wobei der Großteil der microRNAs in diesen Vergleichen - entgegen der gängigen Annahme in Krebserkrankungen [135] - signifikant hochreguliert werden. Alle weiteren Vergleiche, mit nicht *MYCN* amplifizierten Stadien, zeigen die erwartete Runterregulation der microRNAs. Die Ergebnisse der Targetanalyse zeigen zudem, dass aus den ermittelten microRNA-Target Beziehungen eine geringfügig größerer Teil herunterreguliert wird oder zumindest gleich viele microRNAs in ihrer Expression nach oben oder unten verändert sind.

Zusammengenommen spiegelt sich anhand der analysierten Daten in dieser Arbeit der bereits erwähnte globale Trend zur Runterregulation von microRNAs auch im Neuroblastom wider. Es ist jedoch deutlich erkennbar, dass die *MYCN* Amplifikation einiger Tumore einen weitreichenden Einfluss auf die verstärkte Expression von microRNAs in Stadien dieser Tumore zur Folge hat. Betrachtet man dazu nun wesentliche Faktoren der microRNA Biogenese und Aktivität

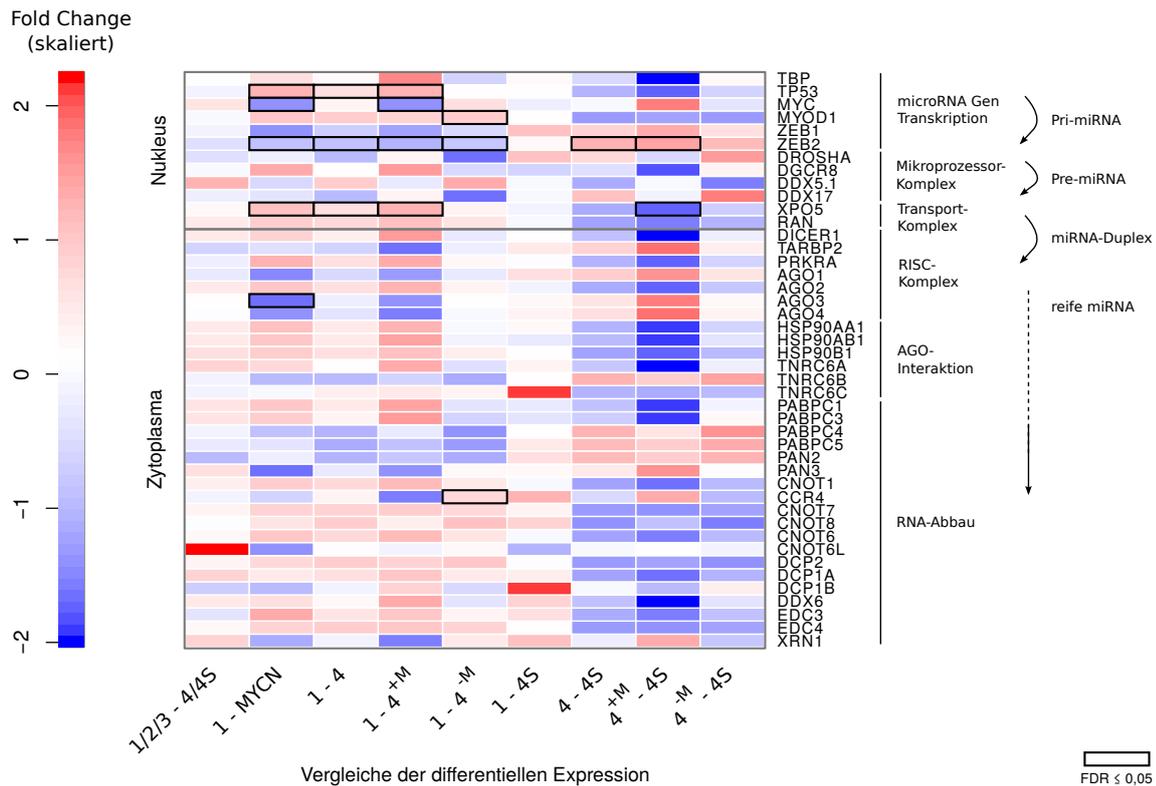


Abbildung 3.30.: Heatmap der differentiellen Expression von Faktoren, die im Wesentlichen an der microRNA Biogenese und Aktivität beteiligt sind. Dargestellt sind die Veränderungen der Faktoren aus den differentiellen Vergleichen der Neuroblastomstadien durch die skalierten logFC-Werte. Die Faktoren sind nach ihrer Lokalisation in der Zelle sowie nach ihrer Wirkungsreihenfolge in der Biogenese und Aktivität geordnet. Unveränderte Faktoren sind weiß, hochregulierte rot und herunterregulierte Faktoren sind blau dargestellt. Signifikante Veränderung eines Faktors ($FDR \leq 0,05$) sind mit einem schwarzen Rechteck markiert. Die Skalierung der logFC-Werte fand zeilenweise für jeden Faktor statt, so dass ein Mittelwert von 0 und eine Standardabweichung von 1 erreicht wurde.

fällt zunächst auf, dass nur wenig signifikante Veränderungen in den differentiellen Vergleichen vorliegen, jedoch lassen sich neben einigen signifikant veränderte Faktoren auch klare Tendenzen für eine Regulierung erkennen. Die hierbei untersuchten Faktoren sind nach ihrer Lokalisation in der Zelle sowie nach ihrer Wirkungsreihenfolge in der microRNA Reifung und Aktivität angeordnet (vgl. Abb. 3.30). Die signifikantesten Veränderungen zeigen Faktoren die an der nukleären Prozessierung beteiligt sind, wobei die davon am häufigsten auftretenden aus *MYCN* amplifizierten Stadien des Vergleiches 1 – MYCN stammen.

Veränderungen in der microRNA Biogenese Die Transkription und Prozessierung des primären microRNA-Transkripts sowie der anschließende Export in das Zytoplasma bilden die ersten Schritte der microRNA Biogenese [81,268]. Dabei wird die Transkription im Wesentlichen

von der RNA-Polymerase II (Pol-II) durchgeführt. Da Pol-II neben den microRNAs auch diverse andere RNAs prozessiert, sollen an dieser Stelle Transkriptionsfaktoren (TF) betrachtet werden, die bekanntermaßen an der Transkriptions-Regulation von microRNAs beteiligt sind und die Prozessierung zum primären microRNA-Transkript (Pri-miRNA) beeinflussen können. Zu diesen Faktoren gehören TP53, MYC, MYOD1, ZEB1 und ZEB2 [81, 135] sowie TBP, einem TF mit wesentlicher Bedeutung bei der Initiation des Transkriptionskomplexes der Pol-II [288]. Signifikante Veränderungen können anhand der analysierten Daten in dieser Arbeit für die Faktoren TP53, MYC und ZEB2 beobachtet werden, vor allem in *MYCN* amplifizierten Tumoren der Vergleiche 1 – *MYCN* und 1 – 4^{+M} (vgl. hierzu Abb. 3.30). TBP als Initiationsfaktor der Pol-II zeigt keine signifikanten Veränderungen, dennoch zeigt sich eine Hochregulation unter *MYCN* Amplifikation, was für eine erhöhte Aktivität der Pol-II sprechen könnte. Unter den Faktoren der Transkriptions-Regulation von microRNAs werden TP53 und MYOD1 in den aggressiven Tumoren hochreguliert, wobei bei TP53 eine Verstärkung unter *MYCN* Amplifikation auftritt. Beiden TF haben Einfluss auf die tumorsuppressiven microRNA Cluster miR-1 und miR-34 und verstärken die Pol-II vermittelte Transkription der microRNAs [81, 135]. ZEB1 und ZEB2 zeigen eine Runterregulation in den aggressiven Tumoren, MYC wird hingegen nur in *MYCN* amplifizierten Tumoren runterreguliert. Während womöglich ZEB1 und ZEB2 eine negative Regulation auf die microRNA Transkription aufweisen, können durch MYC respektive *MYCN* diverse microRNAs positiv oder negativ reguliert werden [81, 135]. Für MYC ist jedoch bekannt, dass es gewebespezifisch exprimiert wird und es vor allem im neuronalen Gewebe zu einer vorwiegenden Expression des zur MYC-Familie gehörenden onkogenen Transkriptionsfaktors *MYCN* kommt [92]. Zusammenfassend bedeutet dies eine Veränderung wesentlicher Faktoren der Transkription dahingehend, dass eine positive Regulation des ersten Reifungsschrittes durch eine erhöhte Expression der Faktoren in den aggressiven, vor allem aber in *MYCN* amplifizierten, Tumoren stattfindet. Dies könnte tendenziell auf eine erhöhte microRNA Expression durch eine verstärkte Transkription hindeuten und infolgedessen auf eine verstärkte Aktivität der microRNA vermittelte RNA-Stilllegung andeuten.

Im Anschluss der Transkription kommt es zur nukleären Prozessierung der gebildeten Pri-miRNA durch den Mikroprozessor-Komplex, welcher sich aus den Faktoren DROSHA und DGCR8 zusammensetzt und dessen Aktivität durch die interagierenden RNA-Helikasen DDX5 und DDX17 verstärkt werden kann [81, 214]. Dieser Reifungsschritt dient der Prozessierung der Pri-miRNA zur Vorläufer microRNA mit ihrer charakteristische Haarnadelstruktur, auch

als *Precursor Hairpin* oder Pre-miRNA bezeichnet. DROSHA nimmt dabei eine zentrale Rolle bei der Abspaltung des 5' und 3' Armes der Pri-miRNA ein, wobei DGCR8 für eine stabile Interaktion mit der Pri-miRNA und für die präzise Abspaltung der Arme sorgt [268]. Anhand der Abb. 3.30 ist zu erkennen, dass die differentielle Expression beider Faktoren vor allem in *MYCN* amplifizierten Stadien eine bei DROSHA leichte und bei DGCR8 stärkere Hochregulierung aufzeigen. Ansonsten scheinen beide Faktoren in den aggressiven Tumoren einer eher runterregulierten oder kaum veränderten Expression aufzuweisen. Auffällig dabei ist, dass beim Vergleich von Stadium 4S mit Stadium 1 und 4 (Vergleiche 1 – 4S und 4 – 4S) DROSHA hoch- und DGCR8 runterreguliert wird. Hingegen zeigen beide Faktoren in 4^{+M} – 4S eine gemeinsame Runterregulation, während sie im Vergleich 4^{-M} – 4S eine Hochregulation aufweisen. Dies könnte darauf hinweisen, dass die Expression von DROSHA und DGCR8 durch die *MYCN* Amplifikation beeinflusst wird. Besonders zeigt sich dies in Vergleichen zu Stadien mit gemischten Proben, z. B. 1 – 4 oder 4 – 4S, die zu einer unterschiedlichen Veränderung der Faktoren führen. Eindeutig unterteilte Stadien in *MYCN* amplifiziert oder nicht amplifiziert Tumore führen hingegen zu einer einheitlichen Veränderung. Dabei unterliegen beide Faktoren im Stadium 4S einer eher positiven Regulation, verglichen mit Stadium 4S ohne *MYCN* Amplifikation. Ist eine Amplifikation vorhanden, scheint dies in den amplifizierten Tumoren zu einer höheren Expression als im Stadium 4S zu führen (vgl. 4^{+M} – 4S). Alle Veränderungen der Faktoren sind jedoch nicht signifikant, so dass keine eindeutige Aussage anhand der unterschiedlichen Expressionsdaten getroffen werden kann. Tendenziell zeichnet sich aber ab, dass eine *MYCN* Amplifikation in den Tumoren zu einer verstärkten Expression der mRNA der Faktoren DROSHA und DGCR8 führt. Dies wiederum könnte womöglich unter *MYCN* Amplifikation zu einer verstärkten Pri-miRNA Prozessierung führen, was die erhöhte Anzahl an veränderten microRNAs in den Vergleichen mit Proben einer Amplifikation erklären kann und was demnach zum verstärkten RNA-Silencing in den aggressiven und risikoreichen Tumoren führen könnte. Die genannten Helikasen DDX5 und DDX17 zeigen untereinander eine entgegengesetzte und sehr schwache Veränderung. Die stärkste Veränderung zeigen beide im Vergleich 4^{-M} – 4S, in der DROSHA und DGCR8 ebenfalls eine klare Hochregulation zeigen. Dies weist ebenfalls darauf hin, dass im Stadium 4S eine verstärkte microRNA Prozessierung stattfindet, verglichen mit aggressiven Tumoren ohne *MYCN* Amplifikation.

Nach Bildung der nukleären Pre-miRNA findet ihr Export in das Zytoplasma durch den Transport-Komplex statt, welcher aus den Faktoren XPO5 und RAN gebildet wird [81, 135,

214, 268]. Beide Faktoren sind eindeutig in den aggressiven Tumoren hochreguliert, wie die ersten fünf differentiellen Vergleiche in Abb. 3.30 zeigen. Der Vergleich 1 – 4S zeigt, dass die Faktoren zwischen den Stadien 1 und 4S weitestgehend unverändert sind. Die Umkehrung der Expression in den Vergleichen von 4 und 4S zeigen, dass die Faktoren in den weniger aggressiven und risikoärmeren Stadien runterreguliert werden. Zudem zeigt XPO5 teilweise signifikante Veränderungen, so dass sich für diesen Reifungsschritt ebenfalls eine positive Regulierung der microRNA Biogenese durch einen tendenziell verstärkten Transport der Pre-miRNA ins Zytoplasma andeutet.

Im Anschluss an den nukleären Export findet die zytoplasmatische Prozessierung der Pre-miRNA durch DICER1 und die Ausbildung des RISC-Komplexes (*RNA-induced silencing complex*) statt. DICER1 spaltet dabei die terminale Schleifenstruktur der Pre-miRNA ab und bildet somit die ca. 22 Nukleotid lange reife Duplex-miRNA. Für die DICER1 vermittelte Prozessierung sind zwei begünstigende Faktoren zu nennen, zum einen TARBP2 und zum anderen PRKRA. Beide Faktoren sind nicht zwingend für die Prozessierung notwendig, haben jedoch Auswirkung auf die Stabilität und Aktivität von DICER1 [81, 214, 268]. Die *DICER1* mRNA zeigt in den analysierten Daten eine reduzierte Expression im Stadium 4 ohne *MYCN* Amplifikation (Vergleich 1 – 4^{-M}) sowie in allen Vergleichen von Stadium 4 Tumoren mit Stadium 4S. Vor allem in *MYCN* amplifizierten Stadien ist eine verstärkte Hochregulation von DICER1 erkennbar, die stärkste Runterregulation zeigt sich im Vergleich 4^{-M} – 4S. Die beiden Cofaktoren zeigen untereinander eine entgegengesetzte Expression. Während PRKRA sich identisch zur Expression von DICER1 verhält, zeigt TARBP2 eine Runterregulation in den aggressiven Tumoren mit stärkster Veränderung im 1 – 4^{+M}. Es ist bekannt, dass eine Veränderung von TARBP2 zur Destabilisierung von DICER1 führen kann, was negative Auswirkungen auf die microRNA Expression nach sich zieht [81, 214, 268]. Tendenziell könnte jedoch die erhöhte DICER1 Expression ein verstärkte Prozessierung der Pre-miRNA, vor allem bei *MYCN* amplifizierten Tumoren, des Neuroblastoms zur Folge haben. Dies geht auch mit den bisherigen Befunden der vorliegenden Arbeit einher, dass in amplifizierten Tumoren mehr microRNAs Prozessiert werden könnten und möglicherweise ein verstärkter Export in das Zytoplasma stattfindet.

Die durch DICER1 generierte Duplex-miRNA wird in ATP-Abhängigkeit zur reifen microRNA entwunden und es findet unterstützt durch das Hitzeschockprotein HSP90 die Beladung eines der Argonautenproteine (AGO1-4) mit einem der microRNA Stränge (*Functional Guide Strand*)

statt. Durch diesen Prozess wird, mittels reifer microRNA, DICER1, TARBP2, PRKRA und AGO, der RISC-Komplex ausgebildet [81, 135, 158, 214]. Beim Menschen ist dabei im Wesentlichen AGO2 katalytisch aktiv, während die Funktion der anderen Argonautenproteine noch ungeklärt sind [100,158]. Betrachtet man, neben den bereits untersuchten Komponenten des RISC-Komplexes, wie DICER1, TARBP2 und PRKRA, die AGO-Proteine und die Varianten der Hitzeschockprotein so ist erkennbar, dass AGO2 in den aggressiven Tumoren eine Hochregulation der Expression aufweist. Im Vergleich Stadium 1 zu 4S ist AGO2 unverändert und zeigt beim Vergleichen Stadium 4S mit aggressiver Tumor eine deutliche Runterregulation. Die anderen, im Menschen nicht katalytisch aktiven AGO-Proteine, weisen in den aggressiven Tumoren eine Herunterregulierung und im Stadium 4S eine erhöhte Expression auf. Die drei Varianten der Hitzeschockprotein zeigen eine vergleichbare Expression zu AGO2. Der gesamte RISC-Komplex, zusammen mit einigen Interaktionspartnern, zeigt demnach in den analysierten Daten dieser Arbeit eine verstärkte Expression in den aggressiven und risikoreichen Tumoren, vor allem in Anwesenheit einer *MYCN* Amplifikation. Somit wird auch anhand der zytoplasmatischen Prozessierung bei der microRNA-Reifung angedeutet, dass eine verstärkte Prozessierung an microRNAs in *MYCN* amplifizierten Tumoren statt finden könnte und dies möglicherweise darauf hindeutet, dass in diesen Tumore eine größere Anzahl an microRNAs gebildet wird. Ergänzend zu den bisherigen Befunden der erhöhten differentiellen Expression, der nukleären Prozessierung und des Transports ins Zytoplasma kann dies somit zu einer möglichen Verstärkung des RNA-Abbaus durch die microRNA vermittelten RNA-Stilllegung führen. Zusammen betrachtet lassen diese Befunde den Schluss zu, dass der Transkriptionsfaktor MYC respektive MYCN zu einer enormen Veränderung der microRNAs Expression führen und dies auf allen Ebenen der microRNA Biogenese zu verfolgen ist. Tumore mit einer hohen *MYCN* Abundanz führen dabei sichtlich zu einer Erhöhung der microRNA Expression, was mögliche Folgen für die microRNA Aktivität und RNA-Stilllegung nach sich ziehen kann, was im nachfolgenden Abschnitt näher betrachtet werden soll.

Veränderungen in der microRNA Aktivität Findet somit eine verstärkte Biogenese von microRNAs in *MYCN* amplifizierten Tumoren des Neuroblastoms statt, so würden dadurch mehr microRNAs gebildet und mehr mRNAs durch Prozesse der RNA-Stilllegung abgebaut werden können. Um die Aktivität der microRNAs diesbezüglich zu untersuchen wurden weitere Faktoren betrachtet, die bei der RNA-Stilllegung eine wesentliche Rolle spielen. Die Regulierung

von mRNAs erfolgt dabei durch die Kombination verschiedener Mechanismen, wie beispielsweise transkriptionelle Hemmung, Deadenylierung und 5' zu 3' Abbau. Der beschriebene RISC-Komplex vermittelt die posttranskriptionelle Stilllegung von mRNAs, die eine partielle oder vollständig komplementäre Sequenz zu der im Komplex enthaltenen microRNA aufweist. Dabei werden bei vollständiger Basenpaarung die mRNAs durch das katalytisch aktive AGO2, partiell komplementäre Sequenzen durch die Rekrutierung zusätzlicher AGO-Interaktionspartner still gelegt [100, 158]. Bei der RNA-Stilllegung spielen sogenannte GW-Proteine eine wesentliche Rolle, im Menschen sind dies insbesondere TNRC6A-C. Diese Proteine interagieren mit den Argonautenproteinen und stellen die Verbindung zwischen den AGOs und Komponenten der Stilllegung her. Zu diesen Komponenten gehören beispielsweise das Poly(A)-bindenden Protein (PABPC) sowie der PAN2-PAN3- und der CCR4-NOT-Komplex, welche die Deadenylierung der Target mRNA katalysieren [100, 158]. Eine zentrale Stellung bei der Deadenylierung nimmt der CCR4-NOT-Komplex ein, der den Abbau des Poly(A)-Schwanzes katalysiert und neben der Deadenylierung an der translationellen Repression und der Degradation der microRNA-Targets beteiligt ist. Der Komplex besteht aus CCR4 sowie diversen unabhängigen Komponenten, die mit CNOT1 in Verbindung stehen, wie beispielsweise CNOT6, CNOT6L, CNOT7 und CNOT8 [100]. Im Anschluss an die Deadenylierung findet der Kappen-Abbau (*Decapping*) der deadenylierten mRNAs durch DCP2 statt, welches für seine vollständige Aktivität weitere Cofaktoren benötigt, dazu zählen DCP1, DDX6, EDC3 und EDC4. Nach der Deadenylierung und Entfernung der Kappen-Struktur findet letztendlich der Abbau der mRNAs durch den zytoplasmatische 5' zu 3' Exoribonuklease XRN1 statt [100].

Die Daten aus Abb. 3.30 zeigen, analog zu den bisherigen Befunden, dass der überwiegende Teil der untersuchten Faktoren des RNA-Abbaus einer Hochregulation in den Tumoren unterliegen. Vor allem in den *MYCN* amplifizierten Stadien ist eine verstärkte Expression zu erkennen. Ausnahmen bilden die Poly(A)-bindenden Proteine PABPC4-5, PAN2 und PAN3 sowie das CCR4 des CCR4-NOT-Komplexes und XRN1. Für die Poly(A)-bindenden Proteine ist nicht genau spezifiziert, welche der Varianten an der microRNA vermittelten Regulierung beteiligt sind bzw. ist nicht bekannt, welche davon im Neuroblastom exprimiert werden. Dies könnte die unterschiedlichen Expression der einzelnen Varianten in den analysierten Daten erklären. Des Weiteren zeigen CCR4 und XRN1, welche wesentliche Aufgaben bei der Stilllegung übernehmen, eine Runterregulationen in den Vergleichen mit *MYCN* amplifizierten Stadien, während jedoch CNOT1 als zweite essentielle Komponente des CCR4-NOT-Komplexes eine durchgängige

Hochregulation aufweist. Zudem ist weiterhin auffällig, dass die Faktoren der Stilllegung beim Vergleich 1 – 4S eine verstärkte Hochregulation aufweisen, während die betrachteten Faktoren der vorherigen Reifungsschritte eine eher geringe oder unveränderte Expression zeigten. Dies legt eine verstärkte Aktivität der microRNA vermittelten RNA-Stilllegung im weniger aggressiven und risikoärmeren Stadium 4S des Neuroblastoms im Vergleich zum Stadium 1 nahe, was einen verstärkten mRNA-Abbau nach sich ziehen könnte. Für das Stadium 4S konnte bereits bei der nukleären microRNA Prozessierung durch DROSHA, DGCR8 und den beteiligten Helikasen eine verstärkte microRNA Prozessierung beobachtet werden. Wodurch sich auch die marginalen Unterschiede in der differentiellen Expression und die daraus resultierenden geringen Mengen an differentiell exprimierten mRNAs aus dem Vergleich 1 – 4S erklären ließen (20 hoch- und 1 herunterregulierte mRNA, vgl. hierzu Tab. 3.2).

Letztendlich können die analysierten Expressionswerte der mRNAs in der vorliegenden Arbeit nur einen Hinweis auf mögliche Veränderungen der RNA-Menge geben. Diese Daten lassen jedoch keinen Rückschluss auf die tatsächlich gebildeten und vorliegenden Proteinmengen in den Tumorproben zu, so dass sich auf RNA- und Proteinebene die Menge der Faktoren in den Tumoren unterscheiden können bzw. unterscheiden werden. Neben den betrachteten RNA-Abundanzen bleiben somit Mutationen und Proteinmengen unbekannt. Zudem bieten RNA-bindenden Proteine, sogenannte RBPs, neben der Biogenese und Modulierung der microRNA vermittelten Regulierung auch die Möglichkeit zum Schutz vor dem mRNA Abbau durch microRNAs. Dabei können RBPs ähnliche oder gar identische Sequenzen der RNA binden, um somit einen entgegengewirkenden Effekt zu erzielen indem sie die Bindung der microRNA und somit die Aktivität und den Abbau durch die microRNA verhindern [99, 139]. Dennoch kann zusammenfassend nach Betrachtung wesentlicher Faktoren der microRNA Biogenese und der microRNA vermittelten RNA-Stilllegung postuliert werden, dass die untersuchten *MYCN* amplifizierten Tumor ein deutlich verändertes Bild der Expression dieser Faktoren aufzeigen. Entgegen der Vergleiche 1 – *MYCN* und 1 – 4^{+M} , welche die aggressiven und risikoreichsten Stadien des Tumors mit dem risikoarmen Stadium 1 vergleichen, zeigt sich in den Stadienvergleichen nicht amplifizierter Proben, wie beispielsweise 1 – 4^{-M} oder 1 – 4S, eine runterregulierte oder unveränderte Expression der Faktoren. Anhand dieser Befunde wird deutlich, dass die in dieser Arbeit untersuchten Neuroblastomproben in *MYCN* amplifizierten Tumoren eine auf mRNA-Ebene betrachtete positive Regulation der nukleären und zytoplasmatischen microRNA Prozessierung aufweisen und sich ein verstärkter Export von

Pri-mRNAs in das Zytoplasma andeutet. Darüber hinaus zeigen große Teile des RISC-Komplexes und untersuchter Faktoren der RNA-Stillegung eine verstärkte Expression innerhalb dieser Tumore. Tendenziell sprechen diese Ergebnisse der Veränderungen in der microRNA Biogenese und Aktivität für eine erhöhte microRNA Expression und der damit einhergehenden Möglichkeit zur verstärkten microRNA vermittelten RNA-Stillegung in den aggressiven, vor allem aber in *MYCN* amplifizierten, Tumoren.

4. Diskussion

Analyse neuronaler Zellmorphologie und Proteinverteilungen

Im Rahmen der hier vorliegenden Arbeit wurde eine Applikation zur automatischen Auswertung der Zellmorphologie und Proteinverteilung innerhalb neuronaler Zellen in Fluoreszenzmikroskopbildern entwickelt. Dabei stehen eine einfache Benutzbarkeit sowie ein möglichst automatischer Ablauf der Auswertung im Vordergrund. Um eine einfache Nutzbarmachung der Applikation für den Anwender zu ermöglichen und diese auf einem einfachen Wege bereitzustellen, wurde die entwickelte Applikation in das MiToBo - Projekt [168] integriert und steht öffentlich für diverse Computerplattformen zur Verfügung.

Der `NeuronAnalyzer2D` kann auf Grundlage von 2D-Bildern die darin abgebildeten fluoreszenzmarkierten Einzelzellen automatisch detektieren. Dabei wird neben der eigentlichen Kontur der neuronalen Zellen verschiedene morphologische Merkmale ermittelt. Zudem werden einzelne strukturelle Bestandteile der Neuronen erkannt, aus denen weitere morphologische Eigenschaften extrahiert werden. Basierend auf den detektierten Strukturen ist der `NeuronAnalyzer2D` in der Lage, die Lokalisation markierter Proteine innerhalb der Zelle zu erfassen und die Verteilung der Proteine entlang der Neuriten zu quantifizieren und anhand eines grafischen Verlaufes darzustellen. Verwendet wurde der `NeuronAnalyzer2D` zur Ermittlung diverser morphologischer Merkmale neuronaler Zellen. Dazu zählen beispielsweise die Flächen der markierten Zellen und deren Zellkerne, die Anzahl der vorhandenen Neuriten und deren Verzweigungspunkte. Die Ermittlung der strukturellen Bestandteile der neuronalen Zellen stellt ein Novum zu herkömmlichen Anwendungen dar, in denen häufig nur Teile der zuvor genannten Merkmale extrahiert werden. Mittels der erkannten Strukturen wird es möglich, die neuronalen Zellen in detailliertere Bereiche zu unterteilen und spezifische Merkmale zu generieren. Neben der Zelle als Ganzes und den markierten Zellkernen können somit spezifische neuronale Bereiche wie das Soma oder die Neuriten identifiziert werden, um weitere Bereiche wie den Neuritenast oder die Wachstumskrone im Detail zu analysieren. Über die Quantifizierung der

Neuritenlängen und Verzweigungen hinaus bietet dies die Möglichkeit zur Erfassung weiterer morphologischer Merkmale wie z. B. Neuritenbreiten, Anzahl an Vorstülpungen (Spines) im Neuritenschaft und der Wachstumskrone sowie deren Größe, die in Wachstums-, Stimuli- oder Differenzierungsanalysen von Relevanz sind. Durch eine detailliertere Betrachtung der einzelnen Strukturen konnte innerhalb dieser Arbeit die Lokalisation von Proteinen untersucht und deren Verteilung entlang der Neuriten quantifiziert werden. Anhand der beschriebenen Lokalisation von F-Aktin und α -Tubulin, welche für die Aktindynamik und den Mikrotubuliaufbau der Neuriten verantwortlich sind, konnte gezeigt werden, dass die erfassten Verteilungen die beschriebenen biologischen Lokalisationen dieser Proteine abbilden und bestätigten. Die Detektionsergebnisse der neuronalen Zellen wurden anhand von manuell markierten Zellgrenzen sowie mit den Ergebnissen der öffentlich verfügbaren Anwendungen NeuriteTracer [196] und Vaa3D-Neuron [192] verglichen. Der in dieser Arbeit entwickelte NeuronAnalyzer2D zeigte dabei eine hohe Übereinstimmung der extrahierten Zellregionen auf den manuell markierten Zellen und erzielte insgesamt genauere Detektionsergebnisse im Vergleich zu den anderen getesteten Anwendungen.

Im Rahmen der Validierung der Detektionsleistung wurde deutlich, dass der NeuronAnalyzer2D aufgrund der Fluoreszenzmarkierung basierenden Detektion an sehr dünnen Neuriten zu Fehlern in der Detektion tendieren kann. Diese Neuriten enthalten oftmals geringe Mengen an markiertem Protein, welches bei langen Neuriten zudem über einen großen Bereich verteilt sein kann. Dadurch können geringere Fluoreszenzintensitäten entlang dieser Neuriten auftreten, die zu einer fehlerhaften Detektion durch die schwache Färbung führen. Im Fall der hier vorliegenden fixierten Zellen könnten ergänzende Färbungen in den Experimenten Anwendung finden, die diese Problematik umgehen, wie beispielsweise Markierung von Zellmembran- und/oder zytoplasmatischer Proteine. Damit könnte die Markierung der Zellen vervollständigt werden, um eine bessere Basis für die Detektion zu ermöglichen. Diese Ergänzungen sind mit zusätzlichem Kosten- und Zeitaufwand verbunden, könnten jedoch zu einer wesentlichen Verbesserung der Detektion beitragen und sollten in Abhängigkeit der experimentell notwendigen Proteinmarkierungen in Betracht gezogen werden. Alternativ könnte mittels Durchlichtmikroskopie eine gesamtheitliche Abbildung der Zellen erfolgen. Abhängig von der Art der zu untersuchenden Zellen führen diese Techniken jedoch besonders bei tubulären Strukturen, wie Neuriten, oft zu einer Unschärfe bzw. ungenaueren Abbildung, so dass spezifische Verfahren wie Phasenkontrast (PC) oder Differentialinterferenzkontrast (DIC) angewendet

werden müssen. Diese Verfahren ermöglichen durch Helligkeitsunterschieden die Erzeugung einer dreidimensional wirkenden Darstellung zur verbesserten Visualisierung der Zellen und Strukturen, aufgrund spezieller Beleuchtungs- und Schattenverläufe. Eine automatische Analyse der akquirierten Bilder erfordert jedoch eine nicht triviale Bildprozessierung oder Rekonstruktion der Zellen, bei der die dreidimensional Darstellung nachträglich und möglichst physikalisch korrekt zurückgerechnet werden muss, um eine genaue Lokalisation der Zellgrenzen in den zweidimensionalen Bildern zu ermöglichen. Die Durchlichtmikroskopie bietet im Vergleich zur Fluoreszenzbildern zudem einen geringeren Kontrast zwischen den abgebildeten Zellen und dem Hintergrund. Neben den intensitätsbedingten Abweichungen in der Zelldetektion wurde in nachfolgenden Betrachtungen festgestellt, dass bei der Erkennung der strukturellen Bestandteile der Neuronen die Abgrenzung der Wachstumskronen formabhängig ist. Längliche Formen der Wachstumskronen führen hierbei zu einer frühzeitigen Abtrennung innerhalb des Neuritenschaftes. Mögliches Entgegenwirken könnte hier mittels Einbeziehung weiterer Annahmen, die aus dieser länglichen Form resultieren, in die Wavelet basierte Zerlegung der strukturellen Bestandteile erzielt werden. So könnte beispielsweise der automatisch bestimmten Skalenbereich des ermittelte Modell-Peaks für den Bereich der Wachstumskrone angepasst werden, um die Grenzen entsprechend einer länglichen Wachstumskrone zu verschieben und dadurch eine genauere Abgrenzung zu ermöglichen.

Die verwendeten Parameter, welche im Abschnitt 2.1 beschrieben wurden, sind an die zugrunde liegenden Fluoreszenzbilder hippocampaler Neuronen angepasst und sind teilweise an die verwendeten optischen Gegebenheiten des Mikroskops gekoppelt (z. B. Vergrößerung/Auflösung des verwendeten Objektivs). Es wurde weitestgehend darauf geachtet, den Großteil der verwendeten Parameter automatisch und anhand der Bilddaten zu ermitteln. Darüber hinaus existieren weitere biologisch relevante Parameter, die anwendungsabhängig für die abgebildeten Neuronen angepasst werden sollten. Dazu gehören beispielsweise die Zellkerngröße oder die zu definierende Länge eines Spines. Anhand der Automatisierung der Applikation sollten notwendige Anpassungen jedoch unter geringen Aufwand zu testen sein. Manuelle Schwellwerte für Fluoreszenzintensitäten, wie sie z. B. im NeuriteTracer nötig sind, wurden vermieden. Generell sollten Veränderungen in den optischen Einstellungen oder der Untersuchung anderer Zelltypen bei der Bildakquise eigener Experimente berücksichtigt und die Parameter dementsprechend angepasst werden. Eine starke Sättigung fluoreszenzmarkierter Bereiche sollte vermieden werden, da somit strukturelle Informationen verloren gehen und dies eine

ungenauere Detektion, beispielsweise durch die Vergrößerung der Zellfläche, zur Folge hat. Die zu analysierenden Zellen sollten zudem möglichst vereinzelt und vollständig abgebildet sein. Bei längerem Neuritenwachstum ist zu beachten, dass der NeuronAnalyzer2D nicht in der Lage ist, sich kreuzende Neuriten zu trennen. In diesem Fall würden beide Neuronen herausgefiltert und nicht mit in die Analyse einbezogen werden.

Quantifizierung von Proteinprofilen: Lokalisation des Zipcode Binding Protein 1 (ZBP1)

Mithilfe des in dieser Arbeit entwickelten NeuronAnalyzer2D konnte erstmals die vollständige Lokalisation des *Zipcode Binding Protein 1* (ZBP1) in primären hippocampalen Neuronen ermittelt und dessen Verteilung innerhalb der Neuriten quantifiziert werden. Der Hippocampus fungiert als Informationsfilter für die Erzeugung neuer Erinnerung mittels neuronaler Veränderungen bzw. Entwicklung. Die Fähigkeit der Neuronen auf neue Anforderungen oder Reize der Umgebung, beispielsweise durch Neuritenwachstum, zu reagieren ist dabei von biologischer Relevanz [57]. Der Mechanismus der lokalen Translation von mRNAs spielt bei dieser Reaktion eine zentrale Rolle für das Auswachsen neuer Wachstumskronen, der Ausrichtung der Neuriten und die Entwicklung des gesamten Neurons [57, 109]. Das RNA-bindende Protein ZBP1 sowie dessen humanes Ortholog IGF2BP1 weisen ein onkofötales Expressionsmuster auf und werden aktuell mit diversen Krebserkrankungen in Verbindung gebracht [13, 180, 234]. Die Expression beeinflusst wesentliche Aspekte in der Zelle, wie beispielsweise die Migration, Morphologie und Differenzierung, und deutet auf eine zentrale Rolle in der Differenzierung von Zellen und im Zusammenhang mit dem Neuroblastom hin [14, 180]. Die Lokalisation von IGF2BP1 sowie dessen Veränderungen in den Neuroblastomstadien sollten somit anhand der hier entwickelten Methoden in weiterführenden Arbeiten genauer untersucht werden. Veränderungen der Zellmorphologie als auch des Transkriptomts zwischen Zelllinien verschiedener Neuroblastomstadien als auch innerhalb von Tumoren der Erkrankung könnten somit korreliert und auf ihre biologische Relevanz im Neuroblastom getestet werden.

Das mittels NeuronAnalyzer2D ermittelte Proteinprofil, bildet den Verlauf von ZBP1 vollständig innerhalb der Neuriten ab und bestätigt die beschriebene Lokalisation anhand bekannter funktionaler Wirkungsorte des Proteins in Neuriten. Anhand der Proteinprofile kann die Lokalisation von Proteinen untersucht und deren Verteilung entlang neuronaler Strukturen quantifiziert werden. Die Erfassung bekannter Proteinverteilungen der Proteine α -Tubulin und F-Aktin wurden zur Validierung der erzeugten Proteinverteilung herangezogen. Dabei

konnte gezeigt werden, dass sich die beschriebene Lokalisation beider Proteine abbilden und bestätigen lässt. Die hierbei erzeugte graphische Darstellung des Proteinprofils ermöglicht über die Quantifizierung hinaus einen schnellen und einfachen Weg, die Verteilung verschiedener Proteine zu vergleichen und strukturbezogen zu beschreiben. Zum Erhalt eines aussagekräftigen und validen Proteinprofils sollte, so es experimentell möglich ist, eine Vielzahl von Neuronen analysiert werden. Wie bereits erwähnt können beim NeuronAnalyzer2D geringen Abweichungen in der Detektion sehr dünner Neuriten auftreten, die einen negativen Einfluss auf das Proteinprofil haben können. Eine entsprechend hohe Zahl an untersuchten Neuriten sollte hierbei zu validen Ergebnissen des erzeugten Proteinprofils führen. Weiterhin sollten die optische Eigenschaften der Bildakquise so gewählt werden, dass eine sinnvolle Vergrößerung gewählt wird, die experimentell möglichst viele und vollständige Neuronen abbildet, sich jedoch nicht negativ auf die Strukturen der Neuriten auswirkt. Zudem beeinträchtigt eine zu geringe Auflösung die Quantifizierung der Fluoreszenzintensität und deren Verläufe entlang der untersuchten Strukturen. Eingesetzte Antikörper und Farbstoffe sollten im Vorfeld bezüglich Menge und Intensität getestet werden, um starke Schwankungen der Fluoreszenzintensitäten innerhalb der Experimente und zwischen den markierten Proteinen zu vermeiden. Weniger spezifische Antikörper würden hierbei, beispielsweise in zu hoher Menge eingesetzt, zu einer Verfälschung des tatsächlichen Proteinverlaufs führen. Dabei sollten abweichende Belichtungszeiten oder Laserintensitäten auf technischer Seite ebenso berücksichtigt werden und, abhängig vom Aufnahmegerät, entsprechend eingestellt und angepasst werden. Wie in Abschnitt 2.1.3 beschrieben, wurden Polynome 2. Ordnung zur Glättung verwendet, um eine vereinfachte Interpretation der Proteinprofile zu ermöglichen und eine Überanpassung des mittleren Intensitätsprofils zu vermeiden. Dadurch kann der generellen Verlauf der Proteine besser dargestellt werden, während Polynome höheren Grades größere Schwankungen im Verlauf hervorrufen und eine Interpretation erschweren. Dennoch kann in Abhängigkeit der untersuchten Proteine eine Anpassung des Polynomgrades durchaus sinnvoll und auch notwendig sein. Die ist beispielsweise der Fall, wenn das Protein keinen einheitlichen Verlauf, sondern Akkumulationen in bestimmten Regionen der untersuchten Strukturen aufweist und infolgedessen eine vereinfachte Glättung den Verlauf nicht korrekt widerspiegeln kann.

Differentielle Expression im Neuroblastom

Um die pathologischen Befunde, die zur klinischen Einteilung des Neuroblastoms herangezogen werden, durch molekularbiologische Aspekte zu ergänzen, wurden transkriptomische Veränderungen zwischen den primären Tumorproben der eingeteilten Neuroblastomstadien analysiert. Anhand differentieller Expressionsanalysen zwischen einzelnen Stadien, erfolgte in der vorliegenden Arbeit eine Quantifizierung signifikant veränderter Transkripte auf Grundlage von RNA- und microRNA-Sequenzierungsdaten. Als Ergebnis dieser Analysen konnte eine Vielzahl differentiell exprimierter Transkripte ermittelt werden, von denen Protein-codierende mRNAs und microRNAs genauer betrachtet und weiterführend untersucht wurden. Mittels in der Literatur beschriebener onkogener und tumorsuppressiver Gene ließen sich bekannte Transkripte der ermittelten differentiellen mRNAs und microRNAs (DE-mRNAs bzw. DE-microRNAs) bestätigen und die Ergebnisse der Analyse hinsichtlich ihrer Plausibilität validieren. Vergleiche mit dem Stadium 4S zeigten bei den Analysen die geringsten Veränderungen beim Test auf differentiell exprimierte Transkripte. Das Stadium ist aus molekularbiologischer Sicht unzureichend untersucht und zeigt z. B. bei der microRNA-Target Analyse noch nicht beschriebene Befunde (siehe übernächster Abschnitt), die in weiterführenden Untersuchungen detaillierten betrachtet werden sollten. Im Stadium 4 und in Proben mit einer *MYCN* Amplifikation weisen die vorliegenden Daten in dieser Arbeit eine große Anzahl differentiell exprimierten mRNAs und microRNAs auf. Innerhalb der risikoreicheren Stadien liegt anhand der analysierten Daten eine verstärkte Veränderung gegenüber anderen Stadien vor. Hervorzuheben ist hierbei das hoch risikoreiche und aggressive Neuroblastom, welches eine *MYCN* Amplifikation aufweist. Sowohl durch den Transkriptionsfaktors selbst als auch durch die Co-Amplifikation liegt eine starke Regulation diverser Ziel-Gene vor, gefolgt von einer starken Veränderungen des Transkriptoms sowie daran gekoppelter Signalwege. Co-amplifizierte als auch bekannte *MYCN* regulierte Gene können anhand der differentiellen Expressionsanalysen in dieser Arbeit identifiziert werden und bestätigen die beschriebene Expression diverser mRNAs, z. B. für *DKK3* oder *DDX1* [51, 123]. Vermehrt treten signifikant veränderte mRNAs auf, die in Wechselwirkung mit dem Wnt-Signalweg⁵⁸ stehen und diesen inhibieren oder aktivieren können. Gene mit einer inhibierenden Wirkung auf den Signalweg sind z. B. *DKK3*, *SOST* und *SFRP2*. In den aggressiven und risikoreichen Stadien zeigt sich beim Vergleich zu den weniger aggressiven

⁵⁸Vgl. hierzu auch https://www.genome.jp/kegg-bin/show_pathway?hsa04310 .

Stadien eine erhöhte Expression von Genen, die zur Aktivieren und damit Verstärkung des Signalweges führen, wie z. B. *BAMBI*, *FZD5*, *LRP5*, *WNT3* und *WNT3A*.

Die Analyse der microRNAs zeigt in der differentiellen Expression vergleichbare Ergebnisse, bei denen die risikoreicheren Stadien 4 und Proben mit *MYCN* Amplifikation die Mehrzahl an veränderten Transkripten in den Vergleichen aufweisen. Stark *MYCN* abhängige microRNAs als auch microRNA Cluster bestätigen anhand der durchgeführten Analysen die beschriebene Expression. Beispielsweise weisen die microRNAs des *MYCN* induzierten miR-17~92 Clusters sowie dessen paralogen Clustern miR-106a~363 und miR-106b~25 die stärkste Expression in Proben des Stadiums 4 mit *MYCN* Amplifikation auf. Weitere microRNAs sind z. B. *hsa-miR-21*, *hsa-miR-181* und *hsa-miR-204*, die innerhalb des risikoreichen Neuroblastoms bekannte Veränderungen zeigen [155, 216]. Neben der Bestätigung postulierter Befund lassen sich bekannte Transkripte identifizieren, die jedoch keine signifikant Veränderung oder eine entgegengesetzte Veränderungsrichtung aufweisen. Zum einen kann hierbei die Heterogenität der Expressionsdaten innerhalb der primären Tumorproben zu diesen Abweichungen führen. Zum anderen sind die unterschiedlichen Methoden zur Bestimmung der Expressionen in den publizierten Daten zu berücksichtigen. So zeigt z. B. die microRNA *hsa-miR-21-5p* eine hochregulierte Expression in *MYCN* amplifizierten Proben, weist jedoch im Test auf differentielle expression keine signifikante Veränderung auf. Ähnliches zeigt sich für die *hsa-miR-15a* welche in metastasierenden Tumoren hochreguliert ist, jedoch keine signifikante Veränderung in den vorliegenden Daten aufweist. Gleichwohl können, beim Vergleich mit publizierten Daten, unveränderte oder entgegengesetzte Expressionen aus den genannten Gründen auftreten, wie es beispielsweise bei der *hsa-miR-21-5p* oder der *hsa-miR-380-5p* zu beobachten ist. Die genannten microRNAs sind hierbei als hochregulierte microRNAs mit onkogenem Potential im Neuroblastom beschrieben [155]. Durch die Ergebnisse dieser Arbeit konnten mRNAs und microRNAs ermittelt werden, die nicht mit dem Neuroblastom, aber in verschiedenen Fällen mit anderen Krebserkrankungen assoziiert wurden. Die signifikante Veränderung in den vorliegenden Daten gibt einen Hinweise auf die biologische Relevanz im Neuroblastom, erbringt jedoch keinen Nachweis dafür. Um eine biologische Relevanz in den Stadien des Neuroblastoms nachzuweisen, sollten weiterführende Untersuchungen durchgeführt werden. Für eine valide und zuverlässige Aussage sollten umfangreiche Menge von Tumorproben analysiert werden, um die Varianzen in den Expressionsdaten zu verringern, bedingt durch die Heterogenität von primären Tumorprobe. Die Analyse von Einzelzellen mittels *Single-Cell RNA sequencing* aus den Tumorproben kann

ebenso zur Verbesserung der Analyse und Verringerung der Heterogenität beitragen. Außerdem sollten, unabhängig von der Methodik der Ermittlung der Expression und deren Analyse, Experimente zur quantitativen RNA-Bestimmung gezielt für einige der ermittelten RNAs realisiert werden, beispielsweise durch Methode wie qPCR (quantitative Echtzeit-PCR) oder smFISH (Einzelmolekül Fluoreszenz-in-situ-Hybridisierung).

Eine gesamtheitliche Betrachtung der Ergebnisse der differentiellen Expression zeigt, dass sowohl auf mRNA- als auch microRNA-Ebene in den Vergleichen der aggressiven und risikoreichen Stadien zu den risikoärmeren Stadien eine starke transkriptomische Veränderung vorliegt. Während bei den mRNAs innerhalb aller Vergleiche der Großteil der differentiell exprimierten mRNAs hochreguliert wird, zeigen die microRNAs eine tendenzielle Runterregulierung oder ähnlich große Veränderungen in beide Richtungen. Ausnahme bilden die Vergleiche, bei denen Proben enthalten sind, die eine *MYCN* Amplifikation aufweisen. Diese Vergleiche zeigen beim Test auf differentielle Expression größere Menge an hochregulierten microRNAs. Aus der Literatur konnte entnommen werden, dass beim Vergleich verschiedener Krebserkrankungen gegen Normalgewebe eine globale Reduzierung in der Expression von microRNAs vorliegt und eine Störung oder Regulierung in der microRNA Biogenese hierfür verantwortlich sein könnte [135]. Einige Biogenesekomponenten zeigen jedoch in verschiedenen Erkrankungen unterschiedliche Expressionen auf, so dass keine eindeutige Aussage diesbezüglich getroffen werden kann [135]. Die in dieser Arbeit erzielten Ergebnisse der differentiellen Expressionsanalyse zeigen, dass beim Vergleich der Stadien 1 – MYCN und 1 – 4^{+M} - entgegen der gängigen Annahme in Krebserkrankungen - der Großteil der microRNAs signifikant hochreguliert wird. Der Transkriptionsfaktor MYCN hat hierbei einen weitreichenden Einfluss auf die verstärkte Expression von microRNAs in den Tumoren dieser Stadien. Basierend auf dem vorliegenden Probenmaterial wurden sowohl mRNAs als auch microRNAs innerhalb jeder Probe analysiert. Damit war es möglich einen direkten Vergleich der Expression auf Ebene der mRNAs und microRNAs vorzunehmen. Genauere Betrachtungen von microRNA Biogenesefaktoren zeigen, dass die überwiegende Hochregulation von microRNAs in *MYCN* amplifizierten Tumoren aus einer verstärkten Bildung an microRNAs resultieren könnte. Dabei wurden in dieser Arbeit Faktoren unterschiedlicher Biogeneseschritte betrachtet, die insgesamt eine erhöhte - wenngleich nicht immer signifikante - Expression in der microRNA Regulation, dem Transport und der zytoplasmatischen Prozessierung aufweisen. Diese Befunde gehen mit der Beobachtung einer erhöhten Anzahl an differentiell exprimierten microRNAs aus der Expressionsanalyse in den

MYCN amplifizierten Proben einher, weisen zugleich jedoch abweichende Resultate zu bisher veröffentlichten Ergebnissen der Literatur auf [135]. Festzuhalten ist hierbei, dass die Daten der vorliegenden Arbeit aus primären Tumorproben erkrankter Patienten stammen und die Vergleiche innerhalb der Tumorstadien stattfinden. Die in [135] beschriebene Arbeiten basieren hingegen zum Teil auf Experimenten in Zelllinien, Geweben und/oder aus dem Vergleich von Tumoren zu Normalgeweben in Mäusen. Zudem erfolgen darin durchgeführte Quantifizierung der microRNA Expression oftmals mittels Microarrays oder PCR-Methoden. Die ausgewählten oder mittels Quantifizierung ermittelten microRNAs der Studien stellen meist nur einen Bruchteil der in dieser Arbeit untersuchten und aktuell annotierten Menge an microRNAs dar. In [142] wird eine globale Reduzierung von microRNAs in verschiedenen Tumoren (Brust, Lunge, Niere etc.) verglichen zu Normalgeweben postuliert, wobei von 217 betrachteten microRNAs 129 herunterreguliert sind. Für das Neuroblastom wird in [23] anstelle eines globalen Effekts eine allgemeinen Dysregulation der microRNA Expression aufgezeigt. Anhand von 145 untersuchten Proben werden von 37 differentiell exprimierten microRNAs, aus dem Vergleich zwischen *MYCN* amplifizierten und nicht amplifizierten Tumoren, 23 herunterreguliert. Bei detaillierter Betrachtung einiger Studien ist nicht eindeutig nachzuvollziehen, ob hierbei von einer globalen Reduzierung gesprochen werden kann. Ein Nachweis für die oftmals getroffene allgemeine Aussage bleibt überwiegend aus oder beschränkt sich nur auf Teilmengen bekannter microRNAs. Transkriptomweite Studien zeigen hingegen eine generelle Dysregulation an microRNAs, legen sich jedoch auf keine eindeutige globale Richtung der Expressionsveränderungen fest. Ähnlich der hier gewonnen Resultate sind dabei mehr hochregulierte microRNAs oder keine vorrangige Expressionsrichtung erkennbar [154, 224]. Wenige der Arbeiten betrachten dabei die Veränderungen der ablaufenden microRNA Biogeneseprozesse, wie es in der hier vorliegenden Arbeit erfolgte. In [133] wurden 66 Neuroblastomproben untersucht, wobei neben der microRNA Reduzierung eine geringe Expression der microRNA Biogenesefaktoren DROSHA und DICER für hochrisikoreiche Tumore postuliert wird. Dies unterstützt die Hypothese der erwähnten globalen Reduzierung von microRNAs in fortgeschrittenen Erkrankungen. Bei genauer Betrachtung von [133] zeichnet sich jedoch bei der Untersuchung *MYCN* amplifizierten Tumore keine eindeutige Expression dieser Faktoren ab. DICER weist in nicht amplifizierten Tumoren eine höhere Expression auf, während DROSHA eine geringere Expression in amplifizierten Tumoren erkennen lässt. Neben der Biogenese sind jedoch auch die Synthese- und Abbauraten der microRNAs entscheidend, so dass DROSHA und DICER keine ausreichenden

Faktoren für die Untersuchung einer globalen Veränderung darstellen. In der hier vorliegenden Arbeit deuten die untersuchten Komponenten der microRNA Biogenese auf eine verstärkten Bildung an microRNAs in *MYCN* amplifizierten Tumore hin, was den Befund einer erhöhten Anzahl an differentiell exprimierten microRNAs in diesen Tumoren bestätigt. Die Ergebnisse dieser Arbeit stehen somit nicht im Widerspruch zu anderen Arbeiten, Unterschiede zur Literatur sollten jedoch mittels experimenteller Versuche weiterführend untersucht werden. Die *MYCN* Amplifikation sollte tiefergehend analysiert werden, da ihr eine wesentliche Rolle sowohl in der microRNA Biogenese als auch der microRNA vermittelten Regulation zukommt. Zudem weisen die amplifizierte Tumore starke Varianzen in den ermittelten Resultaten dieser Arbeit im Vergleich mit publizierten Arbeiten, aber auch innerhalb dieser Arbeiten auf.

Gensignaturen

Zur Ergänzung der pathologischen Befunde der Tumorstadien durch weitere molekularbiologische Aspekte wurden Gensignaturen auf Basis der differentiell exprimierten mRNAs zwischen den Stadien ermittelt. Die resultierenden Signaturen enthalten für die Vergleiche der Stadien repräsentative Mengen an signifikant hoch- und herunterregulierter Genen aus den Vergleichen der differentiellen Expressionsanalyse. Diese Mengen an Genen können hilfreich für eine spezifischere Einteilung der Stadien sein und könnten im Hinblick auf zielgerichtete diagnostische und therapeutische Ansätze in einem frühen Stadium der Erkrankung Anwendung bei der Untersuchung stadienspezifischer Biomarker finden. Die Analyse der ermittelten Gensignaturen hinsichtlich bekannter biologischer Prozesse lies auch innerhalb der Gensignaturen einen Zusammenhang zu Wechselwirkungen mit dem Wnt-Signalweg erkennen. Dabei konnten indirekte Assoziationen über die Regulierung der Adipogenese ermittelt werden, welche auf der Modifikation von Histonen beruht und mithilfe der Methylierung den Wnt-Signalweg beeinflussen können. Direkte Zusammenhänge wurden aus den überrepräsentierten Gen-Sets der hier durchgeführten Gensignaturanalyse erkennbar, zu denen z. B. die Gen-Sets „LABBE WNT3A TARGETS“, „LABBE TARGETS OF TGFB1 AND WNT3A“, „WNT UP.V1“ und „KEGG WNT SIGNALING PATHWAY“ zählen.

Die aus den Vergleichen ermittelte neun Gensignaturen dieser Arbeit bestehen insgesamt aus 16 Gen-Sets, die sich in neun *Up*- und sieben *Down*-Sets der hochregulierten bzw. herunterregulierten Gene aufteilen. Mittels Validierung anhand öffentlich verfügbarer Datensätze aus Microarray- und RNA-Sequenzierungsanalysen konnte gezeigt werden,

dass fünf der identifizierten Gensignaturen valide Mengen von Gene darstellen, welche die Vergleiche der klinischen Stadien mithilfe molekularbiologischer Aspekte abbilden. Die Signaturen NB metas, NB s4S, NB s4S 4M⁻ und NB s4S 4 zeigen, anhand der erreichten Positionen im Vergleich mit den öffentlichen Daten, dass die Signaturgene die untersuchten Stadien nicht ausreichend gut repräsentieren. Die Sets dieser Signaturen weisen jedoch hierbei die geringste Anzahl an Genen auf (weniger als 30 Gene). Hier kann vermutet werden, dass für eine Charakterisierung der klinischen Stadien des Neuroblastoms eine so geringe Menge von Genen nicht ausreichend ist. Diese Gene scheinen die Variationen der Genexpression zwischen den Stadien nicht vollständig zu repräsentieren und die molekularbiologischen Gegebenheiten gegenüber anderen getesteten Zusammenhängen nicht deutlich abbilden zu können. Die Tatsache, dass aktuell keine umfangreichen stadienbasierten Signaturen existieren, erschwerte letztlich eine Prüfung der hier erzielten Ergebnisse. Weiterführende Arbeiten sollten klären, ob sich die nicht molekularbiologische klinische Einteilung nach dem INSS anhand molekularbiologischer Gegebenheiten in den Stadien abbilden lässt. Zudem sollte versucht werden die hier ermittelten Signaturen dahingehend zu optimieren, dass die Gene aller generierten Signaturen die Unterschiede in den Stadien charakterisieren und bei Vergleichen mit anderen krankheitsbezogenen Signaturen eindeutig auf die Erkrankung des Neuroblastoms und die Variationen zwischen den Stadien hinweisen. Dabei sollten erfasste Veränderungen zwischen den Stadien der Neuroblastomproben auch anhand vergleichbarer Datensätzen durch die hier gewonnenen Signaturen repräsentiert werden. Eine Optimierung dieser Signaturen könnte zum einen durch die Erweiterung der untersuchten Daten anhand weiterer Proben erreicht werden, um die molekularbiologischen Variationen deutlicher zwischen den Stadien herauszustellen und die biologischen Varianzen zu minimieren, welche zwischen den Einzelproben eines Stadiums auftreten. Weiterhin könnte eine Kombination der hier untersuchten Daten mit öffentlich verfügbaren Daten der gleichen oder zwischen verschiedenen Methoden zur Genexpressionsanalyse genutzt werden, um die Unterschiede der stadienbasierten Vergleichen genauer zu charakterisieren.

Putativ neue microRNA-Target Beziehungen

Um das Zusammenspiel von microRNAs und deren Zielgene (Targets) sowie damit verbundener Signalwege innerhalb der klinischen Stadien des Neuroblastoms besser zu verstehen, wurden putative microRNA-Target Beziehungen aus den Sequenzierungsdaten ermittelt, die eine hohe

Wahrscheinlichkeit für eine direkte Regulation der microRNA an die Target mRNA aufweisen. Dabei wurden Beziehungen gefunden, die aktuell nicht mit dem Neuroblastom in Verbindung gebracht werden oder von denen einer der Bindungspartner in diesem Zusammenhang unbekannt ist. Die in dieser Arbeit durchgeführte *in silico* Analyse zur Identifizierung möglicher direkter Bindung erfolgte unter Einbeziehung der gewonnenen Expressionsdaten, den Ergebnissen verschiedener Vorhersagedatenbanken sowie bindungsspezifischer Daten der microRNAs, mRNAs als auch eigener Überlebenszeitanalysen der klinischer Daten der untersuchten Tumorproben, wobei keine Translations-regulierenden Beziehungen innerhalb dieser Arbeit erfasst wurden. Anhand differentiell exprimierter microRNAs, aus den Vergleichen der Tumorstadien und mittels Datenbanken vorhergesagten Targets erfolgte die Ermittlung wahrscheinlicher Kandidaten einer direkten Bindung von microRNAs an differentiell exprimierte mRNAs. Dabei lag die Annahme zugrunde, dass für eine direkt Bindung der microRNA an ein mögliches Target dessen mRNA Expression bei erhöhter microRNA Expression, aufgrund der microRNA vermittelten Regulation, signifikant reduziert wird und umgekehrt. Diese Annahme beinhaltet, dass die differentielle Expression eines betrachteten Targets und somit dessen signifikante Expressionsveränderung im Wesentlichen auf der microRNA-Regulation basiert, wovon nicht auszugehen ist, da weitere Mechanismen (transkriptioneller oder posttranskriptioneller Art, RNA-Modifikationen etc.) sich auf die Expression der Targets auswirken. Eine direkte Bindung *in vivo* kann anhand der hier durchgeführten Analysen nicht nachgewiesen werden. Indirekte Regulationen können zudem nicht ausgeschlossen werden, so dass die gefundenen Beziehungen in weiterführenden Experimenten diesbezüglich untersucht werden sollten. Um die Qualität der ermittelten Beziehungen genauer einordnen zu können, wurde ein Bewertung basierend auf der Kombination unterschiedlicher Kriterien vorgenommen, die angibt, wie wahrscheinlich es ist, dass es sich bei der ermittelten Beziehung um eine direkte Bindung handelt. Vor dem Hintergrund, dass diese Bindungen nur möglicherweise eine direkte Regulation zeigen, konnten anhand der Vielzahl ermittelter microRNA-Target Beziehungen bekannte direkte Regulationen bestätigt werden. Daneben konnten aber auch bekannte Bindungen aufgezeigt werden, deren Beziehung sich anhand der vorliegenden Daten nicht bestätigen lassen. Die Gründe dafür können unterschiedlichen Ursprungs sein. Detaillierter Betrachtungen ergaben hierzu, dass oftmals die microRNA, aufgrund ihrer geringen Abundanz, aus der Analyse ausgeschlossen wird oder das vorhergesagte Target keine signifikante Veränderung aufzeigt, somit nicht differentiell exprimiert ist. Alternativ könnten hierfür weniger strikte Kriterien für die Selektion der

microRNAs und Targets verwendet werden. Dadurch würde sich die zu untersuchende Menge an möglichen Beziehungen jedoch stark vergrößern ohne zwingend positiv auf eine differentielle Veränderung hin getestet zu werden oder die Anzahl resultierender Bindungen aus der Analyse zu erhöhen. Zudem müssten diese zusätzlichen Beziehung ebenfalls validiert werden, was durch die schwächeren Kriterien nicht vereinfacht wird.

Von den bis zu zehn Besten der mehr als 30000 ermittelten Beziehungen aus den verschiedenen Vergleichen der Tumorstadien konnten die putativen Beziehungen ermittelt werden, die neu im Zusammenhang mit dem Neuroblastom stehen oder gänzlich unbekannt im Bezug auf Krebserkrankungen sind. Detaillierte Recherchen in der Literatur wurden durchgeführt, um im Rahmen dieser Arbeit ein Teil der ermittelten putativ neuen microRNA-Target Beziehungen zu untersuchen und wenn möglich zu bestätigen. Dabei wurden 67 dieser Beziehungen genauer betrachtet, wovon sieben microRNA:mRNA Bindungen identifiziert wurden, von denen weder die microRNA noch das ermittelte Target aktuell mit dem Neuroblastom assoziiert sind. Diese Bindungen deuten somit auf eine mögliche neue microRNA-Target Regulation in der Erkrankung hin. Die drei untersuchten Bindungen, *hsa-miR-744-5p:HJURP*, *hsa-miR-130b-3p:NDEL1* und *hsa-miR-1246:MATN2*, lassen eine verstärkte biologische Relevanz im Neuroblastom vermuten. Die microRNA *hsa-miR-744-5p* steht im Zusammenhang mit Wnt-Signalweg, welcher bereits innerhalb der zuvor betrachteten Analysen mehrfach in Erscheinung getreten ist.

Die erkannten und geschlussfolgerten Zusammenhänge aus der Literatur geben letztlich nur einen Hinweis auf eine mögliche Bindung und deren Relevanz in der Krebserkrankung. Die gewonnenen Befunde dieser Arbeit sollten daher als Ansätze für weiterführende Analysen dienen, um die Relevanz im Neuroblastom oder in anderen Krebserkrankungen zu bestätigen oder zu widerlegen. Die hier ermittelte Regulationen können helfen biologische Prozesse zu identifizieren, um die Erkrankung besser zu verstehen und neue Ansätze für die Prognose, Diagnose und Therapie zu entwickeln. Eine experimentelle Validierung der Regulationen sollte dabei im Vordergrund stehen, um in einem ersten Schritt eine direkte Regulation der mRNA durch die microRNA nachzuweisen. Dabei könnten Methoden wie beispielsweise die Immunpräzipitation von Komponenten des RISC-Komplexes, z. B. von Argonautenproteine (RIP, CLIP, CLASH), die Biotin-Markierung der identifizierten microRNAs oder die Affinitätsaufreinigung der microRNAs (miTRAP) angewendet werden [22, 85, 247].

5. Zusammenfassung und Ausblick

Im Rahmen der vorliegenden Arbeit wurden diverse Verfahren entwickelt, um die automatische Verarbeitung und Analyse von Bild- als auch Sequenzierungsdaten zu realisieren. Basierend auf zweidimensionalen Fluoreszenzbildern wurden Methoden zur Morphologieanalyse und zur Lokalisation von intrazellulär markierten Proteinen in neuronalen Zellen entwickelt. Zudem wurde mittels eigener Sequenzierungsdaten der Arbeitsgruppe das Transkriptom primärer Tumorproben von Neuroblastompatienten in verschiedener Stadien der Krebserkrankung untersucht und weiterführend analysiert. Die in der vorliegenden Arbeit untersuchten Gebiete, Morphologie und Genexpression, stellen zusammen wesentliche Faktoren bei der Untersuchung von Tumorerkrankungen dar, hinsichtlich diverser Aspekte wie Migration, Invasivität, Differenzierung oder Wachstum des Tumors, und sind von starker biologischer Relevanz für die Diagnose, Prognose und Therapie der Erkrankung. Die entwickelten Methoden zur automatischen Analyse der Bilddaten lieferten ein schnelles Verfahren zur Erzeugung einer Vielzahl morphologischer Daten, wie z. B. Zellflächen, Neuritenlängen, Anzahl der Neuriten und Verzweigungen. Diese würden beispielsweise einen einfachen Abgleich morphologischer Kennzahlen zwischen Zellen unter verschiedenen Bedingungen wie Stress, Wachstums und Differenzierung ermöglichen. Des Weiteren wurden die strukturellen Bestandteile neuronaler Zellen, wie z. B. Neuritenschaft, Wachstumskrone, sowie die Lokalisation angereicherter intrazellulärer Proteine erfasst. Um die Ergebnisse des Verfahrens der Zelldetektion zu validieren, wurden diese mittels manuell markierter Zellen als auch mit den Ergebnissen existierender Methoden, zur Erkennung neuronaler Zellen in Fluoreszenzbildern verglichen. Dabei zeigte sich, dass die entwickelte Methodik verlässliche Ergebnisse liefert, robust arbeitet und qualitativ bessere Ergebnisse bei der Evaluierung mit vergleichbaren Methoden generiert. Auf den detektierten Zellen basierend war es möglich strukturelle Kompartimente automatisch zu identifizieren und markierte Proteine innerhalb der Strukturen zu lokalisieren. Die erkannten Strukturen ermöglichten die Erfassung weiterer Kennzahlen, die in verfügbaren Anwendung

oftmals nicht ermittelt werden, wie beispielsweise Neuritenbreiten, Fläche des Somas oder der Wachstumskronen. Entlang der identifizierten Strukturen wurden Proteinverteilungen quantitativ erfasst und grafisch visualisiert. Dadurch können Verteilungen von Proteinen innerhalb der Zellen und zwischen verschiedenen experimentellen Bedingungen verglichen und gegenüber gestellt werden. Zudem ermöglichen die Strukturen eine detailliertere Eingrenzung der Lokalisation und bieten damit die Möglichkeit zur Erfassung des Verlaufes eines Proteins innerhalb der untersuchten Zellen und deren Strukturen. Hierzu konnte in dieser Arbeit mittels bekannter Proteinlokalisationen gezeigt werden, dass das entwickelte Verfahren verlässliche Ergebnisse liefert und sich zur quantitativen Darstellung von Proteinverteilungen in neuronaler Zellen eignet. Weiterhin konnte in der vorliegenden Arbeit erstmals die Verteilung des Proteins ZBP1 in primäre Mausneuronen beschrieben und der Verlauf des Proteins entlang der Neuriten vollständig dargestellt werden. Die Lokalisation des Proteins entspricht dabei der aus der Literatur bekannten funktionalen Relevanz des Proteins und bekräftigt die postulierte Funktion von ZBP1 im Zusammenhang mit der lokalisierten Translation von β -Aktin während des neuronalen Zellwachstums. Die entwickelten Verfahren zur Detektion neuronaler Zellen in Fluoreszenzmikroskopbilder sowie die sich daran anschließende Extraktion neuronaler Strukturen und die Lokalisation von markierten Proteinen wurde in die *Microscope Image Analysis Toolbox*, einer Erweiterung für die populäre Bildanalyseplattform ImageJ, integriert und öffentlich verfügbar gemacht.

Im weiteren Teil der Arbeit fand die molekularbiologische Analyse der Sequenzierungsdaten der untersuchten primären Proben des Neuroblastoms statt, einer Krebserkrankung im Kindesalter, die in vielen Fällen mit einer schlechten Prognose verbunden ist. Grundlage dieser Arbeit bilden mehr als 60 untersuchte Tumorproben, einschließlich pathologischer Befunde und klinischer Daten. Die Aufbereitung der Proben erlaubte es innerhalb dieser Arbeit, die transkriptomischen Veränderungen zwischen den verschiedenen Tumorstadien sowohl auf mRNA- als auch auf microRNA-Ebene zu untersuchen und eine kombinierte Auswertung der Expressionsanalysen zu entwickeln. Durch die Analyse der Sequenzierungsdaten konnte die Expression aktuell annotierter Transkripte in den einzelnen Proben bestimmt werden. Die verarbeiteten Daten sind über die *R2 - Genomics Analysis and Visualization Platform* zugänglich und für weitere Untersuchungen nutzbar. Auf diesen Daten basierend wurden Expressionsanalysen zur Identifizierung signifikant veränderter Transkripte zwischen den untersuchten Tumorstadien durchgeführt. Die Betrachtung von Protein-codierenden mRNAs

und microRNAs ergab eine Vielzahl veränderter Transkripte zwischen den Stadien, woraus über 200 differentiell exprimierten microRNAs und 2000 mRNAs ermittelt werden konnten. Eine genauere Betrachtung der Transkripte zeigte sowohl bekannte als auch mögliche neue mRNA bzw. microRNA Transkripte mit biologischer Relevanz im Neuroblastoms auf. Mittels Literatur konnten die Daten der differentiellen Expression anhand von bekannten Onkogen oder Tumorsuppressoren evaluiert werden. Aufwendige Literaturrecherchen konnten nachfolgend putativ neue Transkripte in Verbindung mit dem Neuroblastom identifizieren. Hierzu zählen beispielsweise mögliche onkogen oder tumorsuppressiv wirkende Transkripte, wie z. B. *ASPM*, *KIF14*, *hsa-miR-18b* respektive *RAB3B*, *hsa-miR-125a-5p* und *hsa-miR-148a*, die auf eine biologische Relevanz in der Diagnose oder im Bezug auf therapeutische Maßnahmen hindeuten. Auf Grundlage der ermittelten Veränderungen der Genexpressionen zwischen den Neuroblastomstadien fanden weiterführenden Analyse der verarbeiteten Sequenzierungsdaten statt. Ein Ergebnis dieser intensiveren Untersuchungen war die Erarbeitung spezifischer Gensignaturen aus den Vergleichen der Stadien des Neuroblastoms, um Mengen von charakteristischen Genen zwischen den Stadien zu bestimmen. Damit konnten die pathologischen Befunde, zur Einteilung der klinischen Stadien unter molekularbiologischen Aspekten betrachtet und um diese Aspekte ergänzt werden. Im Rahmen dieser Arbeit wurden neun Signaturen mit bis zu 2081 enthaltenen Genen generiert. Diese Gensignaturen wurden anhand öffentlicher Datensätze validiert und hinsichtlich ihrer Relevanz zu bereits annotierten biologischen Prozessen in anderen Krebserkrankungen analysiert. Dabei zeigte sich, dass fünf dieser Signaturen valide Mengen charakteristischer Gene der transkriptomischen Veränderungen zwischen den Stadienvergleichen enthalten, welche die pathologischen Stadien anhand molekularbiologischer Aspekte abbilden. Die durch die vorliegende Arbeit erbrachte Verknüpfung der pathologischen und molekularbiologischen Daten stellt eine wesentliche Erweiterung der bisherigen pathologisch vermittelten Befunden dar. Zukünftig sollten diese kombinierten Informationen im Hinblick auf ein spezifischere Einteilung der Stadien und damit einhergehenden für eine zielgerichtete Diagnose oder Therapie in einem frühen Stadium der Erkrankung Anwendung finden. Die enthaltenen Gene der Signaturen wurden nachfolgend bezüglich bekannter und unbekannter biologischer Zusammenhänge und Prozesse untersucht. Neben der Bestätigung bekannter und beschriebener Vorgänge konnten diverse neue Zusammenhänge mit dem Neuroblastom assoziiert werden. Dazu zählen beispielsweise Prozesse die an Entzündungsreaktionen, Kanalaktivitäten oder der Reaktion auf UV Bestrahlung beteiligt

sind. Besonders hervorzuheben sind jedoch Assoziation zu Prozessen der Methylierung sowie der Adipositas und Diabetes. Untersuchungen hierzu verfestigten die Annahme, dass die Regulierung der Adipogenese, durch Modifikation von Histonen und mithilfe von Methylierungsreaktionen, den Wnt-Signalweg beeinflussen und somit allem Anschein nach am Prozess der Differenzierung im Neuroblastom beteiligt sein könnten. Hier bietet sich ein möglicher Ansatz für die Therapie der Tumorerkrankung, da das reduzierte Differenzierungspotential vor allem in den aggressiven Stadien für die schlechte Prognose mitverantwortlich ist und eine Behandlung in zahlreichen Therapien zielführend, aber nicht immer erfolgreich ist. Die im Rahmen dieser Arbeit entwickelten Methoden zur Erstellung und Auswertung von spezifischen Gensignaturen sollte sich dabei auf jedwede Art von Tumorerkrankung übertragen lassen und kann sowohl neue Erkenntnisse als auch Ansätze für die Diagnose bzw. Therapie der Erkrankung hervorbringen. An dieser Stelle sei erwähnt, dass während der molekularbiologische Analysen der Sequenzierungsdaten in allen untersuchten Bereichen, sowohl bei der differentiellen Expression, der Gensignaturanalyse als auch in den nachfolgend beschriebenen Ergebnissen der microRNA-Targetanalyse, ein immer wiederkehrender Bezug zum Wnt-Signalweg erkennbar ist. Die damit einhergehenden Befunde deuten an, dass der Signalweg im Neuroblastom eine starke biologische Relevanz aufweist und vorwiegend in den aggressiven und risikoreichen Tumoren eine erhöhte Expression der Wnt Genfamilie vorliegt. Inwieweit wichtige Funktionen oder Mechanismen innerhalb der Tumorerkrankung durch den Signalweg reguliert werden ist zur Zeit nicht vollends untersucht und wurden im Rahmen dieser Arbeit nicht weiter verfolgt. Wenige Publikationen deuten jedoch ein Zusammenhang des Signalweges mit dem Neuroblastom an und sollten in weiterführenden Studien eingehender betrachtet werden.

Anhand der Probenaufbereitung war es im letzten Teil dieser Arbeit möglich eine kombinierte Auswertung der Expression von mRNAs als auch microRNAs hinsichtlich der Veränderungen in den Tumorstadien auf mRNA- und microRNA-Ebene gleichermaßen zu untersuchen und miteinander zu korrelieren. Bei der hierfür entwickelten Targetanalyse wurden microRNA-Target Beziehungen ermittelt, die eine putativ direkte Bindung zwischen einer differentiell exprimierten microRNA und deren möglicherweise regulierten mRNA Targets erkennen lassen. Anhand beschriebener Regulationen im Neuroblastom konnten diverse bekannte Regulationen bestätigt werden. Dabei zeigt sich die miR-17 Familie als eine der stärksten exprimierten microRNA Familien in den aggressiveren und risikoreichen Tumoren des Neuroblastoms und weist dazu zahlreiche Targets auf, die durch ihre microRNA vermittelte negative Regulation

mit der Aggressivität und der schlechten Prognose der Tumorerkrankung in Verbindung stehen, wie beispielsweise *DKK3* oder *TGFBR2*. Insgesamt konnten mittels Targetanalyse 69 verschiedene microRNAs und 2808 verschiedene potentielle Target mRNAs ermittelt werden, wovon 67 putativ neue microRNA:mRNA Bindungen genauer untersucht wurden. Unter diesen ermittelten Bindungen befanden sich sieben, die auf eine mögliche neue microRNA-Target Regulation im Neuroblastoms hindeuten könnten. Hervorzuheben ist hierbei die ermittelte Bindung *hsa-miR-130b-3p:NDEL1*, welche Auswirkungen auf die Spindelfaserorientierung und den NOTCH-Signalweg hat und somit relevante Prozesse des Wachstums, der Differenzierung sowie der Tumorgenese zu beeinflussen scheint. Dies könnte in weiterführenden Arbeiten interessante Befunde hinsichtlich therapeutischer Ansätze im Neuroblastom oder anderer Krebserkrankungen liefern.

Die Korrelation auf mRNA- und microRNA-Ebene ermöglichte es zudem, Komponenten der microRNA Biogenese und Aktivität innerhalb dieser Arbeit detaillierter zu betrachten. Dabei zeigte sich, dass vornehmlich in *MYCN* amplifizierten Tumoren des Neuroblastoms eine verstärkte Abundanz an microRNAs vorliegt. Insgesamt konnte anhand der analysierten Daten beobachtet werden, dass es entgegen einer globalen Reduktion, wie sie teilweise in der Literatur für einige Krebserkrankungen beschrieben ist, eine verstärkte Biogenese und damit eine vermehrte Reifung von microRNAs erkennbar ist. Evaluieren lässt sich dies an der erhöhten differentiellen Expression und den in dieser Arbeit betrachteten Faktoren der einzelnen Reifungsschritten der microRNA Biogenese, die mitunter alle auf eine Verstärkung der Biogenese hindeuten, sowohl in der nukleären Prozessierung, dem Transport ins Zytoplasma und der sich anschließenden zytoplasmatischen Prozessierung zur reifen microRNA. Diese Befunde lassen darauf schließen, dass der Transkriptionsfaktor MYC respektive MYCN im Neuroblastom zu einer enormen Veränderung der microRNA Expression führt, was auf allen Ebenen der Biogenese anhand der analysierten Daten dieser Arbeit nachvollzogen werden kann. Vermittelt durch die verstärkte Reifung von microRNAs liegt der Schluss nahe, dass dies ebenfalls zu einer möglichen Verstärkung des RNA-Abbaus durch die microRNA vermittelten RNA-Stillegung führen kann. Untersuchungen hierzu haben in dieser Arbeit gezeigt, dass in Tumoren mit einer hohen *MYCN* Abundanz eine verstärkte microRNA Aktivität vorliegen könnte. Wesentliche Komponenten der RNA-Stillegung weisen hierbei eine erhöhte Expression innerhalb dieser Tumore auf. Insgesamt deuten die Ergebnisse aus den Betrachtungen an, dass *MYCN* amplifizierte Tumor eine erhöhte microRNA Expression und damit einhergehend die

Möglichkeit zur verstärkten microRNA vermittelten RNA-Stillegung zeigen.

Die im Rahmen dieser Arbeit generierten Gensignaturen und microRNA-Target Beziehungen sollen perspektivisch auf der *Molecular Signatures Database*⁵⁹ zugänglich gemacht werden, um damit erstmals Signaturen zu den verschiedenen Neuroblastomstadien öffentlich zur Verfügung zu stellen. MicroRNAs gewinnen in aktuellen Studien der Krebsforschung immer mehr an Bedeutung. Die Funktionalität und Beziehungen der microRNAs werden dabei meist in Form von (Target- oder Vorhersage-) Datenbanken nutzbar gemacht, weisen jedoch häufig fehlerbehaftete Sequenzen, Bindestellen oder unvollständige Vorhersagen auf. Im Gegensatz zu den beschriebenen öffentlichen Gensignaturen liegen Signaturen, basierend auf microRNAs, nur in geringem Maße vor, mit deren Hilfe sich relevante Prozesse oder Mechanismen in den Erkrankungen analysieren ließen. Oftmals sind neben wenigen krankheitsbezogenen microRNAs nur vereinzelte microRNA-Familien oder *in silico* vorhergesagte Targets beschrieben, die folglich als Grundlage für weiterführende Analysen dienen und deren experimentelle Nachweise zur Bindung oder Funktion oftmals noch ausstehen. Aus diesen Gründen wäre es wünschenswert, valide microRNA-Signaturen zu erstellen und vorhandene *Gene Set Enrichment Analyses* für deren Auswertungen zu erweitern oder anzupassen, um die meist microarraybasierten Signaturen zu ergänzen und mittels moderner Techniken genierten microRNA-Expressionsdaten in ähnlicher Form der Gensignaturen bereitzustellen und zu validieren. Zusätzlich zu den in dieser Arbeit analysierten Tumorproben wurden bereits 30 weitere Proben des Neuroblastoms in gleicher Weise sequenziert. Diese sollen in die aktuell vorliegenden Daten integriert werden, um die Variabilität innerhalb der einzelnen Stadien besser zu kompensieren sowie mögliche neue Erkenntnisse zu erzielen. Dazu sollen zudem Betrachtungen auf DNA-Ebene durchgeführt werden, die in Form von Genomanalysen (*Whole Genome Sequencing, WGS*) Auskunft über genomische Veränderungen, wie z. B. Stärke der *MYCN* Amplifikation, innerhalb der Proben und zwischen den Stadien geben. Neben primären Tumorproben sollen zudem Untersuchungen in diversen Zelllinien stattfinden. Dabei soll geprüft werden, inwiefern sich die in dieser Arbeit ermittelten Signaturen auf die Zellsysteme übertragen lassen und ob bzw. in wie weit die Zelllinien noch ihren ursprünglichen Tumorstadien entsprechen.

Durch die entwickelten Verfahren in der vorliegenden Arbeit konnte erstmalig eine vollständige Quantifizierung der Verteilung von ZBP1 in neuronalen Zellen ermittelt und diese Verteilung entlang der Neuriten grafisch dargestellt werden. Im Hinblick auf neueste Untersuchungen im

⁵⁹<http://software.broadinstitute.org/gsea/msigdb/index.jsp>

Zusammenhang mit dem humanen ZBP1 Ortholog IGF2BP1 im Neuroblastom und anderen Krebserkrankungen [13, 14, 171] ermöglichen die entwickelten Verfahren und gewonnenen Erkenntnisse in dieser Arbeit eine Grundlage zur Verknüpfung von Lokalisierungs- sowie morphologischen und transkriptomischen Analysen. Zukünftig können somit IGF2BP1 vermittelte (und anderer proteinabhängige) Veränderungen im Bezug auf Differenzierung, Wachstum, Metastasierung oder Tumorgenese im Neuroblastom auf morphologischer und transkriptomischer Ebene gemeinsam analysiert und miteinander kombiniert werden. Die in Form der identifizierten Gensignaturen ermittelten Mengen veränderter Gene aus den stadienbasierten Vergleichen dieser Arbeit ergänzen die pathologischen Befunde die zur klinischen Einteilung der Stadien des Neuroblastoms verwendet werden. Diese molekularbiologische Ergänzung kann für eine spezifischere Einteilung und zur Ermittlung genauerer Prognosen verwendet werden. Die im Rahmen dieser Arbeit ermittelten Zusammenhänge zur Methylierung und damit verbundenen Auswirkung auf den Wnt-Signalweg sowie der microRNA vermittelten Regulation der Spindelfaserorientierung und des NOTCH-Signalwegs bieten erste direkte Ansatzmöglichkeiten relevante Prozesse des Wachstum, der Differenzierung und der Tumorgenese im Neuroblastom gezielt zu untersuchen. Weiterführenden Analysen und experimentelle Nachweise der Relevanz dieser Zusammenhänge könnten bestehende Therapien ergänzen oder zur Entwicklung neuer Therapieansätze beitragen.

Literatur

- [1] ACKERMANN, SANDRA et al.: *Polo-Like Kinase 1 is a Therapeutic Target in High-Risk Neuroblastoma*. *Clinical Cancer Research*, 17(4):731–741, 2011.
- [2] ALEXIOU, PANAGIOTIS et al.: *Lost in Translation: An Assessment and Perspective for Computational microRNA Target Identification*. *Bioinformatics*, 25(23):3049–3055, 2009.
- [3] ANDERS, SIMON und HUBER, WOLFGANG: *Differential Expression Analysis for Sequence Count Data*. *Genome Biology*, 11(10):1–12, 2010.
- [4] AURENHAMMER, FRANZ und KLEIN, ROLF: *Voronoi Diagrams**. In: SACK, J.-R. und URRUTIA, J. (Herausgeber): *Handbook of Computational Geometry*, Kapitel 5, Seiten 201–290. North-Holland, Amsterdam, Netherlands, 2000.
- [5] BAHRAMI, ARMITA et al.: *Pigment-Synthesizing Melanocytic Neoplasm with Protein Kinase C Alpha (PRKCA) Fusion*. *JAMA Dermatology*, 152(3):318–322, 2016.
- [6] BALAMURUGAN, KUPPUSAMY und STERNECK, ESTA: *The Many Faces of C/EBP δ and their Relevance for Inflammation and Cancer*. *Int J Biol Sci*, 9(9):917–933, 2013.
- [7] BASSETT, ANDREW R. et al.: *Understanding Functional miRNA-Target Interactions In Vivo by Site-Specific Genome Engineering*. *Nature Communications*, 5:4640, 2014.
- [8] BATES, DOUGLAS M. und WATTS, DONALD G.: *Nonlinear Regression Analysis and its Applications*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 1988.
- [9] BAUER, CHRISTIAN und BISCHOF, HORST: *A Novel Approach for Detection of Tubular Objects and its Application to Medical Image Analysis*. In: RIGOLL, GERHARD (Herausgeber): *Pattern Recognition, 30th DAGM Symposium, Munich, Germany, June 10-13, 2008, Proceedings*, Band 5096, Seiten 163–172. Springer, 2008.
- [10] BEARD, JORDAN A., TENGA, ALEXA und CHEN, TAOSHENG: *The Interplay of NR4A Receptors and the Oncogene-Tumor Suppressor Networks in Cancer*. *Cellular Signalling*, 27(2):257–266, 2015.
- [11] BECKERS, ANNELEEN et al.: *MYCN-Driven Regulatory Mechanisms Controlling LIN28B in Neuroblastoma*. *Cancer letters*, 366(1):123–132, 2015.
- [12] BECKERS, ANNELEEN et al.: *MYCN-Targeting miRNAs are Predominantly Downregulated During MYCN-Driven Neuroblastoma Tumor Formation*. *Oncotarget*, 6(7):5204–5216, 2015.
- [13] BELL, JESSICA L. et al.: *Insulin-Like Growth Factor 2 mRNA-Binding Proteins (IGF2BPs): Post-Transcriptional Drivers of Cancer Progression?* *Cellular and Molecular Life Sciences*, 70(15):2657–2675, 2013.
- [14] BELL, JESSICA L. et al.: *IGF2BP1 Harbors Prognostic Significance by Gene Gain and Diverse Expression in Neuroblastoma*. *Journal of Clinical Oncology*, 33(11):1285–1293, 2015.
- [15] BENJAMINI, YOAV und HOCHBERG, YOSEF: *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.
- [16] BEWICK, VIV, CHEEK, LIZ und BALL, JONATHAN: *Statistics Review 12: Survival Analysis*. *Critical Care*, 8(5):389–394, 2004.
- [17] BHATTACHARYA, SOURAV et al.: *Serum miR-30e and miR-223 as Novel Noninvasive Biomarkers for Hepatocellular Carcinoma*. *The American Journal of Pathology*, 186(2):242–247, 2016.
- [18] BLIN, MARYLINE et al.: *NR4A2 Controls the Differentiation of Selective Dopaminergic Nuclei in the Zebrafish Brain*. *Molecular and Cellular Neuroscience*, 39(4):592–604, 2008.
- [19] BOCKHORN, JESSICA et al.: *MicroRNA-30c Inhibits Human Breast Tumour Chemotherapy Resistance by Regulating TWF1 and IL-11*. *Nature Communications*, 4(1393):1–14, 2013.

- [20] BORGEFORS, GUNILLA: *Distance Transformations in Digital Images*. Computer Vision, Graphics, and Image Processing, 34:344–371, 1986.
- [21] BOYERINAS, BENJAMIN et al.: *The Role of let-7 in Cell Differentiation and Cancer*. Endocrine-Related Cancer, 17(1):F19–F36, 2010.
- [22] BRAUN, JULIANE et al.: *Rapid Identification of Regulatory microRNAs by miTRAP (miRNA Trapping by RNA in Vitro Affinity Purification)*. Nucleic Acids Research, 42(8):e66, 2014.
- [23] BRAY, ISABELLA et al.: *Widespread Dysregulation of miRNAs by MYCN Amplification and Chromosomal Imbalances in Neuroblastoma: Association of miRNA Expression with Survival*. PLoS ONE, 4(11):e7850, 2009.
- [24] BRESLER, SCOTT et al.: *ALK Mutations Confer Differential Oncogenic Activation and Sensitivity to ALK Inhibition Therapy in Neuroblastoma*. Cancer Cell, 26(5):682–694, 2014.
- [25] BRODEUR, GARRETT M.: *Spontaneous Regression of Neuroblastoma*. Cell and Tissue Research, 372(2):277–286, 2018.
- [26] BUECHNER, JOCHEN et al.: *Tumour-Suppressor microRNAs let-7 and mir-101 Target the Proto-Oncogene MYCN and Inhibit Cell Proliferation in MYCN-Amplified Neuroblastoma*. British Journal of Cancer, 105(2):296–303, 2011.
- [27] BULLARD, JAMES H. et al.: *Evaluation of Statistical Methods for Normalization and Differential Expression in mRNA-Seq Experiments*. BMC Bioinformatics, 11(1):1–13, 2010.
- [28] BUTRYM, ALEKSANDRA et al.: *Low Expression of microRNA-204 (miR-204) is Associated with Poor Clinical Outcome of Acute Myeloid Leukemia (AML) Patients*. Journal of Experimental & Clinical Cancer Research, 34(1):1–5, 2015.
- [29] CABUK, AYTEKIN D., ALPAY, ERDENAY und ACAR, BURAK: *Detecting Tubular Structures via Direct Vector Field Singularity Characterization*. Conf Proc IEEE Eng Med Biol Soc, 1:4801–4804, 2010.
- [30] CANNON, ROBERT C. et al.: *An On-Line Archive of Reconstructed Hippocampal Neurons*. Journal of Neuroscience Methods, 84(1–2):49–54, 1998.
- [31] CAPOWSKI, J. J.: *An Automatic Neuron Reconstruction System*. Journal of Neuroscience Methods, 8(4):353–364, 1983.
- [32] CASTELLA, MARIA et al.: *FANCI Regulates Recruitment of the FA Core Complex at Sites of DNA Damage Independently of FANCD2*. PLoS Genetics, 11(10):1–27, 2015.
- [33] CEKAITE, LINA et al.: *MiR-9, -31, and -182 Deregulation Promote Proliferation and Tumor Cell Survival in Colon Cancer*. Neoplasia, 14(9):868–881, 2012.
- [34] CHAKRAVARTI, INDRA MOHAN, LAHA, RADHA GOVIND und ROY, JOGABRATA. In: *Handbook of Methods of Applied Statistics*, Seiten 392–394. John Wiley and Sons, New York, NY, USA, 1967.
- [35] CHAN, TONY F. und VESE, LUMINITA A.: *Active Contours Without Edges*. IEEE Transactions on Image Processing, 10(2):266–277, 2001.
- [36] CHANG, SUKMOON, METAXAS, DIMITRIS N. und AXEL, LEON: *Scan-Conversion Algorithm for Ridge Point Detection on Tubular Objects*. In: ELLIS, RANDY E. und PETERS, TERRY M. (Herausgeber): *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003, 6th International Conference, Montréal, Canada, November 15-18, 2003, Proceedings, Part II*, Band 2879, Seiten 158–165. Springer, 2003.
- [37] CHANG, TINGJIE et al.: *MicroRNA-30a Promotes Extracellular Matrix Degradation in Articular Cartilage via Downregulation of Sox9*. Cell Proliferation, 49(2):1365–2184, 2016.
- [38] CHEN, YONGXIN und STALLINGS, RAYMOND L.: *Differential Patterns of microRNA Expression in Neuroblastoma are Correlated with Prognosis, Differentiation, and Apoptosis*. Cancer Research, 67(3):976–983, 2007.
- [39] CHEN, YUYAN et al.: *Oncogenic Mutations of ALK Kinase in Neuroblastoma*. Nature, 455(7215):971–974, 2008.

- [40] CHENG, CHAO, BHARDWAJ, NITIN und GERSTEIN, MARK: *The Relationship Between the Evolution of microRNA Targets and the Length of their UTRs*. *BMC Genomics*, 10(1):431, 2009.
- [41] CHENG, CHUN-WEN et al.: *MicroRNA-30a Inhibits Cell Migration and Invasion by Downregulating Vimentin Expression and is a Potential Prognostic Marker in Breast Cancer*. *Breast Cancer Research and Treatment*, 134(3):1081–1093, 2012.
- [42] CHERVONA, YANA und COSTA, MAX: *Histone Modifications and Cancer: Biomarkers of Prognosis?* *American Journal of Cancer Research*, 2(5):589–597, 2012.
- [43] COCK, PETER J. A. et al.: *The Sanger FASTQ File Format for Sequences with Quality Scores, and the Solexa/Illumina FASTQ Variants*. *Nucleic Acids Research*, 38(6):1767–1771, 2010.
- [44] COHN, SUSAN L. et al.: *The International Neuroblastoma Risk Group (INRG) Classification System: An INRG Task Force Report*. *Journal of Clinical Oncology*, 27(2):289–297, 2009.
- [45] COLNAGHI, LUCA et al.: *Patient-Derived C-Terminal Mutation of FANCI Causes Protein Mislocalization and Reveals Putative EDGE Motif Function in DNA Repair*. *Blood*, 117(7):2247–2256, 2011.
- [46] CONCEPCION, CARLA P., BONETTI, CIRO und VENTURA, ANDREA: *The miR-17-92 Family of microRNA Clusters in Development and Disease*. *Cancer Journal*, 18(3):262–267, 2012.
- [47] CORTÉS-CANTELI, MARTA et al.: *CCAAT/Enhancer-Binding Protein β Plays a Regulatory Role in Differentiation and Apoptosis of Neuroblastoma Cells*. *Journal of Biological Chemistry*, 277(7):5460–5467, 2002.
- [48] DAI, WEIJUN et al.: *A Post-Transcriptional Mechanism Pacing Expression of Neural Genes with Precursor Cell Differentiation Status*. *Nature Communications*, 6(7576):1–11, 2015.
- [49] DAS, ANI V. und PILLAI, RADHAKRISHNA M.: *Implications of miR Cluster 143/145 as Universal Anti-OncomiRs and their Dysregulation During Tumorigenesis*. *Cancer Cell International*, 15(92):1–12, 2015.
- [50] DAUBECHIES, INGRID: *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
- [51] DE BROUWER, SARA et al.: *Dickkopf-3 is Regulated by the MYCN-Induced miR-17-92 Cluster in Neuroblastoma*. *International Journal of Cancer*, 130(11):2591–2598, 2012.
- [52] DEHMELT, LEIF et al.: *NeuriteQuant: An Open Source Toolkit for High Content Screens of Neuronal Morphogenesis*. *BMC Neuroscience*, 12(100), 2011.
- [53] DENT, ERIK W. und GERTLER, FRANK B.: *Cytoskeletal Dynamics and Transport in Growth Cone Motility and Axon Guidance*. *Neuron*, 40(2):209–227, 2003.
- [54] DI GIOVINE, PAOLO et al.: *Structure of Herpes Simplex Virus Glycoprotein D Bound to the Human Receptor Nectin-1*. *PLoS Pathogens*, 7(9):1–13, 2011.
- [55] DILLIES, MARIE-AGNÈS et al.: *A Comprehensive Evaluation of Normalization Methods for Illumina High-Throughput RNA Sequencing Data Analysis*. *Briefings in Bioinformatics*, 14(6):671–683, 2013.
- [56] DOBSON, JASON R. et al.: *Hsa-mir-30c Promotes the Invasive Phenotype of Metastatic Breast Cancer Cells by Targeting NOV/CCN3*. *Cancer Cell International*, 14(73):1–14, 2014.
- [57] DOYLE, MICHAEL und KIEBLER, MICHAEL A.: *Mechanisms of Dendritic mRNA Transport and its Role in Synaptic Tagging*. *The EMBO Journal*, 30(17):3540–3552, 2011.
- [58] DU, PAN, KIBBE, WARREN A. und LIN, SIMON M.: *Improved Peak Detection in Mass Spectrum by Incorporating Continuous Wavelet Transform-Based Pattern Matching*. *Bioinformatics*, 22(17):2059–2065, 2006.
- [59] DURINCK, STEFFEN et al.: *BioMart and Bioconductor: A Powerful Link Between Biological Databases and Microarray Data Analysis*. *Bioinformatics*, 21(16):3439–3440, 2005.
- [60] DUSEK, RACHEL L. und ATTARDI, LAURA D.: *Desmosomes: New Perpetrators in Tumour Suppression*. *Nature Reviews Cancer*, 11(5):317–323, 2011.

- [61] EKIMLER, SEMIH und SAHIN, KANIYE: *Computational Methods for microRNA Target Prediction*. Genes, 5(3):671–683, 2014.
- [62] ERZ, GREGOR et al.: *A Region and Contour Based Technique for Automatic Detection of Tomato Roots in Minirhizotron Images*. Journal of Applied Botany and Food Quality, 79:83–88, 2005.
- [63] FEINBERG, ANDREW: *DNA Methylation in Cancer: Three Decades of Discovery*. Genome Medicine, 6(5):36, 2014.
- [64] FELEKKIS, KYRIACOS et al.: *MicroRNAs: A Newly Described Class of Encoded Molecules that Play a Role in Health and Disease*. Hippokratia, 14(4):236–240, 2010.
- [65] FLAHAUT, M. et al.: *The Wnt Receptor FZD1 Mediates Chemoresistance in Neuroblastoma Through Activation of the Wnt/ β -Catenin Pathway*. Oncogene, 28(23):2245–2256, 2009.
- [66] FLYNN, KEVIN C.: *The Cytoskeleton and Neurite Initiation*. Bioarchitecture, 3(4):86–109, 2013.
- [67] FRANSSON, SUSANNE, MARTINSSON, TOMMY und EJESKÄR, KATARINA: *Neuroblastoma Tumors with Favorable and Unfavorable Outcomes: Significant Differences in mRNA Expression of Genes Mapped at 1p36.2*. Genes, Chromosomes and Cancer, 46(1):45–52, 2007.
- [68] FRICKE, ALBA et al.: *Identification of a Blood-Borne miRNA Signature of Synovial Sarcoma*. Molecular Cancer, 14(151):1–13, 2015.
- [69] FUJITA, TOMOYUKI et al.: *CHD5, a Tumor Suppressor Gene Deleted From 1p36.31 in Neuroblastomas*. Journal of the National Cancer Institute, 100(13):940–949, 2008.
- [70] GATTA, GEMMA et al.: *Childhood Cancer Survival in Europe 1999-2007: Results of EURO CARE-5 - a Population-Based Study*. The Lancet Oncology, 15(1):35–47, 2014.
- [71] GE, KAI: *Epigenetic Regulation of Adipogenesis by Histone Methylation*. Biochimica et Biophysica Acta, 1819(7):727–732, 2012.
- [72] GEYIK, ESRA et al.: *Investigation of the Association Between ATP2B4 and ATP5B Genes with Colorectal Cancer*. Gene, 540(2):178–182, 2014.
- [73] GLASER, EDMUND M. und LOOS, H. F. MACHIEL VAN DER: *A Semi-Automatic Computer-Microscope for the Analysis of Neuronal Morphology*. IEEE Transactions on Biomedical Engineering, 12(1):22–31, 1965.
- [74] GRIFFITHS-JONES, SAM: *The microRNA Registry*. Nucleic Acids Research, 32:D109–D111, 2004.
- [75] GRIFFITHS-JONES, SAM et al.: *miRBase: microRNA Sequences, Targets and Gene Nomenclature*. Nucleic Acids Research, 34:D140–D144, 2006.
- [76] GRIFFITHS-JONES, SAM et al.: *miRBase: Tools for microRNA Genomics*. Nucleic Acids Research, 36:D154–D158, 2008.
- [77] GRIFFITHS-JONES, SAM et al.: *Two Members of the TRiC Chaperonin Complex, CCT2 and TCP1 are Essential for Survival of Breast Cancer Cells and are Linked to Driving Oncogenes*. Experimental Cell Research, 332(2):223–235, 2015.
- [78] GU, JIAN-JUN et al.: *MicroRNA-130b Promotes Cell Proliferation and Invasion by Inhibiting Peroxisome Proliferator-Activated Receptor- γ in Human Glioma Cells*. International Journal of Molecular Medicine, 37(6):1587–1593, 2016.
- [79] GU, YIXUE et al.: *MicroRNA-493 Suppresses Tumor Growth, Invasion and Metastasis of Lung Cancer by Regulating E2F1*. PLoS ONE, 9(8):e102602, 2014.
- [80] GUIMARAES, GEOFFREY J. und DELUCA, JENNIFER G.: *Connecting with Ska, a Key Complex at the Kinetochore-Microtubule Interface*. The EMBO Journal, 28(10):1375–1377, 2009.
- [81] HA, MINJU und KIM, V. NARRY: *Regulation of microRNA Biogenesis*. Nature Reviews Molecular Cell Biology, 15(8):509–524, 2014.
- [82] HAGEMANN, CARSTEN et al.: *Expression Analysis of the Autosomal Recessive Primary Microcephaly Genes MCPH1 (Microcephalin) and MCPH5 (ASPM, Abnormal Spindle-Like, Microcephaly Associated) in Human Malignant Gliomas*. Oncology Reports, 20(2):301–308, 2008.

- [83] HAMFJORD, JULIAN et al.: *Differential Expression of miRNAs in Colorectal Cancer: Comparison of Paired Tumor Tissue and Adjacent Normal Mucosa using High-Throughput Sequencing*. PLoS ONE, 7(4):e34150, 2012.
- [84] HANAHAH, DOUGLAS und WEINBERG, ROBERT A.: *Hallmarks of Cancer: The Next Generation*. Cell, 144(5):646–674, 2011.
- [85] HAUSSER, JEAN und ZAVOLAN, MIHAELA: *Identification and Consequences of miRNA-Target Interactions beyond Repression of Gene Expression*. Nature Reviews Genetics, 15(9):599–612, 2014.
- [86] HEERBOTH, SARAH et al.: *EMT and Tumor Metastasis*. Clinical and Translational Medicine, 4(6):1–13, 2015.
- [87] HENRICH, KAI-OLIVER et al.: *CAMTA1, a 1p36 Tumor Suppressor Candidate, Inhibits Growth and Activates Differentiation Programs in Neuroblastoma Cells*. Cancer Research, 71(8):3142–3151, 2011.
- [88] HO, SHINN-YING et al.: *NeurphologyJ: An Automatic Neuronal Morphology Quantification Method and its Application in Pharmacological Discovery*. BMC Bioinformatics, 12(1):230, 2011.
- [89] HORVATH, STEVE et al.: *Analysis of Oncogenic Signaling Networks in Glioblastoma Identifies ASPM as a Molecular Target*. Proceedings of the National Academy of Sciences of the United States of America, 103(46):17402–17407, 2006.
- [90] HU, SHAO WEI et al.: *Expression of Zinc-Fingers and Homeoboxes 2 in Hepatocellular Carcinogenesis: a Tissue Microarray and Clinicopathological Analysis*. Neoplasma, 54(3):207–211, 2007.
- [91] HU, ZIHUA: *Insight into microRNA Regulation by Analyzing the Characteristics of their Targets in Humans*. BMC Genomics, 10(1):594, 2009.
- [92] HUANG, MILLER und WEISS, WILLIAM A.: *Neuroblastoma and MYCN*. Cold Spring Harbor Perspectives in Medicine, 3(10):1–22, 2013.
- [93] HUANG, WEIHUA, LI, HUIFEN und LUO, RONGCHENG: *The microRNA-1246 Promotes Metastasis in Non-Small Cell Lung Cancer by Targeting Cytoplasmic Polyadenylation Element-Binding Protein 4*. Diagnostic Pathology, 10(1):127, 2015.
- [94] HUNG, PEI-FANG et al.: *The Motor Protein KIF14 Inhibits Tumor Growth and Cancer Metastasis in Lung Adenocarcinoma*. PLoS ONE, 8(4):e61664, 2013.
- [95] HÜTTELMAIER, STEFAN et al.: *Spatial Regulation of Beta-Actin Translation by Src-Dependent Phosphorylation of ZBP1*. Nature, 438:512–515, 2005.
- [96] INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM: *Initial Sequencing and Analysis of the Human Genome*. Nature, 409(6822):860–921, 2001.
- [97] JIA, XIAOTING: *MiR-493 Mediated DKK1 Down-Regulation Confers Proliferation, Invasion and Chemo-Resistance in Gastric Cancer Cells*. Oncotarget, 7(6):7044–7054, 2016.
- [98] JIANG, LILI et al.: *Hsa-miR-125a-3p and hsa-miR-125a-5p are Downregulated in Non-Small Cell Lung Cancer and have Inverse Effects on Invasion and Migration of Lung Cancer Cells*. BMC Cancer, 10(1):1–13, 2010.
- [99] JIANG, PENG und COLLIER, HILARY: *Functional Interactions Between microRNAs and RNA Binding Proteins*. MicroRNA, 1(1):70–79, 2012.
- [100] JONAS, STEFANIE und IZAURRALDE, ELISA: *Towards a Molecular Understanding of microRNA-Mediated Gene Silencing*. Nat Rev Genet, 16(7):421–433, 2015.
- [101] JUÀREZ-MÉNDEZ, SERGIO et al.: *Splice Variants of Zinc Finger Protein 695 mRNA Associated to Ovarian Cancer*. Journal of Ovarian Research, 6(61):1757–2215, 2013.
- [102] JUN, GYUNGAH et al.: *PLXNA4 is Associated with Alzheimer Disease and Modulates Tau Phosphorylation*. Annals of Neurology, 76(3):379–392, 2014.
- [103] KAPLAN, EDWARD L. und MEIER, PAUL: *Nonparametric Estimation from Incomplete Observations*. Journal of the American Statistical Association, 53(282):457–481, 1958.

- [104] KAPUSCINSKI, JAN: *DAPI: A DNA-Specific Fluorescent Probe*. Biotechnic Histochemistry, 70(5):220–233, 1995.
- [105] KARANTZA, VASSILIKI: *Keratins in Health and Cancer: More than Mere Epithelial Cell Markers*. Oncogene, 30(2):127–138, 2011.
- [106] KASS, MICHAEL, WITKIN, ANDREW P. und TERZOPOULOS, DEMETRI: *Snakes: Active Contour Models*. International Journal of Computer Vision, 1(4):321–331, 1988.
- [107] KIGEL, BOAZ et al.: *Plexin-A4 Promotes Tumor Progression and Tumor Angiogenesis by Enhancement of VEGF and bFGF Signaling*. Blood, 118(15):4285–4296, 2011.
- [108] KIM, DAEHWAN et al.: *TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions*. Genome Biology, 14(4):R36, 2013.
- [109] KIM, HAK HEE et al.: *Different Motif Requirements for the Localization Zipcode Element of β -Actin mRNA Binding by HuD and ZBP1*. Nucleic Acids Research, 43(15):7432–7446, 2015.
- [110] KIM, KWANG-MIN, SON, KILHO und PALMORE, G. TAYHAS R.: *Neuron Image Analyzer: Automated and Accurate Extraction of Neuronal Data from Low Quality Images*. Scientific Reports, 5(17062):1–12, 2015.
- [111] KIRBAS, CEMIL und QUEK, FRANCIS K. H.: *A Review of Vessel Extraction Techniques and Algorithms*. ACM Computing Surveys, 36(2):81–121, 2004.
- [112] KNELSON, ERIK H. et al.: *Stromal Heparan Sulfate Differentiates Neuroblasts to Suppress Neuroblastoma Growth*. The Journal of Clinical Investigation, 124(7):3016–3031, 2014.
- [113] KO, CHIUNG-YUAN et al.: *Epigenetic Silencing of CCAAT/Enhancer-Binding Protein δ Activity by YY1/Polycomb Group/DNA Methyltransferase Complex*. The Journal of Biological Chemistry, 283(45):30919–30932, 2008.
- [114] KOPPEN, ARJEN et al.: *Dickkopf-3 Expression is a Marker for Neuroblastic Tumor Maturation and is Down-Regulated by MYCN*. International Journal of Cancer, 122(7):1455–1464, 2008.
- [115] KOZOMARA, ANA und GRIFFITHS-JONES, SAM: *miRBase: Integrating microRNA Annotation and Deep-Sequencing Data*. Nucleic Acids Research, 39:D152–D157, 2011.
- [116] KOZOMARA, ANA und GRIFFITHS-JONES, SAM: *miRBase: Annotating High Confidence microRNAs using Deep Sequencing Data*. Nucleic Acids Research, 42:D68–D73, 2014.
- [117] KUHL, CARSTEN et al.: *CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets*. Analytical Chemistry, 84(1):283–289, 2011.
- [118] KWAK, SEO-YOUNG et al.: *Ionizing Radiation-Inducible miR-30e Promotes Glioma Cell Invasion through EGFR Stabilization by Directly Targeting CBL-B*. The FEBS Journal, 282(8):1512–1525, 2015.
- [119] LAMERS, FIEKE et al.: *Targeted BCL2 Inhibition Effectively Inhibits Neuroblastoma Tumour Growth*. European Journal of Cancer, 48(16):3093–3103, 2012.
- [120] LAMOUILLE, SAMY, XU, JIAN und DERYNCK, RIK: *Molecular Mechanisms of Epithelial-Mesenchymal Transition*. Nature Reviews Molecular Cell Biology, 15(3):178–196, 2014.
- [121] LANEVE, PIETRO et al.: *The Interplay Between microRNAs and the Neurotrophin Receptor Tropomyosin-Related Kinase C Controls Proliferation of Human Neuroblastoma Cells*. Proceedings of the National Academy of Sciences of the United States of America, 104(19):7957–7962, 2007.
- [122] LANGMEAD, BEN und SALZBERG, STEVEN L.: *Fast Gapped-Read Alignment with Bowtie 2*. Nature Methods, 9(4):357–359, 2012.
- [123] LASTOWSKA, MARIA et al.: *Identification of Candidate Genes Involved in Neuroblastoma Progression by Combining Genomic and Expression Microarrays with Survival Data*. Oncogene, 26(53):7432–7444, 2007.
- [124] LÄTHÉN, GUNNAR, JONASSON, JIMMY und BORGA, MAGNUS: *Blood Vessel Segmentation using Multi-Scale Quadrature Filtering*. Pattern Recognition Letters, 31(8):762–767, 2010.

- [125] LE, XIAO-FENG et al.: *The Roles of microRNAs in the Cancer Invasion-Metastasis Cascade*. *Cancer Microenvironment*, 3(1):137–147, 2010.
- [126] LEE, MINNKYONG et al.: *GNL3 and SKA3 are Novel Prostate Cancer Metastasis Susceptibility Genes*. *Clinical & Experimental Metastasis*, 32(8):769–782, 2015.
- [127] LEWIS, BENJAMIN P., BURGE, CHRISTOPHER B. und BARTEL, DAVID P.: *Conserved Seed Pairing, often Flanked by Adenosines, Indicates that Thousands of Human Genes are microRNA Targets*. *Cell*, 120(1):15–20, 2005.
- [128] LI, HENG et al.: *The Sequence Alignment/Map Format and SAMtools*. *Bioinformatics*, 25(16):2078–2079, 2009.
- [129] LI, MIAOXIN et al.: *PMCA4 (ATP2B4) Mutation in Familial Spastic Paraplegia*. *PLoS ONE*, 9(8):1–5, 2014.
- [130] LI, WEIDONG et al.: *Decreased Expression of miR-204 is Associated with Poor Prognosis in Patients with Breast Cancer*. *International Journal of Clinical and Experimental Pathology*, 7(6):3287–3292, 2014.
- [131] LIANG, JIANMING, MCINERNEY, TIM und TERZOPOULOS, DEMETRI: *United Snakes*. *Medical Image Analysis*, 10(2):215–233, 2006.
- [132] LIAO, YANG, SMYTH, GORDON K. und SHI, WEI: *featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features*. *Bioinformatics*, 30(7):923–930, 2014.
- [133] LIN, RUEY-JEN et al.: *MicroRNA Signature and Expression of Dicer and Drosha can Predict Prognosis and Delineate Risk Groups in Neuroblastoma*. *Cancer Research*, 70(20):7841–7850, 2010.
- [134] LIN, SHIH-YEH et al.: *ASPM is a Novel Marker for Vascular Invasion, Early Recurrence, and Poor Prognosis of Hepatocellular Carcinoma*. *Clinical Cancer Research*, 14(15):4814–4820, 2008.
- [135] LIN, SHUBIN und GREGORY, RICHARD I.: *MicroRNA Biogenesis Pathways in Cancer*. *Nat Rev Cancer*, 15(6):321–333, 2015.
- [136] LIN, ZIYING et al.: *Abnormal miRNA-30e Expression is Associated with Breast Cancer Progression*. *Clinical Laboratory*, 62(1–2):121–128, 2016.
- [137] LIU, X. et al.: *Deregulated Wnt/ β -Catenin Program in High-Risk Neuroblastomas Without MYCN Amplification*. *Oncogene*, 27(10):1478–1488, 2007.
- [138] LIU, XINRAN, GONG, HAO und HUANG, KUN: *Oncogenic Role of Kinesin Proteins and Targeting Kinesin Therapy*. *Cancer Science*, 104(6):651–656, 2013.
- [139] LOFFREDA, ALESSIA et al.: *RNA-Binding Proteins in the Regulation of miRNA Activity: A Focus on Neuronal Functions*. *Biomolecules*, 5(4):2363–2387, 2015.
- [140] LONGAIR, MARK H., BAKER, DEAN A. und ARMSTRONG, J. DOUGLAS: *Simple Neurite Tracer: Open Source Software for Reconstruction, Visualization and Analysis of Neuronal Processes*. *Bioinformatics*, 27(17):2453–2454, 2011.
- [141] LOWERY, LAURA ANNE und VAN VACTOR, DAVID: *The Trip of the Tip: Understanding the Growth Cone Machinery*. *Nature Reviews Molecular Cell Biology*, 10:332–343, 2009.
- [142] LU, JUN et al.: *MicroRNA Expression Profiles Classify Human Cancers*. *Nature*, 435(7043):834–838, 2005.
- [143] MA, LI et al.: *MiR-9, a MYC/MYCN-Activated microRNA, Regulates E-cadherin and Cancer Metastasis*. *Nature Cell Biology*, 12(3):247–256, 2010.
- [144] MA, MENG-TAO et al.: *MiR-487a Resensitizes Mitoxantrone (MX)-Resistant Breast Cancer Cells (MCF-7/MX) to MX by Targeting Breast Cancer Resistance Protein (BCRP/ABCG2)*. *Cancer Letters*, 339(1):107–115, 2013.
- [145] MA, MENG-TAO et al.: *MiR-487a Promotes TGF- β 1-Induced EMT, the Migration and Invasion of Breast Cancer Cells by Directly Targeting MAGT2*. *International Journal of Biological Sciences*, 12(4):397–408, 2016.
- [146] MALLAT, STÉPHANE: *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.

- [147] MARIONI, JOHN C. et al.: *RNA-Seq: An Assessment of Technical Reproducibility and Comparison with Gene Expression Arrays*. Genome Research, 18(9):1509–1517, 2008.
- [148] MARIS, JOHN M.: *Recent Advances in Neuroblastoma*. The New England Journal of Medicine, 362(23):2202–2211, 2010.
- [149] MARTIN, JOHN et al.: *Post-Transcriptional Regulation of Transforming Growth Factor Beta-1 by microRNA-744*. PLoS ONE, 6(10):e25044, 2011.
- [150] MARTIN, MARCEL: *Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads*. EMBnet.journal, 17(1):10–12, 2011.
- [151] MAYSTADT, ISABELLE et al.: *The Nuclear Factor κ B-Activator Gene PLEKHG5 is Mutated in a Form of Autosomal Recessive Lower Motor Neuron Disease with Childhood Onset*. American Journal of Human Genetics, 81(1):67–76, 2007.
- [152] MAZEH, HAGGI et al.: *The Diagnostic and Prognostic Role of microRNA in Colorectal Cancer - a Comprehensive Review*. Cancer, 4(3):281–295, 2013.
- [153] MAZUMDER, BARSANJIT, SESHADRI, VASUDEVAN und FOX, PAUL L.: *Translational Control by the 3'-UTR: The Ends Specify the Means*. Trends in Biochemical Sciences, 28(2):91–98, 2003.
- [154] MEGIORNI, FRANCESCA et al.: *A Sketch of Known and Novel MYCN-Associated miRNA Networks in Neuroblastoma*. Oncology Reports, 38(1):3–20, 2017.
- [155] MEI, HONG et al.: *The Roles of microRNAs in Neuroblastoma*. World Journal of Pediatrics, 10(1):10–16, 2014.
- [156] MEIJERING, ERIK: *Neuron Tracing in Perspective*. Cytometry, 77(7):693–704, 2010.
- [157] MEIJERING, ERIK et al.: *Design and Validation of a Tool for Neurite Tracing and Analysis in Fluorescence Microscopy Images*. Cytometry, 58(2):167–176, 2004.
- [158] MEISTER, GUNTER: *Argonaute Proteins: Functional Insights and Emerging Roles*. Nat Rev Genet, 14(7):447–459, 2013.
- [159] MESTDAGH, PIETER et al.: *The miR-17-92 microRNA Cluster Regulates Multiple Components of the TGF β -Pathway in Neuroblastoma*. Molecular Cell, 40(5):762–773, 2010.
- [160] MIAO, CHENKUI et al.: *The Significance of microRNA-148/152 Family as a Prognostic Factor in Multiple Human Malignancies: A Meta-Analysis*. Oncotarget, 8(26):43344–43355, 2017.
- [161] MIGNONE, FLAVIO et al.: *Untranslated Regions of mRNAs*. Genome Biology, 3(3):1–10, 2002.
- [162] MISIAK, DANNY et al.: *Automatic Analysis of Fluorescence Labeled Neurites in Microscope Images*. In: *Workshop on Applications of Computer Vision (WACV)*, Snowbird, UT, USA, 2009.
- [163] MISIAK, DANNY et al.: *Extraction of Protein Profiles from Primary Neurons using Active Contour Models and Wavelets*. Journal of Neuroscience Methods, 225:1–12, 2014.
- [164] MIYAMAE, MAHITO et al.: *Plasma microRNA Profiles: Identification of miR-744 as a Novel Diagnostic and prognostic Biomarker in Pancreatic Cancer*. British Journal of Cancer, 113(10):1467–1476, 2015.
- [165] MOGILYANSKY, ELENA und RIGOUTSOS, ISIDORE: *The miR-17/92 Cluster: A Comprehensive Update on its Genomics, Genetics, Functions and Increasingly Important and Numerous Roles in Health and Disease*. Cell Death and Differentiation, 20(12):1603–1614, 2013.
- [166] MOLENAAR, JAN J. et al.: *LIN28B Induces Neuroblastoma and Enhances MYCN Levels via let-7 Suppression*. Nature Genetics, 44(11):1199–1206, 2012.
- [167] MOLENAAR, JAN J. et al.: *Sequencing of Neuroblastoma Identifies Chromothripsis and Defects in Neuritogenesis Genes*. Nature, 483(7391):589–593, 2012.
- [168] MÖLLER, BIRGIT et al.: *MiToBo - A Toolbox for Image Processing and Analysis*. Journal of Open Research Software, 4(1):e17, 2016.
- [169] MONCLAIR, TOM et al.: *The International Neuroblastoma Risk Group (INRG) Staging System: An INRG Task Force Report*. Journal of Clinical Oncology, 27(2):298–303, 2009.

- [170] MORTAZAVI, ALI et al.: *Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq*. Nature Methods, 5(7):621–628, 2008.
- [171] MÜLLER, SIMON et al.: *IGF2BP1 Enhances an Aggressive Tumor Cell Phenotype by Impairing miRNA-Directed Downregulation of Oncogenic Factors*. Nucleic Acids Research, 46(12):6285–6303, 2018.
- [172] MURAKAMI, YOSHIKI et al.: *The Expression Level of miR-18b in Hepatocellular Carcinoma is Associated with the Grade of Malignancy and Prognosis*. BMC Cancer, 13(1):1–11, 2013.
- [173] MURRAY, MATTHEW J. et al.: *Solid Tumors of Childhood Display Specific Serum microRNA Profiles*. Cancer Epidemiol Biomarkers Prev, 24(2):350–360, 2015.
- [174] NAGEL, STEFAN et al.: *Transcriptional Deregulation of Homeobox Gene ZHX2 in Hodgkin Lymphoma*. Leukemia Research, 36(5):646–655, 2012.
- [175] NASS, DVORA et al.: *MiR-92b and miR-9/9* are Specifically Expressed in Brain Primary Tumors and can be used to Differentiate Primary from Metastatic Brain Tumors*. Brain Pathology, 19(3):375–383, 2009.
- [176] NEUMANN, STEFFEN, THUM, ANDREA und BÖTTCHER, CHRISTOPH: *Nearline Acquisition and Processing of Liquid Chromatography-Tandem Mass Spectrometry Data*. Metabolomics, 9(1):84–91, 2012.
- [177] NEUZILLET, CINDY et al.: *Targeting the TGF β Pathway for Cancer Therapy*. Cancer and Metastasis Reviews, 147:22–31, 2015.
- [178] NEWMAN, ERIKA A. und NUCHTERN, JED G.: *Recent Biologic and Genetic Advances in Neuroblastoma: Implications for Diagnostic, Risk Stratification, and Treatment Strategies*. Seminars in Pediatric Surgery, 25(5):257–264, 2016.
- [179] NIBLACK, WAYNE: *An Introduction to Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1986.
- [180] NICASTRO, GIUSEPPE et al.: *Mechanism of β -Actin mRNA Recognition by ZBP1*. Cell Reports, 18(5):1187–1199, 2017.
- [181] NINOMIYA, MASASHI et al.: *Distinct microRNAs Expression Profile in Primary Biliary Cirrhosis and Evaluation of miR505-3p and miR197-3p as Novel Biomarkers*. PLoS ONE, 8(6):1–11, 2013.
- [182] NISHIDA, NAOHIRO et al.: *MicroRNA-125a-5p is an Independent Prognostic Factor in Gastric Cancer and Inhibits the Proliferation of Human Gastric Cancer Cells in Combination with Trastuzumab*. Clinical Cancer Research, 17(9):2725–2733, 2011.
- [183] ONG, KOK HAUR et al.: *NeuronCyto II: An Automatic and Quantitative Solution for Crossover Neural Cells in High Throughput Screening*. Cytometry A, 89(8):747–754, 2016.
- [184] OSHER, STANLEY J. und FEDKIW, RONALD P.: *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 1 Auflage, 2002.
- [185] OSHLACK, ALICIA, ROBINSON, MARK D. und YOUNG, MATTHEW D.: *From RNA-Seq Reads to Differential Expression Results*. Genome Biology, 11(12):1–10, 2010.
- [186] OTA, KAZUSHIGE et al.: *The H3K27 Demethylase, Utx, Regulates Adipogenesis in a Differentiation Stage-Dependent Manner*. PLoS ONE, 12(3):e0173713, 2017.
- [187] OTSU, NOBUYUKI: *A Threshold Selection Method from Gray-Level Histograms*. Trans. on Systems, Man, and Cybernetics (SMC), 9(1):62–66, 1979.
- [188] PAJTLER, KRISTIAN W. et al.: *Expression of NTRK1/TrkA Affects Immunogenicity of Neuroblastoma Cells*. International Journal of Cancer, 133(4):908–919, 2013.
- [189] PANI, GIUSEPPE et al.: *MorphoNeuroNet: An Automated Method for Dense Neurite Network Analysis*. Cytometry A, 85(2):188–199, 2014.
- [190] PAUL, PRITHA et al.: *Gli1 Transcriptional Activity is Negatively Regulated by AKT2 in Neuroblastoma*. Oncotarget, 4(8):1149–1157, 2013.

- [191] PELLATT, DANIEL F. et al.: *Expression Profiles of miRNA Subsets Distinguish Human Colorectal Carcinoma and Normal Colonic Mucosa*. *Clinical and Translational Gastroenterology*, 7(3):e152, 2016.
- [192] PENG, HANCHUAN et al.: *V3D Enables Real-Time 3D Visualization and Quantitative Analysis of Large-Scale Biological Image Data Sets*. *Nature Biotechnology*, 28(4):348–353, 2010.
- [193] PENG, HANCHUAN, LONG, FUHUI und MYERS, GENE: *Automatic 3D Neuron Tracing using All-Path Pruning*. *Bioinformatics [ISMB/ECCB]*, 27(13):239–247, 2011.
- [194] PENG, RUI et al.: *MiR-30a Inhibits the Epithelial-Mesenchymal Transition of Podocytes through Downregulation of NFATc3*. *International Journal of Molecular Sciences*, 16(10):24032–24047, 2015.
- [195] PETERSON, SARAH M. et al.: *Common Features of microRNA Target Prediction Tools*. *Frontiers in Genetics*, 5(23):1–10, 2014.
- [196] POOL, MADELINE et al.: *NeuriteTracer: A Novel ImageJ Plugin for Automated Quantification of Neurite Outgrowth*. *Journal of Neuroscience Methods*, 168(1):134–139, 2008.
- [197] POWERS, JOHN T. et al.: *Multiple Mechanisms Disrupt the let-7 microRNA Family in Neuroblastoma*. *Nature*, 535(7611):246–251, 2016.
- [198] PRENZLER, FRAUKE et al.: *Functional Analysis of ZFP36 Proteins in Keratinocytes*. *European Journal of Cell Biology*, 95(8):277–284, 2016.
- [199] PUISIEUX, ALAIN, VALSESIA-WITTMANN, SANDRINE und ANSIEAU, STÉPHANE: *A Twist for Survival and Cancer Progression*. *British Journal of Cancer*, 94(1):13–17, 2006.
- [200] QI, LU und DING, YAN-QING: *Involvement of the CREB5 Regulatory Network in Colorectal Cancer Metastasis*. *Yi Chuan*, 36(7):679–684, 2014.
- [201] R CORE TEAM: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [202] RAAIJMAKERS, JONNE A. et al.: *RAMA1 is a Novel Kinetochore Protein Involved in Kinetochore-Microtubule Attachment*. *Journal of Cell Science*, 122(14):2436–2445, 2009.
- [203] RAMANI, PRAMILA et al.: *High Levels of Polo-Like Kinase 1 and Phosphorylated Translationally Controlled Tumor Protein Indicate Poor Prognosis in Neuroblastomas*. *Journal of Neuro-Oncology*, 125(1):103–111, 2015.
- [204] RAMJI, DIPAK P. und FOKA, PELAGIA: *CCAAT/Enhancer-Binding Proteins: Structure, Function and Regulation*. *Biochemical Journal*, 365(3):561–575, 2002.
- [205] RAMM, PETER et al.: *Automated Screening of Neurite Outgrowth*. *Journal of Biomolecular Screening*, 8(1):7–18, 2003.
- [206] RAMOS-ESPINOSA, P., ROJAS, E. und VALVERDE, MAHARA: *Differential DNA Damage Response to UV and Hydrogen Peroxide Depending of Differentiation Stage in a Neuroblastoma Model*. *NeuroToxicology*, 33(5):1086–1095, 2012.
- [207] RATH, OLIVER und KOZIELSKI, FRANK: *Kinesins and Cancer*. *Nature Review Cancer*, 12(8):527–539, 2012.
- [208] RIBEIRO, DIOGO et al.: *Regulation of Nuclear Hormone Receptors by MYCN-Driven miRNAs Impacts Neural Differentiation and Survival in Neuroblastoma Patients*. *Cell Reports*, 16(4):979–993, 2016.
- [209] RIES, JUTTA et al.: *MiR-186, miR-3651 and miR-494: Potential Biomarkers for Oral Squamous Cell Carcinoma Extracted from Whole Blood*. *Oncology Reports*, 31(3):1429–1436, 2014.
- [210] ROBINSON, MARK D., MCCARTHY, DAVIS J. und SMYTH, GORDON K.: *edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data*. *Genome Biology*, 26(1):139–140, 2010.
- [211] ROBINSON, MARK D. und OSHLACK, ALICIA: *A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data*. *Genome Biology*, 11(3):R25, 2010.

- [212] ROBINSON, MARK D. und SMYTH, GORDON K.: *Small-Sample Estimation of Negative Binomial Dispersion, with Applications to SAGE Data*. *Biostatistics*, 9(2):321–332, 2008.
- [213] RODRÍGUEZ-GONZÁLEZ, F. GERMÁN: *MicroRNA-30c Expression Level is an Independent Predictor of Clinical Benefit of Endocrine Therapy in Advanced Estrogen Receptor Positive Breast Cancer*. *Breast Cancer Research and Treatment*, 127(1):43–51, 2011.
- [214] ROMERO-CORDOBA, SANDRA L. et al.: *MiRNA Biogenesis: Biological Impact in the Development of Cancer*. *Cancer Biology & Therapy*, 15(11):1444–1455, 2014.
- [215] RU, YUANBIN et al.: *The multiMiR R Package and Database: Integration of microRNA-Target Interactions Along with their Disease and Drug Associations*. *Nucleic Acids Research*, 42(17):e133, 2014.
- [216] RYAN, JOSEPH et al.: *MicroRNA-204 Increases Sensitivity of Neuroblastoma Cells to Cisplatin and is Associated with a Favourable Clinical Outcome*. *British Journal of Cancer*, 107(6):967–976, 2012.
- [217] SALOMON, DAVID: *Curves and Surfaces for Computer Graphics*. Springer-Verlag, Berlin, Germany / Heidelberg, Germany / London, UK / etc., 2006.
- [218] SAND, MICHAEL et al.: *Expression of microRNAs in Basal Cell Carcinoma*. *British Journal of Dermatology*, 167(4):847–855, 2012.
- [219] SCARPATI, GIUSEPPINA DELLA VITTORIA et al.: *Analysis of Differential miRNA Expression in Primary Tumor and Stroma of Colorectal Cancer Patients*. *BioMed Research International*, 2014:1–8, 2014.
- [220] SCHILLING, FREIMUT H. et al.: *Neuroblastom-Früherkennung im Alter von einem Jahr in Deutschland*. *Dtsch Arztebl International*, 100(25):A–1739–A–1746, 2003.
- [221] SCHMITZ, SABINE K. et al.: *Automated Analysis of Neuronal Morphology, Synapse Number and Synaptic Recruitment*. *Journal of Neuroscience Methods*, 195(2):185–193, 2011.
- [222] SCHRAMM, ALEXANDER et al.: *Exon-Level Expression Analyses Identify MYCN and NTRK1 as Major Determinants of Alternative Exon Usage and Robustly Predict Primary Neuroblastoma Outcome*. *British Journal of Cancer*, 107(8):1409–1417, 2012.
- [223] SCHULTE, EVA C. et al.: *Rare Variants in PLXNA4 and Parkinson's Disease*. *PLoS ONE*, 8(11):1–10, 2013.
- [224] SCHULTE, JOHANNES H. et al.: *Deep Sequencing Reveals Differential Expression of microRNAs in Favorable versus Unfavorable Neuroblastoma*. *Nucleic Acids Research*, 38(17):5919–5928, 2010.
- [225] SHEN, HAILIN et al.: *Automatic Tracking of Biological Cells and Compartments using Particle Filters and Active Contours*. *Chemometrics and Intelligent Laboratory Systems*, 82(1-2):276–282, 2006.
- [226] SHENOUDA, SYLVIA K. und ALAHARI, SURESH K.: *MicroRNA Function in Cancer: Oncogene or a Tumor Suppressor?* *Cancer and Metastasis Reviews*, 28(369–378):1–12, 2009.
- [227] SHOLL, D. A.: *Dendritic Organization in the Neurons of the 19 Visual and Motor Cortices of the Cat*. *Journal of Anatomy*, 87(4):387–406, 1953.
- [228] SHUI, YIFANG et al.: *MiR-130b-3p Inhibits Cell Invasion and Migration by Targeting the Notch Ligand Delta-Like 1 in Breast Carcinoma*. *Gene*, 609:80–87, 2017.
- [229] SHUKLA, GIRISH C., SINGH, JAGJIT und BARIK, SAILEN: *MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions*. *Molecular and Cellular Pharmacology*, 3(3):83–92, 2011.
- [230] SLABY, ONDREJ: *MiR-190 Leads to Aggressive Phenotype of Neuroblastoma Through Indirect Activation of TrkB Pathway*. *Medical Hypotheses*, 80(3):325–326, 2013.
- [231] SOHN, BO HWA et al.: *Functional Switching of TGF- β 1 Signaling in Liver Cancer via Epigenetic Modulation of a Single CpG Site in TTP Promoter*. *Gastroenterology*, 138(5):1898–1908, 2010.
- [232] SPOONER, BRIAN S. und HOLLADAY, CARTER R.: *Distribution of Tubulin and Actin in Neurites and Growth Cones of Differentiating Nerve Cells*. *Cell Motility*, 1(2):167–178, 1981.

- [233] STANGEL, DANIELA et al.: *KIF20A Inhibition Reduces Migration and Invasion of Pancreatic Cancer Cells*. Journal of Surgical Research, 197(1):91–100, 2015.
- [234] STÖHR, NADINE und HÜTTELMAIER, STEFAN: *IGF2BP1: A Post-Transcriptional “Driver“ of Tumor Cell Migration*. Cell Adhesion and Migration, 6(4):312–318, 2012.
- [235] STRIEDER, VERENA und LUTZ, WERNER: *E2F Proteins Regulate MYCN Expression in Neuroblastomas*. Journal of Biological Chemistry, 278(5):2983–2989, 2003.
- [236] SU, ZHENQIANG et al.: *An Investigation of Biomarkers Derived from Legacy Microarray Data for their Utility in the RNA-Seq Era*. Genome Biology, 15(12):523, 2014.
- [237] SUBRAMANIAN, ARAVIND et al.: *Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles*. Proceedings of the National Academy of Sciences, 102(43):15545–15550, 2005.
- [238] SUN, DIANSHUI et al.: *The Predictive Value and Potential Mechanisms of miRNA-328 and miRNA-378 for Brain Metastases in Operable and Advanced Non-Small-Cell Lung Cancer*. Japanese Journal of Clinical Oncology, 45(5):464–473, 2015.
- [239] SUN, ZHAO et al.: *MicroRNA-1246 Enhances Migration and Invasion Through CADM1 in Hepatocellular Carcinoma*. BMC Cancer, 14(1):616, 2014.
- [240] SZABÓ, PETER M. et al.: *Integrative Analysis of Neuroblastoma and Pheochromocytoma Genomics Data*. BMC Medical Genomics, 5(48):1–19, 2012.
- [241] TAM, SHIRLEY, TSAO, MING-SOUND und MCPHERSON, JOHN D.: *Optimization of miRNA-Seq Data Preprocessing*. Briefings in Bioinformatics, 16(9):950–963, 2015.
- [242] TAN, YU-LE et al.: *MiR-744 is a Potential Prognostic Marker in Patients with Hepatocellular Carcinoma*. Clinics and Research in Hepatology and Gastroenterology, 39(3):359–365, 2015.
- [243] TANG, JIE-TING et al.: *MicroRNA 345, a Methylation-Sensitive microRNA is Involved in Cell Proliferation and Invasion in Human Colorectal Cancer*. Carcinogenesis, 32(8):1207–1215, 2011.
- [244] TANG, RUIXUE et al.: *Downregulation of miR-30a is Associated with Poor Prognosis in Lung Cancer*. Medical Science Monitor, 21:2514–252, 2015.
- [245] TAUTENHAHN, RALF, BÖTTCHER, CHRISTOPH und NEUMANN, STEFFEN: *Highly Sensitive Feature Detection for High Resolution LC/MS*. BMC Bioinformatics, 9(1):504, 2008.
- [246] TEPPOLA, HEIDI et al.: *Morphological Differentiation Towards Neuronal Phenotype of SH-SY5Y Neuroblastoma Cells by Estradiol, Retinoic Acid and Cholesterol*. Neurochemical Research, 41:731–747, 2016.
- [247] THOMSON, DANIEL W., BRACKEN, CAMERON P. und GOODALL, GREGORY J.: *Experimental Strategies for microRNA Target Identification*. Nucleic Acids Research, 39(16):6845–6853, 2011.
- [248] TIRUCHINAPALLI, DHANRAJAN M. et al.: *Activity-Dependent Trafficking and Dynamic Localization of Zipcode Binding Protein 1 and β -Actin mRNA in Dendrites and Spines of Hippocampal Neurons*. The Journal of Neuroscience, 23(8):3251–3261, 2003.
- [249] TISCHKOWITZ, MARC und HODGSON, S.: *Fanconi Anaemia*. Journal of Medical Genetics, 40(1):1–10, 2003.
- [250] TIWARI, VAIBHAV et al.: *Role for Nectin-1 in Herpes Simplex Virus 1 Entry and Spread in Human Retinal Pigment Epithelial Cells*. The FEBS Journal, 275(21):5272–5285, 2008.
- [251] TOPOL, AARON et al.: *Dysregulation of miRNA-9 in a Subset of Schizophrenia Patient-Derived Neural Progenitor Cells*. Cell Reports, 15(5):1024–1036, 2016.
- [252] ULIVI, PAOLA et al.: *Peripheral Blood miR-328 Expression as a Potential Biomarker for the Early Diagnosis of NSCLC*. International Journal of Molecular Sciences, 14(5):10332–1034, 2013.
- [253] UNDE, AMIT S., PREMPRAKASH, V. A. und SANKARAN, PRAVEEN: *A Novel Edge Detection Approach on Active Contour for Tumor Segmentation*. In: *2012 Students Conference on Engineering and Systems (SCES)*, Allahabad, Uttar Pradesh, India, 2012.
- [254] UYLINGS, HARRY B. M. und VAN PELT, JAAP: *Measures for Quantifying Dendritic Arborizations*. Network-Computation in Neural Systems, 13(3):397–414, 2002.

- [255] VALENTE, VALERIA et al.: *Modulation of HJURP (Holliday Junction-Recognizing Protein) Levels is Correlated with Glioblastoma Cells Survival*. PLoS ONE, 8(4):e62200, 2013.
- [256] VAN RIJSBERGEN, C. J.: *Information Retrieval*. Butterworth-Heinemann, Newton, MA, USA, 2 Auflage, 1979.
- [257] VANUYTSEL, KIM et al.: *FANCA Knockout in Human Embryonic Stem Cells Causes a Severe Growth Disadvantage*. Stem Cell Research, 13(2):240–250, 2014.
- [258] VARDABASSO, CHIARA et al.: *Histone Variants: Emerging Players in Cancer Biology*. Cellular and Molecular Life Sciences, 71(3):379–404, 2014.
- [259] VITRIOL, ERIC A. und ZHENG, JAMES Q.: *Growth Cone Travel in Space and Time: The Cellular Ensemble of Cytoskeleton, Adhesion, and Membrane*. Neuron, 73(6):1068–1081, 2012.
- [260] WANG, CHUNXI et al.: *EZH2 Mediates Epigenetic Silencing of Neuroblastoma Suppressor Genes CASZ1, CLU, RUNX3 and NGFR*. Cancer Research, 72(1):315–324, 2012.
- [261] WANG, DADONG et al.: *HCA-Vision: Automated Neurite Outgrowth Analysis*. Journal of Biomolecular Screening, 15(9):1165–1170, 2010.
- [262] WANG, LIFENG et al.: *Histone H3K27 Methyltransferase Ezh2 Represses Wnt Genes to Facilitate Adipogenesis*. Proceedings of the National Academy of Sciences, 107(16):7317–7322, 2010.
- [263] WANG, ZHONG, GERSTEIN, MARK und SNYDER, MICHAEL: *RNA-Seq: A Revolutionary Tool for Transcriptomics*. Nature Reviews Genetics, 10(1):57–63, 2009.
- [264] WEBER, AXEL et al.: *Coexpression of Insulin Receptor-Related Receptor and Insulin-Like Growth Factor 1 Receptor Correlates with Enhanced Apoptosis and Dedifferentiation in Human Neuroblastomas*. Clinical Cancer Research, 9(15):5683–5692, 2003.
- [265] WEINBERG, ROBERT A.: *The Biology of Cancer*. Garland Science, New York, 2007.
- [266] WESTERMARK, ULRICA K.: *The MYCN Oncogene and Differentiation in Neuroblastoma*. Seminars in Cancer Biology, 21(4):256–266, 2011.
- [267] WHITEFORD, CRAIG C. et al.: *Credentialing Preclinical Pediatric Xenograft Models using Gene Expression and Tissue Microarray Analysis*. Cancer Research, 67(1):32–40, 2007.
- [268] WINTER, JULIA et al.: *Many Roads to Maturity: microRNA Biogenesis Pathways and their Regulation*. Nat Cell Biol, 11(3):228–234, 2009.
- [269] WITTEN, DANIELA et al.: *Ultra-High Throughput Sequencing-Based Small RNA Discovery and Discrete Statistical Biomarker Analysis in a Collection of Cervical Tumours and Matched Controls*. BMC Biology, 8(58):1–14, 2010.
- [270] WOJCICKA, ANNA et al.: *Next Generation Sequencing Reveals microRNA Isoforms in Liver Cirrhosis and Hepatocellular Carcinoma*. The International Journal of Biochemistry & Cell Biology, 53:208–217, 2014.
- [271] WYNshaw-BORIS, ANTHONY: *Spindle Orientation: Timing is Everything*. Neuron, 79(2):211–213, 2013.
- [272] XIA, ZHIQIANG et al.: *Decreased Expression of miRNA-204-5p Contributes to Glioma Progression and Promotes Glioma Cell Growth, Migration and Invasion*. PLoS ONE, 10(7):1–19, 2015.
- [273] XIE, YUNLI et al.: *The Phosphatase PP4c Controls Spindle Orientation to Maintain Proliferative Symmetric Divisions in the Developing Neocortex*. Neuron, 79(2), 2013.
- [274] XU, CHENYANG und PRINCE, JERRY L.: *Snakes, Shapes, and Gradient Vector Flow*. IEEE Trans. on Image Processing, 7(3):359–369, 1998.
- [275] XU, LI-JUAN et al.: *Parallel mRNA and microRNA Profiling of HEV71-Infected Human Neuroblastoma Cells Reveal the Up-Regulation of miR-1246 in Association with DLG3 Repression*. PLoS ONE, 9(4):e95272, 2014.
- [276] XU, X. T. et al.: *MicroRNA Expression Profiling Identifies miR-328 Regulates Cancer Stem Cell-Like SP Cells in Colorectal Cancer*. British Journal of Cancer, 106(7):1320–1330, 2012.

- [277] XU, XIAOYIN und WONG, STEPHEN T.C.: *Optical Microscopic Image Processing of Dendritic Spines Morphology*. IEEE Signal Processing Magazine, 23(4):132–135, 2006.
- [278] YAMADA, KAZUYA et al.: *ZHX2 and ZHX3 Repress Cancer Markers in Normal Hepatocytes*. Frontiers in Bioscience, 14:3724–3732, 2009.
- [279] YU, YUE et al.: *Down-Regulation of miR-129-5p via the Twist1-Snail Feedback Loop Stimulates the Epithelial-Mesenchymal Transition and is Associated with Poor Prognosis in Breast Cancer*. Oncotarget, 6(33):34423–34436, 2015.
- [280] YUE, XUETIAN et al.: *Zinc Fingers and Homeoboxes 2 Inhibits Hepatocellular Carcinoma Cell Proliferation and Represses Expression of Cyclins A and E*. Gastroenterology, 142(7):1559–1570, 2012.
- [281] YUVA-AYDEMIR, YELIZ et al.: *MicroRNA-9: Functional Evolution of a Conserved Small Regulatory RNA*. RNA Biology, 8(4):557–564, 2011.
- [282] ZHANG, BAOHONG et al.: *MicroRNAs as Oncogenes and Tumor Suppressors*. Developmental Biology, 302(1):1–12, 2007.
- [283] ZHANG, H. L. et al.: *Neurotrophin-Induced Transport of a β -Actin mRNP Complex Increases β -Actin Levels and Stimulates Growth Cone Motility*. Neuron, 31(2):261–275, 2001.
- [284] ZHANG, QIAN et al.: *Role of microRNA-30c Targeting ADAM19 in Colorectal Cancer*. PLoS ONE, 10(3):1–14, 2015.
- [285] ZHANG, T. Y. und SUEN, C. Y.: *A Fast Parallel Algorithm for Thinning Digital Patterns*. Communications of the ACM, 27(3):236–239, 1984.
- [286] ZHANG, YONG et al.: *A Novel Tracing Algorithm for High Throughput Imaging Screening of Neuron-Based Assays*. Journal of Neuroscience Methods, 160(1):149–162, 2007.
- [287] ZHAO, LIFEN et al.: *MiR-493-5p Attenuates the Invasiveness and Tumorigenicity in Human Breast Cancer by Targeting FUT4*. Oncology Reports, 36(2):1007–1015, 2016.
- [288] ZHAO, XUEMEI et al.: *A Shared Surface of TBP Directs RNA Polymerase II and III Transcription via Association with Different TFIIB Family Members*. Molecular Cell, 11(1):151–161, 2003.
- [289] ZHOU, WEI et al.: *MiR-744 Increases Tumorigenicity of Pancreatic Cancer by Activating Wnt/ β -Catenin Pathway*. Oncotarget, 6(35):37557–37569, 2015.
- [290] ZOU, QIONG et al.: *Clinicopathological Features and CCT2 and PDIA2 Expression in Gallbladder Squamous/Adenosquamous Carcinoma and Gallbladder Adenocarcinoma*. World Journal of Surgical Oncology, 11(1):143, 2013.

Anhang

A. Neuronenmorphologie

Information/Merkmal	Beschreibung
file	Dateiname des aktuellen Bildes
neuronID	fortlaufende Nummer der im Bild enthaltenen Neuronen
neuronArea	Fläche des Neurons
neuriteID	fortlaufende Nummer der Neuriten des zugehörigen Neuron
nArea	Fläche des Neuriten
nBranches	Anzahl der Verzweigungspunkte (Branch) des Neuriten
nSpines	Anzahl der Vorstülpungen (Spines) des Neuriten
nEnds	Anzahl der Endpunkte des Neuriten
branchID	fortlaufende Nummer der Verzweigung im zugehörigen Neuriten
bLength	Länge der Verzweigung
bWidth	mittlere Breite entlang der Verzweigung
bsLength	Länge des Neuritenschafts, ohne Wachstumskrone
bsWidth	mittlere Breite des Neuritenschafts, ohne Wachstumskrone
bsArea	Fläche des Neuritenschafts, ohne Wachstumskrone
bcLength	Länge der Wachstumskrone
bcWidth	mittlere Breite der Wachstumskrone
bcArea	Fläche der Wachstumskrone
bcSpines	Anzahl an Vorstülpungen in der Wachstumskrone
scale	Skalierung der Pixel
unit	Einheit der Skalierung

Tabelle A.1.: Beschreibung der Bildinformationen und ermittelten morphologischen Merkmale des *NeuronAnalyzer2D* zur Anwendung der automatischen Analyse neuronaler Morphologie. Dazu gehören Angaben über die Längen, Flächen und Breiten der im Bild detektierten Neuronen, deren Neuriten sowie der Neuritenschäfte und Wachstumskronen. Die Angabe der Werte erfolgt dabei in Pixel.

B. Probeninformationen zu den Neuroblastomdaten

ID	INSS	MYCN-	Diagnose-	Überlebens-	Lauf	Analyse	
		Amplifikation	alter [Tage]	zeit [Tage]		RNA	miRNA
JB.1	1	-	356	2121	1	ja	ja
JB.2	1	-	1197	1456	1	ja	ja
JB.3	1	-	98	776	1	ja	ja
JB.4	1	-	298	590	1	ja	ja
JB.5	1	-	1852	1631	1	ja	ja
JB.6	1	-	95	54	1	ja	ja
JB.7	1	-	42	729	1	ja	ja
JB.8	1	-	254	509	1	ja	ja
JB.9	1	-	0	345	1	ja	ja
JB_10	1	-	32	185	1	ja	ja
JB_11	1	-	179	294	1	ja	ja
JB_31	1	-	589	1648	2	ja	ja
JB_32	1	-	75	1632	2	ja	ja
JB_33	1	-	102	1452	2	ja	ja
JB_12	2	-	300	2509	1	ja	ja
JB_13	2	-	729	704	1	ja	ja
JB_34	2	-	12	1844	2	ja	ja
JB_35	2	-	623	610	2	ja	ja
JB_36	2	-	1726	354	2	ja	ja
JB_37	2	-	492	995	2	-	ja
JB_14	3	-	217	2539	1	ja	ja
JB_15	3	ja	1022	478	1	ja	ja
JB_39	3	-	518	785	2	ja	ja
JB_40	3	-	282	1288	2	ja	ja
JB_41	3	-	271	2067	2	-	ja
JB_42	3	ja	310	103	2	ja	ja
JB_43	3	ja	1137	5548	2	ja	ja
JB_44	3	-	2551	558	2	ja	ja
JB_45	3	ja	1742	2898	2	ja	ja
JB_16	4	ja	447	203	1	ja	ja
JB_17	4	ja	279	889	1	ja	ja
JB_18	4	ja	1575	262	1	ja	ja
JB_19	4	ja	2045	679	1	ja	ja
JB_21	4	-	1443	969	1	ja	ja

JB_23	4	-	1383	1364	1	ja	ja
JB_24	4	-	1812	834	1	ja	ja
JB_25	4	-	1136	194	1	ja	ja
JB_20	4	-	970	1202	2	ja	ja
JB_22	4	-	849	424	2	ja	ja
JB_46	4	ja	196	539	2	ja	ja
JB_47	4	ja	1904	207	2	ja	ja
JB_48	4	-	-	-	2	ja	ja
JB_50	4	-	400	5801	2	ja	ja
JB_51	4	-	476	5298	2	ja	ja
JB_52	4	-	1644	3511	2	ja	ja
JB_53	4	-	619	3357	2	ja	ja
JB_54	4	-	497	2448	2	ja	ja
JB_56	4	-	1537	1235	2	ja	ja
JB_57	4	-	933	1445	2	ja	ja
JB_58	4	ja	279	889	2	ja	ja
JB_59	4	-	2135	1063	2	ja	ja
JB_60	4	-	633	454	2	ja	ja
JB_61	4	-	790	329	2	ja	ja
JB_62	4	ja	447	203	2	ja	ja
JB_26	4S	-	16	2482	1	ja	ja
JB_27	4S	-	95	665	1	ja	ja
JB_28	4S	-	0	560	1	ja	ja
JB_29	4S	-	6	388	1	ja	ja
JB_30	4S	-	65	211	1	-	ja
JB_63	4S	-	308	3935	2	-	ja
JB_64	4S	-	243	2294	2	ja	ja
JB_65	4S	-	41	1338	2	-	ja
JB_66	4S	-	133	627	2	ja	ja
JB_67	4S	-	107	3571	2	ja	ja
JB_68	4S	-	4	2494	2	ja	ja
JB_69	4S	-	297	2050	2	ja	ja

Tabelle B.2.: Übersicht über die einzelnen Neuroblastomproben und deren Merkmale aus den zwei Sequenzierungsläufen, die für die Transkriptomanalyse in dieser Arbeit verwendet wurden. Durch Störung bei der Sequenzierung wurden die Proben 30, 37, 41, 63 und 65 verworfen und gingen nicht mit in die RNA-Sequenzierung ein. Für die Probe 48 liegen keine klinischen Daten vor. Die Nummerierung der Proben ist nicht lückenlos fortlaufend, da Proben im Vorfeld verworfen wurden, so diese bspw. geringe RNA-Mengen oder Qualitäten aufwiesen.

C. Differentielle Genexpression

Auf eine tabellarische Auflistung aller erzeugten Expressionsdaten in dieser Arbeit wurde an dieser Stelle aus Platzgründen verzichtet. Die Tabellen der differentiell exprimierten mRNAs/microRNAs aus den Vergleichen der klinisch relevanten Stadien können aus dem digitalen Anhang dieser Arbeit entnommen werden. Die Dateien `Expression_mRNAs.csv` und `Expression_microRNAs.csv` im Verzeichnis `Transkriptomanalyse/Expression/` beinhalten die normalisierten Expressionswerte sowie die Ergebnisse der differentiellen Expressionsanalyse.

D. Ermittelte Gensignaturen

D.1. Signaturgene

Die Signaturgene der jeweiligen Up- und Down-Sets sind in der nachfolgenden Tabelle D.3 aufgeführt. Der digitalen Anhang enthält zudem eine Datei der Gensignaturen mit der Bezeichnung `Neuroblastom_INSS_Signature.gmt` im Verzeichnis `Transkriptomanalyse/Gensignatur/`.

Signatur	Set	Genes
NB.metas	UP	KRT14, KRT10, KRT5, FAM49A, FLG, DSC3, CCT2, DSP, PERP, CPSF6, LIX1, SEMA3D, WDR43, MDM2, LPIN1, LCLAT1, FAM84A, FRS2, RPL39, RPL17, LIN28B, SEMA3A, IGBP1, PHF6, CTD-2192J16.20, POLA1, RPL31
NB.metas	Down	FAM124A, SPTB, NYAP2, C11orf96, NR4A1, ZFP36, DUOX2, RASD1, MT2A, HMOX1, HEPACAM2, IL6, AC008271.1, VIP, UTS2
NB.MYCN	UP	ODC1, DDX1, PENK, HOXD10, MYCN, FAM49A, TMPRSS15, CCT2, GALNT14, TM4SF4, ANO3, LIX1, NEUROG2, CPSF6, ZNF695, NCAN, SLC6A5, NBAS, WDR43, ZNF560, DNAJC5G, PYDC2, SIX2, YEATS4, HIST2H2AB, C21orf37, LRRC10, HIST2H3D, ALK, CPNE7, VAX2, HIST1H2AJ, HIST1H2AI, SKA3, HOXD11, TRIM54, ADCY8, PRDM8, PTTG1, ASPM, HIST1H3H, HIST1H3C, KIF20A, ADAMTS19, CBLN1, PITX1, KCTD8, HIST1H2AL, HIST1H2BH, SGCZ, BOLL, HIST1H2AH, RPL22L1, CENPF, HIST1H2BM, ST6GAL2, ISLR2, LIPF, CDCA7L, KCNF1, HIST1H2AB, DLX6, GABRA3, HIST1H2BL, TWIST1, CDKN3, HIST1H3G, HIST1H2BI, FAM19A5, IRX6, PHGDH, HIST1H4L, RPS17L, OPRD1, GRB14, PRSS12, UBE2C, DDN, DLX5, ABCA12, CLIP4, TRIM59, NPW, MDM2, HIST1H3B, HIST1H2BC, CHODL, HIST3H2A, SLC30A3, BIRC5, MCM10, KIF4A, MAGEA10, HIST1H1D, HIST1H3D, NOS1, KIF18A, STXBP6, HIST2H2AA4, PRLHR, BUB1B, LPIN1, HIST1H2BE, RPS25, HIST1H1B, MAGEA9, KIF14, LCLAT1, CCNB2, DSCC1, HMMR, HIST1H4C, HIST1H2AG, HIST2H2AA3, HIST1H2BF, UBE2T, EXO1, NPAS4, MAGEA9B, HIST1H3A, HIST1H2BD, JPH1, TEX15, MAD2L1, HIST2H2BF, MTHFD2, TYMS, HIST1H3F, KIF15, GAD1, BZW2, RPP21, NCAPH, CCNB1, CLMP, NCR3LG1, GLRA2, CDT1, PPA1, HIST1H3J, CCNA2, DBF4, PLAG1, DEPDC1B, SHC3, H2AFX, TRIB2, TMEM97, GINS2, TTK, IRX5, XCL1, ARHGAP19-SLIT1, KIAA0101, SH3GL3, ASS1, HIST1H4D, RPL17, FAM72D, CENPU, CCNB1IP1, TLL2, MCM2, CENPH, DHFR, HIST2H4B, HIST1H2BK, SUSD5, HIST2H4A, ANKRD18B, CLEC4F, WDR76, NME1, LRIT3, EBF3, PLK1, HMGB3, HIST2H2AC, HOXD9, TRIP13, HIST1H2BG, CCDC150, NXPH2, RP11-379H8.1, CHEK1, CCT8, HIST1H2AK, NUSAP1, NCAPG, UXT, ORIQ1, HIST1H3I, HIST1H2BB, LPPR5, LBH, TRMT10C, KIF11, HIST2H3A, CRYGC, SLC35F2, CUX2, HIST2H3C, DSCAM, CKS2, FANCI, RPS4X, SLC38A5, RPL10A, HIST1H1C, TOP2A, TPX2, DIAPH3, DTL, CDK1, FRS2, RCOR2, HIST1H2BJ, NTRK2, STRA6, YPEL5, MRPL33, DNAJC25, GJD4, FAM83D, DLL3, RPS24, FAM160A1, NEIL3, PBK, KCTD14, SLC27A2, RBF3, CDCA2, CDC6, SLC16A1, RPS13, MRPL11, GARI, LSM3, ASF1B, SAAL1, E2F3, SGOL1, SLC29A2, DSCAM1L1, AURKB, ECT2, HJURP, RRM2, MPV17L, CKS1B, C19orf48, AHCY, ZNF98, PRTFDC1, GNPAT1, RPS6, BCHE, ASCL1, MYBL2, RPS14, C21orf90, RPL39, MRPL3, GMNN, FAM111B, C8orf31, RPL34, VARS, TMA16, FAM13A, TK1, PUS7, E2F8, RPS11, C4orf46, SEC61G, ZBTB12, GSG2, TMEM126A,

HIST1H2AE, RPL29, CMBL, IRX3, AURKA, LGR5, PLEKHG4B, RPS21, RRM1, LDHB, CAMKV, HIST1H2AC, LAYN, FBXO5, RAD51AP1, HIST1H2BO, MCM6, BHLHA15, HIST1H4I, BUB1, C17orf82, BTBD11, ITPRIPL1, GABBR2, GMD5, HIST1H4B, MELK, RPL23, RPL35A, C14orf39, HIST1H4H, HIST3H2BB, H2AFZ, C8orf59, NPM3, CDC45, HMGB2, RMI2, TENM4, CEP44, RPSAP58, PSAT1, MYOZ3, FEN1, NEK2, CHAF1A, ORC1, ADRA1A, RPL38, DNA2, GTF2A2, POLE2, RPS26, KIF24, CEP152, ADAMTS16, CDK2, FANCD2, GFRA2, EME1, KNSTRN, SORCS3, ZNRD1, KIT, NDC80, ZNF239, APEX1, RAB33A, SGOL2, RPL35, GNB2L1, E2F1, MKI67, CDCA7, MIS18A, EIF2A, ATAD2, GLYATL2, SLC35G1, FANCA, IGF2BP1, DKC1, GAL, RPL32, RPLP2, RPL31, SAPCD2, ERCC6L, BNC2, NETO2, CDCA5, PATE4, PLD6, FOXM1, CPSF3, HIST2H2BE, SLC25A5, RPS7, DLGAP5, PPEF1, RPS3A, RNASEH2A, PTRHD1, MRPS9, RBBP8, RPS29, FBP1, ATP11C, SAP30, POC1A, POLE3, PPAT, TAX1BP3, BAI1, ANKRD37, SKP2, PDK1, EIF3M, RPS5, ALDH18A1, EHD3, GTF3C2, VWCE, BCL11A, C11orf1, FAM72A, KIFC1, CENPA, CHCHD3, TICRR, GPR12, C11orf82, RPL5, RFC4, LNP1, CDC25A, DPY19L2, CENPN, ZNF90, LRRN1, RPL39L, EIF3E, ZNF724P, POLA1, HIST1H4A, NPM1, ZNF280C, PHF6, ADM, TONSL, ACN9, HK2, MCM5, GINS1, IGBP1, EXOSC5, COCH, BRCA2, FABP6, SFXN4, CDC20, KIF2C, SHCBP1, GPR125, ZWILCH, NFXL1, EFS, DARS, RPLP0, EDA2R, LRIG3, MCTS1, ZNF799, PLK4, ZDHHC22, RPL30, BRIP1, ZNF670, MPST, RPS20, PRIM1, C15orf61, FBL, ANLN, FAM64A, PBX3, POLD1, NIFK, DACH1, SCRT2, GSTO2, KCNT2, POP1, SF3B14, ARHGAP11A, MRRF, RSL1D1, HIST1H2BN, SRD5A3, TTF2, TIGD7, LYTLA1, BAMBI, NME7, DDX39A, MCM3, RPL12, LIN9, BRCA1, ACP6, IGF2BP2, ZNF639, ZNF519, SNRPD2, FAM151B, PRR11, SLIRP, SPIN2B, PAICS, PPL3, ACRV1, RACGAP1, MMADHC, ETV1, AIMP2, WDR3, PANK1, XRCC2, CCDC88C, DNASE1L2, DCAF4, ST6GALNAC3, RPS23, THG1L, ADAMTS17, PCNA, NONO, ZWINT, UBA2, GPN1, CHAF1B, WDR74, HIST1H1E, PDSS1, SVIP, MCM4, SLCO5A1, PRC1, ZNF726, SMO, NME1-NME2, ZNF367, RPS3, TRMT112, ZC4H2, RPL13, CBS, EXOSC9, CHML, NDUFB6, RPS19, EIF4EBP1, EEF1B2, KIAA0020, RPL10, BAZ1A, RPS6KA6, PRMT3, TNFRSF10B, CCDC15, IPO5, IQGAP3, COX7C, GGCCT, LYRM1, RPP25L, RPL36A, ABCC4, RFWD3, KARS, CNNM1, TP53, MRPL9, RPS2, RANBP1, HINT1, RUVBL1, HSPD1, THEM6, CKAP2L, RRP9, ZNF32, XRCC4, LRPPRC, RIBC2, HPDL, RASL10B, DNPH1, ZNF649, LAPTM4B, TMEM243, IMPDH2, KM-PA-2, RPL18, RPL36, GPR19, ESPL1, MRPL48, ABCE1, CDCA4, PHB2, TLL12, RPS17, SHMT2, FAM162A, DCTPP1, CSRP2, TRAP1, ACTL6A, NME4, CHCHD7, BTF3, HLTFF, WDR34, CSE1L, EBPL, SSRP1, PYCR1, CDC25C, SIGMAR1, PSMG1, DSN1, CHAC2, F12, POLR3K, RRP1B, C6orf48, CENPI, NHP2, NT5DC3, HMGAI, RFC3, WDR62, FAM72B, NOB1, NUP37, CREB3L4, UCHL3, POLA2, ZNF536, LSM2, GDAP1L1, RBMX, MARCH11, NAP1L1, SLC25A6, NR2C2AP, PTDSS1, CDC7, FBXO17, RNASEH2C, AEN, C2orf47, ILF2, KIF23, NUP107, RAVER2, SFXN2, COX16, UNG, ATIC, PSMG2, NGDN, CCNE1, ZNF581, ALDH1B1, PACRGL, PRMT5, URB2, ACP1, C11orf73, RPL37A, LDHA, SCML2, DNAH14, CCT3, EIF3H, MRPL44, LBR, FAM136A, MASTL, CCT7, CMSS1, VRK1, DAXX, ATP5L, CENPO, PPM1G, CKAP2, NOL10, CENPK, HAUS1, ALKBH2, RHOBTB3, SPC24, FSTL4, SMYD3, MARS2, FH, HELLS, ACBD6, STOML2, ZNF738, CLNS1A, GNL3, NDC1, TKT, CCNG1, FBXO8, ANAPC10, ANTXR2, RPL41, RRP1, RPL27, SSR4, TMEM216, ATP10D, SOX11, CCT6A, SCFD2, COX5A, C4orf21, SMC2, MTAP, PNMA6C, MGME1, DUT, HIGD2A, ZNF850, CTD-2192J16.20, RAD51, BOP1, THOC3, SLC1A5, NELFE, MTHFD1, RPL37, KIAA1524, ALG3, NUDT5, GTSE1, EIF3L, TIMM17B, BLM, GPX7, SPAG16, AIMP1, TOP1MT, CXCR6, CASC5, IARS, TMEM231, KPNA2, ARHGAP11B, NUP35, C14orf166, MRPL52, MIS18BP1, ICT1, ORC6, KDELCL1, PPM1D, NOA1, PCOLCE, MBNL3, METTL9, MRPL50, MRPS2, NABP2, SNRPD1, TACC3, PSRC1, JAGN1, EIF2S3, C14orf93, MSH2, PDIA6, NFYA, NCL, TOMM22, CCT5, SLC16A14, BARD1, KIF18B, NCAPD2, IGF2BP3, EHMT2, POLR3D, PAQR3, PNMA3, RPS18, MRPS27, PRMT1, RPL18A, RABEPK, ZFAND1, PHF5A, POLD2, CDK4, MSH6, TOMM20, TXNL4B, RPEL1, PDE3B, RTTN, C21orf59, RPS15A, TMEM108, OIP5, BEX2, MALSU1, BUB3, STRBP, TARS, SNRPA, CENPM, RPA3, DTNBP1, FANCF, NCBP1, GSTCD, COX41, SNRPB, RUVBL2, TMSB15B, FARSB, TMEM194A, RPL4, RPL7A, EIF3D, INCENP, TRIM28, SAE1, RPS16, METAP1, SLC39A10, LEO1, GLA, ADSL, POLR1E, GART, ANP32A, IMMPP1L, ECSIT, MAP6D1, LMNB1, GPR137B, C11orf48, NDUFA6, SDCCAG3, LY6E, TEX10, SPDL1, FANCM, ZNF121, MRPL51, RSL24D1, LIG1, MCM7, NUP160, LARS, TCF3, LMNB2, PES1, ZNF660, EPRS, GTF3C3, FZD5, HILPDA, ME2, CAD, MDH2, SEHL1, SSR2, HYLS1, RNF150, XRN2, GTF3A, ZNF124, IFT57, DPP7, CCDC77, GEMIN5, CCSER1, IGSF9, RPL27A, SMC4, LSM4, FAM27C, RPL15, GLTSCR2, CEP135, SF3B3, SLC25A3, QTRTD1, NUP93, DDX21, CIRH1A, SLC30A9, TBL3, UBE2S, GJC1, SLC3A2, PPIA, TMEM258, CCT4, SNX17, TOPBP1, INTS12, PHKA1, PNMA6D, IPO7, RPSA, ZNF605, TGIF2, CCDC101, SRSF7, LIAS, WDR46, PPIF, NCAPD3, HOXD3, BOLA2B, GDI2, RSR1, NUDT1, GAL3ST3, NUP155, PRR7, LTB4R, DIEXF, ZNF66, DOC2A, ZIK1, KIF22, CCFN, KATNB1, ZNF219, IDH2, GMPS, POLR2H, NOL8, IL17RD, POLQ, FEZ1, C1orf43, ZNF229, ACAA2, PLEKHG4, EEF1D, NUDT10, MRPL21, WDR53, TP53RK, MRPS11, NUDT21, USP3, NIF3L1, LCMT2, GEN1, ANKS6, FADS1, DPY19L3, RECQL4, SLC4A1AP, SLC25A15, DARS2, CPSF2, RPS15, CHCHD6, ZNF551, ZNF773, WEE1, ZNF512, GREB1, SPRTN, WDR36, EIF4B, SUV39H1, XPO5, AGPAT5, GINS3, TROAP, RPL3, NACA, ZNF682, MCM8, NUCKS1, MED12L, CNOT10, ZNF93, PHKA2, PKMYT1, UTP20, PPP1R14B, NFYB, ZNF620, ZNF169, CLPTM1L, PDHA1, FAM213A, TIAM1, PRKDC, CBX2, STIP1, GABRB3, TMEM206, PPP2R3B, JADE3, ERCC2, NUP205, TIMELESS, HSPA14, MCCC1, PSMD13, BID, C2orf43, COQ2, ZNF77, TAF4B, PHF21B, ST6GAL1, GFM2, IPO11, GNL3L, PA2G4, METTL12, ANAPC11, RNR3, DDX49, DDX31, NCBP2, HADH, MARC1, SRSF1, C10orf2, CHN1, PRR3, ZMYND19, TRAPPC2L, EXOSC7, QSOX2, SCHIP1, TMEM218, ZCCHC7, FKBP10, NOL11, CEP72, RFC5, ATP5G2, LRBA, C18orf21, MRPS6, TSNAX, KIAA0922, TXNL4A, SNX6, NNT, FKBP4, PNCK, INTS7, SET, TFAM, XYLB, BFAR, CSNK2A2, TBC1D30, MRPS7, TDP1, PTRH2, CBRI, ZFP82, VASH2, WDR12, OSBPL1A, DHPS, FAM156B, POLR1B, GPHN, TRA2B, AGBL5, ACPL2, UCK2, PDCD11, FAM156A, CTPS2, DYM, SLC7A5, NFIB, METAP2, MED12, C9orf3, NUBP2, CCP110, PSME4, ARMC6, WDR5, SENP1, PLCXD1, PARP1, SLC37A4, ZNHIT6, HAUS6, GEMIN6, PIH1D1, UMPS, SUPT16H, ZNF195, DNMT1, ZNF852, TOMM40, KXD1, PFAS, HNRNPAB, IAH1, HSPA4, C12orf73, RPL7, NDUFS3, CCDC85C, NSMF, DAP3, EIF3B, MPV17, TOP3B, CECR5, FLAD1, SMARCC1, KNTC1, SRPK1, CEP192, RPUSD4, EXOSC2, ERLIN1, GLB1L2, OCIAD2, STT3A, FTSJ3, HNRNPA1, DCAF13, PPRC1, ABAT, EEF2K, TCEB1, USP39, WDR54, DDX19B, EIF3C, TSTA3, DDX10, RAB40B, NUP210, PASK, NOP14, ZNF140, ZNFAR2, NMF3, LRRC8B, ZNF135, RBM38, RCCD1, SAMM50, CYFIP1, KIAA1958, PFKM, TSEN54, RMDN3, MDN1, EZH2, STRADA, ZNF610, MRPL27, DCLRE1C, INTS4, DGKH, NCAF2, PRPF3, COX15, URB1, WHSC1, PHACTR3, FAM60A, ANAPC1, SCRIB, SH3YL1, OLA1, CREBZF, INVS, FANCC, PTDSS2, NOL6, COQ9, EIF2B4, HEATR1, MRPL10, DHX33, CENPJ

NB.MYCN

Down

LRP1, FLII, STAT2, ACADVL, ZNF641, PLEKHM2, KIAA0754, RORA, THRA, ITSN2, AGO3, SH3PXD2B, ARHGAP31, SEPT4, SMIM14, HECTD3, SSBP3, CSF1R, WIPF2, CCN2L2, EFCAB14, ELK4, MAST2, LHFPL2, MOB3C, MINK1, TENC1, ZBTB48, C12orf23, NOTCH2, VPS13D, JUP, SERPINB6, DHDDS, POLR2A, WIPF1, FLCN, BCAP29, RIN2, SLC43A2, TP53NP2, LRRC32, HIVEP2, TIAF1, MYO1D, RNF19B, SPTB, GDFP5, FAM124A, FGD5, SKI, PPAP2B, ZBTB7A, PTPRF, KIAA0247, DNALI1, EFN2, KDM6B, CNTNAP1, TVP23C, CCDC92, PHLDB1, NLRP1, ZNF362, TPP1, NRBP2,

		<p>PRKACA, RERE, AHI1, PAQR7, KIAA0319L, NDEL1, TRIM67, SLC26A11, ZNF800, ITPKB, ALDH3A2, PPAD, MAPK11, MMP15, MTSS1, SRRM3, MYO1C, NEURL1, RPH3AL, PSD2, PRKAG2, RRBPI1, CTSZ, RASSF4, IL32, STAT3, PLEKHO2, TRIM62, CSF1, CPEB3, PPP1R15A, CD4, ZFAND3, HEYL, SLC9A1, ZEB2, EPAS1, SYNMI, KIF1B, NFIC, TTLL7, CPNE2, KIAA1462, CAMK2N1, ZBTB8A, NFIA, OPTN, KCNK3, CLCN6, PLBD2, EPB41L1, TMCC3, PLXNC1, TMEM130, NR2F2, FAIM2, SYNPO, CAMK2B, ZBTB16, SOBP, SREBF1, ZDHHC18, MERTK, ST3GAL1, PCLO, DAB2, NDRG4, KCNQ3, NCDN, PLEC, FNDC5, LAMA4, C1QC, CAPN5, CHST1, GPR123, TGM2, PBXIP1, FBXO30, VAMP3, FMN1, DMTN, OSBPL3, PGM2L1, ZBTB46, BCL6, CLSTN1, RAP1GAP2, ARSA, WASF2, SRCIN1, EHD2, PER1, TNS1, EPB41L4B, TYROBP, SCPEP1, SLC6A2, CMPK1, RREB1, NRP1, ATXN1, MS4A7, CORO2A, DLC1, PLAT, SPRN, FAM213B, CAPS2, ALPK3, H6PD, CHRN4, CBX7, C10orf10, TSHZ2, TNFRSF1A, TGFB1, TFEB, KLF6, SORBS3, VAT1, DOCK3, SELPLG, MARCH4, ZFP36L2, ZBTB4, PRDM2, SOX15, FAXDC2, ZNF423, TPRG1L, DDC, MATN2, KCNAB2, TGFB2, LTBP4, MPEG1, ZFPM2, ZBTB8B, CD9, WWTR1, VAMP2, TMEM150C, BRINP2, NRCAM, TCF7L2, SLCO2A1, RGMA, CYSTM1, SERPING1, PRELP, TOM1L2, ELN, LYNX1, FOSL2, APLNR, KIAA1683, SLCO2B1, ELL2, HID1, CYR61, DOC2B, MAFB, PEAR1, NFATC2, CLSTN3, MPZ, DARC, KNDCC1, CORO2B, GRASP, KIF5A, REEP6, ARHGAP23, RIMS3, B4GALT1, TBX20, MYADM, HS3ST5, CNR1, ECE1, SYT7, IFITM3, ITGA3, TNFRSF1B, SDK1, SGK1, CAMK2A, AXL, CDC42EP1, SORBS2, AHDC1, PLAGL1, ROBO2, TH, ABLIM3, ARHGEF10L, RAB15, ATP13A2, GPX3, GPNMB, KIAA1522, EFHD2, FGD6, SCN7A, CPEB4, KCNC4, TLN2, SERINC2, S100A11, HIVEP3, CXCL16, VSTM2A, GPS2, MAP7D1, LRUK, CDH6, PQLC2, GPR153, PDGFD, ITGA8, DUSP1, CADM3, ELFN1, CHRNE, NEBL, LMNA, MAF7, CD74, EBF1, NNMT, STSIA2, NXPH1, IFITM2, C1S, SCG2, FBXO44, ALDOC, LRTM1, ADCY1, EPHX1, MMP19, PAPS2, C1R, S100A6, DKK3, TPPP3, SERPINA5, YBX3, APBA1, CRYAB, TFAP2B, THBD, KIAA0513, C1orf134, ZHX2, IGFBP5, PODN, GLIS3, AHR, SYT2, EPB41L3, SLC18A1, GABARAPL1, TRPM3, RGS9, IGFBP4, SLC24A2, TMEFF2, CASP9, GFRA3, SDC3, MAP1A, SOX10, GRIN3A, PIK3CD, SYT5, UNC5D, C8orf4, CAMTA1, PRKCB, NAV3, LATS2, FOXO1, SPSP1, HAP1, CYB561, IL7, LIG3, WNT4, GADD45B, SNTB1, C2CD4C, CLEC3B, TRPV2, SOD3, MYCL, EMP1, MDGA1, MYC, ANXA2, CITED4, MYL9, THBS1, FABP3, FBLN5, CLU, PLEKHG5, PTPRH, KCNJ5, KLF2, KLRK1, ZFP36L1, SLCO3A1, TNFRSF25, RAMP3, SOCS3, KCNC3, ICAM1, HPCAL4, PRUNE2, C2CD4A, PVRL1, NR4A1, CCBE1, CACNA2D3, ADRB2, CALY, POF1B, APOE, LRRTM4, ATP2B4, C5AR1, PTGDS, DNER, SDC4, EPB41L4A, HMOX1, C2CD4B, CCL19, TNF, NT5C1A, XKR4, SOX6, POU4F2, RCAN3, DIRAS3, CRISPLD2, SPOCK2, TMEM255A, ECEL1, IL1R1, NGFR, AMIGO2, RASD1, SLC18A2, ANKFN1, C17orf107, ACHE, C3, CCL2, S100B, DBH, SNCG, HTR1B, FOSB, TNFAIP3, ZFP36, PMP22, NR4A3, CGCR, SV2B, EPHA5, KRT19, PIRT, HOXA5, PLA2G2D, CHD5, NR4A2, NOV, PLXNA4, TMEM132C, HSPB6, RAB3B, APOD, MT2A, DRD2, INSR, C7, NTRK1, MEI4, CYP11A1, IL6, CYP21A2, GSTA1, STAR, CYP11B1, HSD3B2, CYP17A1, UTS2</p>
NB_s4	UP	<p>TMPRSS15, HIST1H2BA, MC5R, ALX1, CD200R1L, TSPY10, PENK, FAM49A, MAGEA9, HOXD10, PSORS1C2, HBB, MAGEA9B, MAEL, DDX1, MAGEA10, CCT2, MYCN, GALNT14, ZNF695, DHRS2, TM4SF4, NEUROG2, TPRX1, CALB2, LRRC10, LIX1, MUC15, ANO3, OC90, FDCSP, NCAN, NOG, NAA11, HIST1H2AJ, NPAS4, HIST1H2AI, FIBCD1, HIST2H2AB, OR52B6, DEFB104A, ZNF560, IAPP, EPPK1, SKA3, CGB2, HIST1H3H, ODC1, ALX3, CPSF6, HIST1H2AL, FAM19A3, AC008271.1, OTOP3, HIST1H3C, HIST1H3B, HIST2H3D, BIRC5, LTA, KIR3DL1, SLC6A5, SIX2, MAGEA11, PTTG1, KCNF1, CDKN3, KCTD8, ASB5, REXO1L10P, TRIM54, REXO1L11P, KIF20A, GTSF1, C1orf65, KRTAP4-9, RGS13, PTEM, VAX2, ALAS2, ADAD2, DDN, WDR43, HIST1H2AG, RBBP8NL, HIST1H4C, HIST1H2AH, TUBA3C, HIST1H2BC, PXT1, DSCR4, BOLL, YEATS4, ASPM, KIF18A, PYY, HIST1H3D, SLC5A12, SP6, HIST1H2BM, C8G, HIST1H4D, HIST1H3J, HIST1H2AB, FGF5, HIST2H2BF, DSCC1, DNAJC5G, HIST1H2BB, OR2T2, HIST2H2AA4, KIF14, IRS4, POTE, TRPM5, TCHH, BX649567.1, TRIM59, CRYM, RBFOX3, USH1G, HIST1H1B, CBLN1, NBAS, MCM10, RP11-379H8.1, KIF4A, HIST1H2BL, ADCY8, ALK, SGCZ, HIST1H2BH, MAGEA4, FAM72D, HIST1H2BE, SCGB2A2, CCDC169, HIST2H2AA3, RPL22L1, TERT, SLC30A8, HIST1H2BI, PPBP, RRM2, HBA2, UBE2C, SLC32A1, HBA1, FOXJ1, MYH6, CLEC4M, DBF4, HIST1H4L, RP11-481A20.11, CENPF, STXBP6, HIST1H3F, BTLA, CPNE7, PRDM8, HIST1H2BJ, ANKRD35, RGS18, MAGEC1, BUB1B, HIST1H2BK, RP11-382J12.1, HIST1H3G, FLJ45079, WDR76, LRRC56, TTK, GNGT2, HIST1H2BF, SEMA3D, GLDN, HIST3H2A, CDCA7L, GLDC, XCL1, HIST3H2BB, HIST2H4B, CCNB2, MMD2, ASF1B, HMMR, UBE2T, COL25A1, HIST2H4A, C11orf85, SH3GL3, AANAT, KRTAP5-5, RPS17L, LRIT3, DIAPH3, DTL, CRB1, THEMIS, RP11-10A14.4, CCNA2, MORN2, AMBP, TMPRSS12, NCAPH, CD3D, KCNJ11, TOP2A, RP11-758M4.1, CHODL, CCDC144NL, PLAG1, MAD2L1, CDC6, HIST1H2BD, CXCL9, TKTL2, GABRA3, HIST1H3A, PRAME, RPP21, GIPR, CDH7, MTTT, CENPH, CENPU, HIST2H2AC, E2F8, LEFTY2, LGSN, FAM111B, ST6GAL2, MCM2, CLEC4F, RSPH1, HIST1H2BG, SEC61G, HIST2H3A, TEX15, TRMT10C, HIST2H3C, FAM53A, GINS2, TYMS, FOXD4L3, UXT, RBM20, DLX5, KIF11, STAB2, HIST1H2AK, CCNB1, KIAA0101, FAM84A, CCDC150, NUSAP1, GMNN, MUC6, KIF15, TRIM7, GRB14, THEGL, TBX1, NDC80, ALLC, JPH1, FAM153C, CHST6, EXO1, ELFN2, TK1, IRGM, NOS1, RPS25, ATP6V1C2, CKS1B, FAM72A, RMI2, CLIP4, XDH, MDM2, SGOL1, LCLAT1, RPL17, HIST1H1C, DSCAM, RDM1, LSM3, CRB2, HIST1H1D, OPRD1, FLJ20306, BZW2, NCAPG, ILDR2, NEK2, MELK, CDCA2, CD79A, SLC35G5, FAM153B, DEPDC1B, PPA1, USP44, ANKRD18B, RP11-383H13.1, ASCL1, SGK3, POLE2, LPIN1, TMEM97, EPS8L2, POTE, TMA16, TRIP13, DHFR, CDT1, GPR114, OR1Q1, ATAD2, EME1, TMEM184A, CKS2, FEN1, RP11-12J10.3, NME1, ORC1, TPX2, RIBC2, SLC6A11, TSPEAR, HIST1H2BO, MRPL11, RAD51AP1, FANCI, HIST1H2AE, PBK, MIS18A, HIST1H4I, ZWILCH, SLC35G1, CCT8, KIF2C, SUSD5, KIF6, CCDC87, PLK1, CDK1, C14orf39, AL355531.2, PHGDH, MTHFD2, SAAL1, GTF2A2, HIST1H2AC, EFCAB10, ZNF98, RPL39, C9orf116, NEIL3, FOXN4, BRIP1, FAM19A5, GARI, GSG2, DNA2, POLA1, ECT2, BRCA1, RPS13, MKRN3, SLC38A5, SYCE2, CMBL, RPS4X, HOXD9, AURKA, RRM1, MRPL3, KIF24, NXT2, POU6F2, TIFA, SKA1, C4orf46, RPL34, GSTO2, HJURP, HMGB3, RHBG, FANC2, SLC27A2, CENPN, FBXO5, HRSF12, KIF4B, RPL10A, C21orf90, HMGB2, CDC45, E2F3, STK33, CDK2, TNFSF4, GLYATL2, CDC20, RCOR2, RIN1, POC1A, DNAJC25, TST, LYPLA1, C17orf82, MCM6, HIST2H2BE, CLMP, SHCBP1, HIST1H3I, CTD-3203P2.2, AURKB, HIST1H3E, ARHGAP19-SLIT1, CDCA5, CHAC2, BUB1, RPS6, FAM83D, ACN9, RPL38, IQGAP3, PRR11, NOTUM, CRYCG, SGOL2, ETV2, H2AFX, LIN9, RPS14, SLC29A2, ZNF367, SLC16A1, GNPAT1, NOL10, MYBL2, CCNB1IP1, TLCD1, RAD51C, ANLN, CPSF3, MRPL33, BAI1, HIST1H4B, CHEK1, ZNF670, CENPA, FOLH1, NPM3, POP1, DLGAP5, PHF6, ZNF280C, CDC25C, C19orf48, RNASEH2A, RPS24, ICAM2, MRPS9, MPV17L, ZNF799, PTRHD1, ZNF639, MYOZ1, KNSTRN, TMEM182, IBA57-AS1, TMA7, RPL29, MKI67, NUP37, ERCC66, ITPRIPL1, IGF2BP1, SLC22A15, CHCHD7, RPS21, NUF2, H2AFX, FRS2, KIFCI, HIST1H4H, BTBD11, ZNF724P, RPS11, FOXM1, LBH, LIN28B, SRD5A3, SF3B14, ZBTB12, RIMBP3C, GFRA2, ALS2CR11, IGBP1, TMEM243, ESCO2, CAMKV, CDC25A, FAM72B, EIF3M, ELP4, GINS1, TAF9B, NDUFB6, RPS26, LNP1, RPSAP58, BCL11A, RBBP8, E2F1, RIMBP3B, SLIRP, POLE3, ACP6, PPM1D, ARHGEF39, DKC1, CHAF1B, SAPCD2, CDCA7, ZNF595, AUNIP, TSPAN3, PRIM1, RACKAP1, STRIP2, CHAF1A, CEP152, TAX1BP3, RPL23, AHCY, NME1-NME2, EIF2A, RPL39L, ANKRD37, TRMT10A, ORC6, NCR3LG1, CKAP2L, WDR62, LYRM1, APEX1, DSN1, MAB21L2, CHCHD3, VARS,</p>

		<p>GPR12, RIMBP3, ZWINT, RPL32, ZNRD1, TIMM17B, FAM13A, HIST1H1E, TRAPPC2P1, RPL35A, CCNE1, SEMA3A, HSD11B2, PANK1, ESPL1, LTB4R, CLSPN, TXNDC9, RPS7, EIF3E, RFC3, RPS29, C5, POLA2, PPAT, HINT1, RPLP2, BRCA2, CEP70, RFWD3, SKP2, XRCCA, UBA2, RFC4, SLC25A5, DNAH14, PRMT3, PLK4, KPNA2, C4orf21, XRCC2, CEP44, NUDT7, HIST1H2BN, WDR34, PSMG2, ILF2, RPL31, PUST7, GGCT, FANCA, SCML2, ZNF284, ZNF519, COCH, RPL5, MCM5, CDC7, RPS5, FSTL4, MRPL9, RPS3A, SMC2, LDHB, CENPI, UCHL3, ACP1, OIP5, GSTCD, RPLP0, TUBD1, KIF23, MRPL50, RPL35, PRC1, ZNF420, SAP30, FAM64A, PCNA, DARS, NGDN, TICRR, ZFAND1, GNB2L1, MCM4, MSH2, DNASE1L2, JMMP1L, NUP107, CEP135, DEPDC1, F12, ZNF90, CTD-2192J16.20, E2F2, CCHCR1, ICT1, POLD1, TOP1MT, VRK1, TONSL, RPA3, C3orf67, TEX14, DTD2, SUV39H2, C11orf48, RPL30, CXCR6, UQCRQ, GTF3C2, TIGD7, MCTS1, PTPMT1, ZNF850, KARS, MPST, THOC3, ZNF239, MRPL40, C2orf47, AKAP5, ZC2HC1C, VWCE, UBXXN10, SPRTN, MRPL44, LSM2, CE1L, ZNF738, CENPO, SPATA18, MCM3, THG1L, MASTL, TRMT112, WDR53, ARHGAP11A, WDR3, KIAA0020, PDSS1, EAPP, IPO5, NONO, MAP6D1, HAUS1, TTF2, TTC33, PACRGL, ALDH18A1, CCSER1, CCT3, AIMP2, CDCA4, BLM, SNRPD2, RANBP17, CCDC15, PDCD2L, ARHGAP11B, ISCA2, TCF19, NFXL1, ATP11C, HSD17B7, SPAG16, HLTF, EBPL, ZNF280B, RAVER2, INTS2, GPR19, PDK1, AC027307.3, MRPL52, NUDT5, NT5DC3, SSRP1, LRPPRC, SPIN2B, CSRFP2, RPP25L, RPS6KA6, RPS20, MRRF, SHQ1, SVIP, MRPL48, RPEL1, GPN1, CENPK, STRBP, MSH6, TMEM216, SMKR1, NIFK, SLC9A7, NDUFS8, ZC4H2, RNASEH2C, ANKRD32, GGACT, AEN, RANBP1, SLC25A19, CCNF, NETO2, ZNF230, CASC5, FOXN2, DPY19L2, SPC24, SFXN4, RPS23, PYCR1, ADAMT517, KIAA1524, RSL1D1, RPS15A, EXOSC5, PDE3B, ITGB3BP, COX16, IFT57, DUT, LBR, SUPT4H1, TOPBP1, APIP, STGALNAC3, PDIA6, PTDSS1, C7orf31, MALSU1, HELLS, AC004381.6, MARS2, GTF3A, LMNB1, THEM6, C12orf61, STIL, TACC3, PTPLB, ZNF649, TSEN54, WDR74, NUP35, CCDC88C, METTL9, NFYB, STOML2, CKAP2, NCAPG2, NABP2, CCDC18, NOL11, TP53RK, RTTN, RUVBL1, PAQR3, SUV39H1, DTNBP1, COPB1, RARS, RPL12, RRP1B, PSRC1, EEF1B2, KCNT2, ZNF620, ORC5, DDX39A, FH, GAL3ST3, GPX7, MRPL21, CCT6A, MYNN, SFXN2, SIGMAR1, PRMT5, LRRN2, CCT7, WEE1, GPR125, KIF18B, FBL, PPM1G, TXNL4B, GTSE1, PHB2, SAE1, RFC5, LRBA, ARL13B, ZIK1, ABCB7, ASRGL1, TMEM108, TMEM231, COQ6, POLD2, SSR4, ZNF530, RSRCL1, C14orf93, UTP18, GEN1, ZNF569, PPP2R3B, ACAA2, EIF3H, DAXX, RAD51, DCAF4, RASL10B, MRPS7, NFYA, TMEM194A, C18orf21, MARCH1, SPDL1, GMCL1, INCENP, UNG, CMSS1, ADSL, IPO11, MRPS11, NPM1, SEH1L, FADS1, TMEM223, ALKBH2, FAM27C, NCBP1, NUP160, LIG1, TBC1D31, SMC4, NUDT1, MCM7, RFC2, ACBD6, NOA1, ZNF773, SSR2, ZNF675, NOB1, MTHFD1, FARSB, COQ2, CLP1, GINS3, POLQ, ZNF605, LAGE3, MRPS33, C21orf59, PHF5A, STAG2, MRPS17, ME2, CLPTMIL, NAP1L1, URB2, ZNF121, JAGN1, ANAPC11, TEX10, GTF3C3, POLE4, GEMIN6, ALG3, SLC16A14, CCP110, RRP1, QTRTD1, GEMIN5, SRSF7, ZNF169, NELFE, STIPI1, PSMD13, DCTPP1, PDHA1, KNTC1, NDUFS3, SF3B3, PDP2, DBR1, SMYD3, MRPS2, IGF2BP3, CCDC47, SMG8, EIF2S3, ZNF77, SLC39A10, HAX1, RUVBL2, ATIC, HSPA14, AXIN2, TTLL12, NUP205, SLC39A11, MRPS6, CDK4, NOL8, GART, MRPS27, NUP93, CEP192, NCAPD2, SPECC1, EIF1AD, LRRC45, GFM2, CIRH1A, ZDHH13, FTSJ3, NIF3L1, INTS12, NUP155, SENP1, SKA2, LEO1, GJC1, KIF22, HILPDA, TMEM206, WDR46, CBX2, DCLRE1C, RAB40B, ZNF660, TP53, AGBL5, GABRB3, SLC4A1AP, LARS, RPS18, C2orf43, PHKA1, FIBP, RPL18A, PSMC3, METTL12, HADH, TRMT11, EHM2, DYM, TBL3, MRPL27, AP4E1, POLR1E, NUDT10, PIK3C2A, SNX6, SCFD2, TAF12, TOP3B, TDP1, SLC25A13, RPL27A, METAP1, CCZ1B, LMNB2, SMURF2, MED12L, NME4, TSTA3, RRN3, PIK3R4, C17orf80, PIH1D1, BRI3BP, TMSB15B, SNRPA, AKIP1, CHCHD6, CHTF8, ACP2, TRAPPC2L, TTF1, CNOT10, TXNL4A, CREB3L4, TMEM101, ATP5J2, GTF2H2, VPS25, CCDC77, MRPL16, USP3, CBRI, ZNF551, ANP32A, RBM38, DIEXF, TIMELESS, C12orf73, PASK, POLR3D, IGHMBP2, NUP98, ABAT, ALG1, ZFP82, DDX19B, SEC11A, MACROD1, ZMYND19, BOLA2B, TROAP, AZI1, SETD4, PHKA2, ADAT1, UMPS, TTC9C, IAH1, GLE1, WDR5, PSME4, CIT, RECQL4, PRR3, DAP3, MCM8, GPHN, DLG3, SLC25A15, DOC2A, CCZ1, EZH2, PET112, TCEB1, PA2G4, CSNK2A2, FBXW7, TRIM4, C18orf25, UTP20, NCBP2, ZFP30, ZNF619</p>
NB_s4	Down	<p>ATXN1, SORBS3, CAMK2B, GRAMD1B, NRBP2, FAXDC2, NDEL1, FAIM2, ZHX2, SRCIN1, ARHGEF10L, CNTNAP1, ALDH3A2, SDC3, SOBP, WSB1, CRTCC, RAI1, SYNPO, PLBD2, HIVEP3, ECE1, APBA1, FAT4, TRIM67, POLR2A, CAPN5, NR2F2, SPTB, MINK1, PLXDC2, EFAS1, ABR, MAP1LC3A, PPP1R15A, KIAA1522, MAP1A, EHD2, PTRF, PTPRN, WWTR1, TCF7L2, ZBTB16, TNS1, FAM124A, KIF5A, SREBF1, C1QC, ZBTB4, KCNQ3, TOM1L2, KRBA2, SPRY1, SGK1, GPR153, KDM6B, ATP2B4, ANXA2, GABARAPL1, MAP7D1, THY1, DOC2B, TNR, MCMAM, ADCY1, OLFML2A, RAB15, SPOCK2, KLF6, FBXO4A, MS4A7, PHLDB1, MDGA1, SH3BP5, PLEKHG5, LDLR, B4GALT1, CNR1, TNFRSF25, CREB5, ARHGAP23, CASP9, BGN, BRINP2, AXL, CYR61, PER1, NFIC, SPRN, C8orf4, ZEB2, BHLHE40, S100A11, SYT2, MAP7, SERPING1, IFITM3, PLACL1, PDGFD, FSTL3, TFAP2B, NXPH1, TIMP1, COL1A1, CTGF, LMNA, CDC42EP1, KIAA1683, IGFBP5, TRPV2, MYADM, ICAM1, INSR, APOE, SCARB1, ST8SIA2, SV2C, ZFP36L1, NNMT, PMP22, SNTB1, DNER, IGFBP4, DUOX2, PAPP, ELL2, C17orf107, SOCS3, C1orf233, SOX10, FOSL2, PLP1, MYL9, THBS1, C11orf96, CRYAB, EMP1, EPHX1, KCNJ5, HES4, CEBPD, FOSB, CHD5, PVRL1, IL1R1, CEBPB, CRISPLD2, FAM166B, EPB41L3, SERPINE1, PLXNA4, PI15, KLRK1, NR4A1, NR4A2, MEDAG, NR4A3, PIRT, LIF, NOV, ZFP36, RASD1, HOXA5, SFRP2, DRD2, HSPB6, PLA2G2A, HMOX1, AF000783.1, CYP21A2, MT2A, MTRNR2L2, IL6, GSTA1, APOD, CYP11A1, STAR, VIP, CYP17A1, CYP11B1, HSD3B2, UTS2</p>
NB_s4M ⁺	UP	<p>PENK, FAM49A, HOXD10, CCT2, DDX1, MYCN, TMPRSS15, GALNT14, ANO3, LIX1, TM4SF4, CPSF6, NEUROG2, SLC6A5, WDR43, NCAN, ZNF695, YEATS4, NBAS, ALK, ZNF560, PYDC2, LRRC10, CPNE7, DNAJC5G, HIST2H3D, ADCY8, ODC1, SGCZ, PRDM8, MDM2, CLIP4, HIST2H2AB, DLX5, LPIN1, HOXD11, VAX2, SIX2, ADAMTS19, HIST1H2AI, HIST1H2AJ, LCLAT1, CDC47L, AC008271.1, CBLN1, CHODL, HIST1H3H, DDN, HBB, PTTG1, HIST1H3C, SKA3, RPS17L, KIF20A, CENPF, HIST1H2BI, SLC30A3, TRIM54, GABRA3, ASPM, TWIST1, DLX6, TRIM59, HIST1H2BH, HIST1H3G, TRIB2, HIST1H2BM, PRSS12, KIF18A, RPS25, OPRD1, RPL22L1, DSCC1, ST6GAL2, HIST1H2BC, HIST1H2AB, HIST1H3B, HIST1H4L, HIST1H2BL, HIST1H4C, HIST1H1D, KIF4A, UBE2C, BUB1B, CDKN3, HIST1H2AL, ISLR2, ABCA12, PHGDH, HIST1H2AH, HIST1H3D, CLMP, HIST1H2BE, MTHFD2, KIF14, HIST3H2A, BZW2, PLAG1, MCM10, HIST1H2BF, LBH, FRS2, FAM19A5, NPW, SHC3, YPEL5, HIST1H1B, RPL17, HIST1H2BD, JPH1, HOXD9, NXPH2, UBE2T, CCNB2, CDT1, BIRC5, CCNB1, NPAS4, PPA1, ASS1, HIST1H3A, LPPR5, NCR3LG1, IRX6, DBF4, TMEM97, CCNB1IP1, MAD2L1, CCNA2, DNAJC25, HIST2H2AA4, BCHE, KIF15, GRB14, H2AFB, HIST1H3F, DSCAM, HIST2H2BF, RPS4X, SH3GL3, CXU2, TYMS, HIST1H4D, HMGB3, GINS2, EXO1, HIST1H3J, HIST2H2AA3, HIST1H2BB, CENPU, CENPH, DEPDC1B, NTRK2, HIST1H2AG, CCT8, RPL10A, HIST1H2BG, CEP44, HMMR, CKS2, PLK1, NCAPH, RPL39, TPX2, TRMT10C, MPV17L, NUSAP1, MRPL33, HIST1H1C, WDR76, DHFR, UXT, CMBL, BNC2, TTK, SUSD5, SLC16A1, HIST1H2BK,</p>

SLC35F2, GNPAT1, RPS24, EHD3, RPS6, TRIP13, MCM2, PSAT1, HIST2H2AC, HIST1H3I, NCAPG, TOP2A, CKS1B, E2F3, FAM72D, ARHGAP19-SLIT1, RPS14, PUST7, HIST2H4B, CAMK7, HIST2H4A, RPS11, RCOR2, HIST2H3A, NME1, AHCY, MRPL3, KIT, SLC27A2, HIST2H3C, RPL34, RPS13, LDHB, FAM13A, FAM83D, ECT2, PBK, TMA16, HIST1H2BJ, CHEK1, KCNF1, TMEM126A, VARS, HIST1H2AK, GSG2, RPL29, NEIL3, KIF11, SAAL1, MRPL11, CCDC150, GLRA2, DTL, FBXO5, HIST1H2BO, C19orf48, RPS7, HIST1H4H, RPL23, ITPRIPL1, ANKRD18B, FANCI, CDK1, RRM2, HJURP, LSM3, DIAPH3, GNB2L1, C4orf46, RPS21, SGOL1, CDCA2, GTF3C2, KIAA0101, RPL35, RPSAP58, RPL32, STRA6, BCL11A, HIST1H2AE, MYBL2, HIST1H4B, CDCA7, MCM6, IGF2BP1, RPS3A, RPL31, BUB1, EIF2A, HIST3H2BB, ZBTB12, RRM1, C17orf82, RPS26, LAYN, RPL38, APEX1, DLL3, RPL35A, RPL5, MIS18A, LRRN1, DKC1, FOXM1, GMDS, AURKA, CDC6, GMNN, HMGB2, PHF6, RBBP8, GFRA2, IGBP1, HIST1H2AC, ASF1B, NPM1, EIF3E, GTF2A2, PLD6, PPAT, RPLP2, MKI67, EIF3M, SCRT2, AURKB, SLC29A2, RPS5, ATP11C, MRPS9, NDC80, EBF3, KNSTRN, NPM3, PDK1, CPSF3, ATAD2, ZNRD1, RPLP0, RAD51AP1, FBL, NETO2, FEN1, RTRD1, ERCC6L, LNP1, ZNF280C, MELK, RSL1D1, E2F8, TENM4, SKP2, POLE3, SLC25A5, LRIG3, E2F1, ORC1, GPR125, RPL12, KIF24, CHCHD3, ZNF239, POLE2, RAB33A, RPL13, ADM, DNA2, HIST2H2BE, FANCA, RPS29, RNASEH2A, CEP152, DARS, RPS23, CHAF1A, ALDH18A1, GSTO2, RMI2, H2AFZ, MPST, COCH, FANCD2, RPS2, CDK2, EEF1B2, SEC61G, PRMT3, GPN1, ANLN, RPL30, FAM111B, RPS3, SFXN4, SAPCD2, DLGAP5, SRD5A3, RPS6KA6, PAICS, NIFK, AIMP2, RPS17, CDC25A, C11orf1, ZNF639, TK1, SLC25A6, C15orf61, RPS20, HK2, HSPD1, TICRR, HIST1H2BN, ZNF724P, POLA1, FBXO8, IPO5, GINS1, RPL10, NEK2, CDC45, RASL10B, PDSS1, EXOSC5, BAIL, HINT1, MRRF, RPL34, RAVER2, KARS, HIST1H1E, ACRV1, ATIC, ZNF649, CDCA5, ZNF581, HIST1H4A, LRPPRC, SAP30, MMADHC, ABCC4, TTF2, LAPTM4B, PHB2, SLC05A1, NOB1, XRCC2, POP1, ST6GALNAC3, RPL36A, RPL36, EFS, RPL18, ARHGAP11A, HMGAI1, BAZ1A, ACP6, NONO, TIGD7, TLL12, KIAA0020, RPS19, CSEIL, CTD-2192J16.20, KCNT2, PTDSS1, SMO, UBA2, EBPL, EDA2R, GPR19, FDXACB1, EIF4EBP1, GDAP1L1, ABCE1, LIN9, ZNF90, THG1L, CBS, FBXO17, BTF3, NAP1L1, IGFBP2, DDX39A, SNRPD2, PPM1G, NFXL1, KIF2C, RFWD3, CDC20, GGCT, WDR3, THEM6, SHMT2, C2orf47, ACN9, RPL37, MCM5, PYCR1, RACGAP1, KIFC1, EIF3H, TRAP1, RRP1B, KM-PA-2, TRMT112, ZC4H2, VWCE, MARS2, MCM3, GPR12, DCAF4, LSM2, RPL41, PLK4, HPDL, SPIN2B, CCT6A, MCM4, NME4, RFC4, RRP9, NT5DC3, IMPDH2, SIGMAR1, PRR11, PANK1, CCNG1, CCT3, COX5A, RBMX, MARCH11, FAPB6, MRPL9, PTRHD1, NOL10, RPL37A, ZNF367, PBX3, ZNF519, MRPL44, TONS1, LSF2, SSRP1, RPS18, IARS, LYPLA1, ZNF726, TKT, GPX7, RPL27, SLC1A5, CHML, PRC1, ELP4, DCTPP1, ZNF799, EIF3L, PSMG2, RPL4, BRIP1, ALDH1B1, DUT, RPL7A, METTL9, RUVBL1, RPL18A, ADAMTS17, CHCHD7, PRMT5, TP53, LDHA, PSMG1, CKAP2L, TNFRSF10B, FAM64A, ACP1, CCT5, NDC1, SSR4, MTAP, ESPL1, NOA1, RPS16, RANBP1, CCT7, BRCA2, SOX11, WDR74, POC1A, NFYA, C6orf48, ALKBH2, FAM136A, DNPH1, ZNF121, BRCA1, KPNA2, ZNF738, HOXD8, LBR, MRPL50, URB2, TOMM20, GNL3, HIGD2A, RHOBTB3, ALG3, CCDC88C, TMEM216, SMC2, PACRGL, BOP1, TXNL4B, IQGAP3, MRPS27, PCNA, CKAP2, AEN, SCFD2, RPS15A, RPL27A, DSN1, CMSS1, PDIA6, MGME1, MBNL3, LYRM1, SLC39A10, RPL15, ME2, CLNS1A, NCBP1, RFC3, PCOLCE, FH, EPRS, HELLS, CENPK, POLD1, CASC5, POLR1E, RNASEH2C, SLC4A1AP, MRPS2, TARS, PAQR3, POLA2, MCTS1, MSH6, CHAF1B, SNX17, CREB3L4, MRPL48, RPSA, TMEM243, FARSB, ZNF512, STRBP, POLR3D, HOXD3, NABP2, CCT4, CDCA4, NGDN, TOMM22, EIF2S3, GDI2, NUP107, GTSE1, CENPO, POLD2, EIF3D, NUP35, GART, CDK4, ADSL, SLC16A14, ACBD6, TEX10, JAGN1, FANCF, PNMA6C, MAP6D1, METAP1, GTF3C3, GTF3A, RABEPK, GSTCD, LARS, SNRPA, NCAPD2, RSL24D1, CCNE1, TRIM28, EME1, LY6E, KIAA1524, AIMP1, CCSER1, ZNF124, HSD17B12, SPDL1, GREB1, RPL8, NUDT5, IPO7, RPS8, NCL, STOML2, ECSIT, SDCCAG3, SAE1, FZD5, LEO1, SMC4, PDE3B, RPL3, TCF3, ZNF850, UNG, SF3B3, XRN2, DPY19L2, NHP2, GEMIN5, NUP160, SLC25A3, TMEM231, PRR7, NR2C2AP, RBMX2, TMEM194A, ZNF605, PTPLB, CIRH1A, UBE2S, NUDT10, COX16, ANP32A, TP53RK, RNF150, IGF2BP3, DDX21, PPM1D, TMEM108, KIAA1430, SSR2, LSM4, PRMT1, DAXX, KDELC1, NOL8, MRPL52, PHF5A, RPS15, TGIF2, WDR36, MTHFD1, ZNF551, TOPBP1, SFXN2, XPO5, PES1, ABCB7, DPP7, DTD1, LMNB2, GLTSCR2, C1orf43, SLC9A7, TACC3, GMPs, KIF18B, EHMT2, DPY19L3, HADH, USP3, MDH2, ZNF93, RTTN, EIF4B, IMMP1L, IGSF9, SNRPB, THOC3, SRSF7, FEZ1, CDC25C, MCM7, TAF4B, CHCHD6, PRKDC, CAD, GPR137B, ARHGAP11B, CSRP2, CENPF, PRR3, NACA, INCENP, DTNBP1, FAM27C, ZIK1, SLC3A2, TMEM33, C14orf93, WDR53, PHKA1, KATNB1, NELFE, CHN1, IL17RD, WDR62, SLC7A5, QTRTD1, RRN3, NUP93, DNAJA2, TMEM206, PAPPSS1, HNRNPA1, FADS1, SLC25A15, PA2G4, ZNF620, ST6GAL1, EEF1D, NUP155, RPL6, BFAR, PHKA2, GNL3L, GABRB3, LIAS, LRBA, RUVBL2, ZNF229, HDDC2, CDH24, PSRC1, GSR, SPAG16, TXNL4A, WDR12, C18orf21, WEE1, PLEKHG4, HILPDA, TMSB15B, UTP14A, FKBP10, UTP20, RPL7, POLR2H, RRP1, CCDC77, XYLB, DIEXF, GJC1, C9orf3, SEH1L, CSNK2A2, SMYD3, PPP1R14B, PSME4, FERMT2, DYM, C2orf43, INTS12, ZFP82, ACPL2, WDR46, NFYB, TIAM1, NOL11, FKBP4, DAP3, ANKS6, TBL3, SRPK1, TBC1D30, CTSP2, SCHIP1, GARNL3, LIG1, DOC2A, NCBP2, NCAPD3, CLPTMIL, LRRC8B, AGPAT5, ZNF77, CBX2, EI24, NNT, MPV17, IAH1, TROAP, SEMA5A, NOP58, CEP135, CNOT10, RECS, SMARCC1, ZDHHC9, POLR1B, C20orf96, WDR5, SUPT16H, VASH2, MED12L, SRSF1, MCM8, CCP110, KIAA0368, UCK2, EXOSC4, PARP1, ZNHIT6, HAUS6, KIAA1958, DDX31, C12orf73, HSPA4, CCNF, MRPL21, EIF3CL, NDUFS3, KXD1, EIF3C, OCIAD2, ABAT, IPO11, MARC1, ASNS, MED12, DGKH, DDX19B, NSMF, TRA2B, ZMYND19, GPHN, POLR1D, LRRC8D, QSOX2, AGLB5, SENP1, DCAF13, EIF3B, KIAA0922, KDELR1, FLAD1, EIF2B4, ZNF195, USP39, DNMT1, ARMC6, C10orf2, MCCC1, RPS27, OLA1, MDN1, STT3A, PFKM, KNTC1, ZNF548, PPRC1, PDCD11, RMDN3, ZNF852, RPUSD4, NUP210, SLC37A4, SAMM50, CECR5, CAMKK2, EEF2K, CYFIP1, URB1, ABCC1, PHACTR3, HEATR1, PRPF3, DES1, FBXL14, FANCC, KDM3A, ZNF496, EIF3F, AHCTF1, ANAPC1, SORD, DHX33, NOP2, LRP5, POLR1A, POLE, ARHGEF2, FTSJ3, U2AF1, PTDSS2, NOL6

NB_s4M⁺

Down

FOXJ2, ATPAF2, SMG6, MCF2L, KLHL29, EVC, ANKFY1, DHDDS, TSPAN9, ABCB8, WIPF2, ITSN2, VPS13D, HECTD3, CCDC149, PLEKHM1, ETV6, MKNK1, ZFAND3, MROH1, DMPK, RIN2, ZDHHC18, ZNF800, CLTCL1, ZBTB22, EPB41L1, BCAP29, KIAA0319L, GPN2, WIPF1, PLBD2, TTC7A, SOBP, WASF2, TPP1, DNALI1, PNPLA2, ABCA5, KIF1C, SLC15A3, UTRN, MINK1, MOB3C, PPAR2, STAT2, PLXDC2, LRRC32, FAIM2, NRBP2, PRDM2, PHLDB1, NCDN, ZBTB48, CDYL2, ZBTB20, CLCN6, NDRG4, FCHSD1, HEYL, KIRREL, PLEKHO2, KIAA1462, DMTN, BCAR3, MYO1D, CPBE3, MMRN2, CSF1R, KIAA0247, STAT6, MAPKAPK2, LHFPL2, SYNPO, IRF1, ALDH3A2, DUOX1, OPTN, ADAP2, MAPK11, TMBIM1, FMNL1, TMEM234, CAMK2B, PRKAG2, ZBTB8A, ZNF423, LAMA4, CCND2, ACSL6, LAIR1, TPRG1L, H6PD, NDEL1, PSD2, AC135178.1, PNMAL2, RRPB1, MYO1C, EPB41L4B, RAP1GAP2, CLSTN3, MVP, FAM124A, SRCIN1, SHROOM4, NLRP1, ATXN1, LGALS9, ZBTB46,

		<p>GDPD5, SLC26A11, TENC1, PLEC, SLC9A1, LCP1, KIF5A, TSHZ2, EHD2, PER1, APLNR, FLCN, ABLIM2, TRIM62, FGD5, ZNF438, KIAA1683, CACNA1D, TVP23C, ABCA1, WWTR1, SYNM, KCNQ3, S1PR3, NOTCH2, KLHL36, ECM1, ZEB2, ANO1, ATP2A3, SPRN, SLC7A8, LAPTM5, KCNAB2, LRRRC48, TNS1, KLF6, ABCG4, CTSZ, TMEM173, ABLIM3, CCNJL, CPEB4, SPTB, SEPT4, SHISA4, PVRL2, LCP2, FNDC5, GGT5, CMKLR1, CES4A, VSTM4, ZBTB4, DAB2, RAB15, FAXDC2, DOCK3, SORBS3, PCSK5, PGPEP1, CTSH, FAM214B, ALOX12, HID1, TGFBI1, HIVEP3, ATP13A2, RPH3AL, CAPS2, BCL6, SYT7, MYO7A, CAPN5, ARHGAP23, SLC6A2, ST8SIA2, C1QL1, TGM2, SLCO2A1, GPR153, TOM1L2, ECE1, CADM3, TNFSF12, TMEM150C, AHDC1, MATN2, C4B, FXYS5, KLHDC8A, CORO2B, ADC, GABARAPL1, CYSTM1, PDGFA, ASIC3, BHLHE40, CNR1, SREBF1, PSD4, FBXO44, ARSA, LTBP4, FREM1, SDK1, SORBS2, CD4, CD9, FOLR2, OLFML2A, IFITM2, HK3, NXPB1, TRIM56, REEP6, MAMLD1, DUSP1, FAM111A, MAFB, MAP7D1, ST3GAL1, C1QC, IFITM3, FGD6, GRASP, GPX3, CD97, CCDC24, ROBO2, PID1, ARHGEF10L, IGFBP4, EFHD2, KIAA1522, ITPKB, TNFRSF1A, PQLC2, ADCY1, ELF4, LMNA, LFNG, VGLL3, CGNL1, ITGA3, CAMTA1, MPEGI1, SOX15, B4GALT1, MERTK, APBA1, BTBD19, SAMD9L, IGFBP5, SDC3, LMCD1, CLU, MS4A7, GPNMB, C10orf10, ARHGEF15, MYCL, SLCO2B1, ZFPF2, NFATC2, SLC46A3, TCF7L2, CNTNAP3, MAP1A, PRR5, EPHX1, SYT2, ZBTB8B, CBX7, RREB1, ALPK3, PDGFC, DKK3, SCPEP1, SGK1, SPSB1, KRBA2, ZBTB16, WNT4, LRRRC71, PRUNE2, C1QTNF1, PLCH2, APOC1, IRF8, UNC13D, FABP3, GPR123, CHRNE, EPB41L3, APOL6, ELL2, KIAA1755, PBXIP1, PTH1R, SYT5, MYL9, S100A6, PLAT, AXL, HES1, TNFRSF1B, GAST, MYC, CD74, SLCO3A1, AHR, PRKCB, FBLN7, KNDCl, CXCL16, CDC42EP1, RASSF4, TPPP3, NFKB2, IL32, PLAGL1, SELPLG, ANXA2, CDH22, FAM131C, ITGAX, COL9A2, PTPRH, CYB561, IL6R, TH, TFEB, PEAR1, UNC5D, RBM47, C2CDC4, HPCAL4, MICALL2, CACNA2D3, HCLS1, S100A11, C1S, PODN, NR4A1, ICOSLG, ITGA8, FOXO1, ECEL1, FSTL3, KCNC4, BCL3, SGK2, SPOCK2, CSRNP1, ALS2CL, ZFP36L1, TYROBP, ZFR2, KIAA0319, KIAA0513, GLIS3, NAV3, EMP1, CYTIP, ZHX2, PIK3CD, SLC8B1, PVRL1, ELFN1, NEBL, LYNX1, ATP2B4, CITED4, ZCCHC24, CD300LF, ZC3H12A, UPP1, C10orf134, LRRTM4, ZBTB42, IGF1, NNMT, FBLN5, C2CD4B, TMEFF2, SPN, CASP9, C1R, C5AR1, LATS2, PRELP, SLC22A2, PLEKHG5, APOE, DNER, TIMD4, RASD1, SLC18A2, CAMK2A, PDE6G, GADD45G, PLA2G4D, CCBE1, CRYAB, ITGA7, KNCN3, TRPV2, NT5C1A, CCDC36, NTN1, PIRT, GFPT2, RAMP3, RGS9, ANPEP, RCAN3, MDGA1, CRIP3, SDC4, GPS2, MMP19, SERPINA1, ABCA9, EPB41L4A, PTPRZ1, LGI3, LSMEM1, LGI4, GADD45B, ACHG, ATP1A2, AMIGO2, CCL2, RGM4, SNCG, DBH, CALY, DIRAS3, SOD3, ADRB2, TNR, SOCS3, FOSB, PLP1, ZFP36, PMP22, RP11-386G21.1, SNTB1, SLC47A1, NR4A3, TMEM200C, TRPM3, RHBDL3, SOX6, ARSF, HAP1, TNFAIP3, NGFR, RAB3B, MEDAG, NLRP3, KCNJ5, TMEM132C, RGN, CRISPLD2, EPHA5, C10orf107, C3, TMEM255A, IL1R1, TNFRSF25, DFNB31, ICAM1, PLXNA4, BEAN1, PTGDS, XKR4, HMOX1, NR4A2, S100B, C7, LIF, SOX10, CACNG3, XCR1, POU4F2, KIRREL3, SV2B, PLA2G2D, HSPB6, CHD5, AC074389.6, GPRC5A, NTRK1, MT2A, HTR1B, ANKFN1, HCRTR2, CGGR, DRD2, HOXA5, INSR, MEI4, NOV, APOD, PLA2G2A, GSTA1, IL6, STAR, CYP11A1, CYP21A2, CYP11B1, HSD3B2, UTS2, CYP17A1</p>
NB_s4M ⁻	UP	<p>HIST1H2BA, MC5R, ALX1, FAM197Y1, CD200R1L, HELT, TMPRSS15, AC006156.1, TSPY10, SPINK2, TSPY8, MAGEA9, TSPY3, MAEL, TSPY4, MAGEA10, MAGEA9B, PSORS1C2, TSPY2, OR6Q1, HBB, CHMP4C, DHRS2, TSPY6P, OR2G2, FDCSF, NAA11, OC90, NOG, TPRX1, CALB2, FIBCD1, C6ORF163, CGB2, MUC15, DEFB104A, OR52B6, EPPK1, OR6K3, PAGE2, OR14C36, FAM19A3, OTOP3, NPAS4, OR4D6, AL050303.1, LTA, AL390778.1, DEFB104B, ALX3, CGB8, REXO1L10P, REXO1L11P, SSX1, IAPP, GTSF1, HIST1H2AJ, OR6C76, AC017028.1, KRTAP4-9, CSN3, LRRCL10, RNASE8, C10orf65, ADAD2, AC092850.1, ASB5, PRM2, TUBA3C, RBPP8NL, POTE, RGS13, PXT1, KIR3DL1, GALR2, HIST1H2AI, SLC22A12, IFITM5, CST6, NANOS2, OR5P2, PSMA8, C7orf62, ZNF695, FGG, PYY, BTNL8, FGF3, MAGEA11, IRX4, KCTD8, HIST1H2AL, HIST2H2AB, ACTL9, SLC6A18, SKA3, SLC5A12, LGALS7, HIST1H3H, TMEM78, LY6D, AC020907.1, KCNFI1, LGALS7B, ZIC3, LRRRC3C, CGB1, SPANXN1, RGR, LEUTX, PYDC1, USH1G, DBX1, TRPM5, MUC5AC, POTE, OR5M1, BIRC5, FOXA1, OR2T2, AC008060.7, OR2J1, BOLL, MKX, CRYM, TCHH, FGF5, HIST1H3B, RP13-279N23.2, IRS4, MRGPRX4, C8G, WNT9B, GGT6, SP6, DSCA3, OR10Z1, SLC6A3, ALAS2, GPR142, SCGB2A2, MAGEA4, FEAR2, WNT3A, SLC22A11, C10orf99, CST9L, C10orf64, OR6K6, ODC1, AC008271.1, LCN10, TERT, SPDEF, KRTAP4-16P, C9orf141, KIAA1024L, MYH6, SLC30A8, BX649567.1, C11orf86, KLC3, TMEM235, DEFB103B, DEFB103A, HIST1H3C, ETNPPL, ISL2, FOXJ1, MAGEC1, OR13H1, OR6K2, OR4M2, LGALS16, HIST1H2AM, ANKRD2, RP11-520P18.5, SLC32A1, CGB5, HIST1H2AG, NCAN, CLEC4M, LINC01098, ALDH3B2, CDKN3, RP11-481A20.11, CT47A2, APOH, ANKRD35, FAM83A, VGLL2, RBFOX3, RGS18, FLJ45079, OR4C46, BTLA, MIOX, KRTAP4-3, KRTAP29-1, KRTAP10-4, CATX-2, FAM181A, LRRRC56, CGB7, MUC8, C8orf87, CCKAR, AC104057.1, SPATA31E1, C4orf22, IL19, PIGR, RP11-67H2.1, CST11, NKX3-2, AC104809.3, GNGT2, DNAI2, CCDC169, SLC25A47, KRTAP10-5, GLDN, MOG, CTD-2203A3.1, REN, DUXA, CIB3, OR10H4, RP11-382J12.1, OR10K2, LHX5, CXCL9, RP11-47I22.3, C11orf85, LLGL2, IDO2, PPBP, OR14K1, AVPR1B, GBP6, MMD2, SEMA3D, LIPF, MSLN, TRIM54, KRT32, EFCAB3, POM121L2, PTTG1, COL25A1, MYH14, OR6P1, TMEM210, TRIM10, OR52E6, CTRC, HIST1H4D, VSX2, HIST1H3J, OR10K1, CAPN13, FGFR3, AANAT, KRT82, UGT1A4, RP11-379H8.1, SLC22A8, NCCRP1, NPHS1, HBD, GPR152, KIF20A, HIST1H2AH, C19orf69, CD3D, KRT4, OR56A3, CDH7, RP11-477N12.3, TRIM60, KRTAP10-2, HIST1H2BB, NBPF4, OR1L8, ANO9, UMODL1, KLK15, MORN5, THEMIS, TFAP2C, ALOX12B, GSX2, HIST2H2BF, KLK7, CCDC42, RP11-113D6.10, AHSF, TRIML2, CFHR4, S100A12, SIX2, ZNF560, C4BPB, LHX9, TJP3, TMPRSS12, RP11-41O4.1, APOA2, MISP, OR56A4, C2orf53, CELA3A, HIST1H4C, MNX1, CCDC178, KIF19, KRTAP5-5, RRM2, TDRD15, HIST2H3D, CEACAM16, PEBP4, ELF3, OR2T35, HIST1H2BC, FGF8, OR13C8, HIST2H2AA4, TMEM8C, FAM72D, PBOV1, ENTPD8, SYT8, MOGAT2, MTPP, HIST1H3D, FOXI2, CXCL13, CPA2, GSG1L, HOXB1, GDF2, WNT7B, SLC25A21-AS1, SP8, CCDC182, HAVCR1, RSPH1, LEFTY2, MYOD1, SPAG11A, PAX9, CRB1, POM121L7, NBPF6, TTR, GIPR, LGSN, RP11-10A14.4, SPPL2C, CT47A1, GAD1, TEX37, DLX4, MESP2, C3orf30, PCSK9, OLIG1, GPM6A, OTP, NPPB, HRASL55, WBSR28, GPR78, CCDC144NL, CLEC4D, NTSR1, KIR3DX1, FMO6P, OR2C3, ASCL5, FOXD4L3, AMBP, FAM153C, OR5AC2, USH1C, DSCR8, SPAG11B, TBC1D28, ZSCAN10, ENPP7, KCNJJ1, MUC6, DMBX1, HBA1, F2, FGF22, SPACA1, HIST2H2AA3, CT47A8, HDGFL1, HIST1H2BJ, HBA2, GPR144, CCDC63, DDN, CLDN19, SLC6A19, TNNC2, LINC00955, KIF18A, CATSPERD, ELFN2, ADRB3, BPIFB4, PRAME, PRAMEF13, UTS2R, GNAT1, C5orf47, AL513478.1, DUX4L3, SMR3A, DMP1, AIPL1, FOXD4L6, CCDC83, KIF14, GHSR, STAB2, A1CF, TAC1, ONECUT1, TRIM6, KIRREL2, ALLC, RASSF6, ASPM, HIST1H1B, HIST1H2AB, MAGEA8, SGCA, SYCN, FAM84A, OR52L1, CT47A6, KIF12, FAM25C, LRIT3, MRGPRX2, NTS, MMRN1, SMYD1, TKTL2, MAGEB1, THEGL, GLDC, ATP1B4, COL9A3, DYTN, IDI2, WDR38, REXO1L1P, GFRA4, VAX2, FAM53A, EXD1, B3GNT3, AQP8, CNBD2, MAL2, MCM10, SPATS1, KLF1, INS, C10orf106, PRPS1L1, ASF1B, POMC, SPNS3, PHOSPHO1, DNAJC22, KRT3, TBX4, HIST1H2BK, HIST1H2BM, AC011997.1, TCF23, TRIM7, DSCC1, HIST3H2BB, CREB3L3, SMTNL2, CT47A10, AMDHD1, RPP21,</p>

AXDND1, FAM153B, GLP2R, AGXT, EFCAB1, STXBP6, CD79A, C9orf62, RP11-863K10.7, CLCNKA, PIFO, AQP10, IRGM, ANXA13, ILDR2, TBX1, S100A7A, IZUMO2, LRR9, SEPT12, TFF3, RBM20, MYBPH, ORL4, GHRHR, CD3G, TPRX2P, CT47A3, MYRN2, CHST6, FLJ20306, HIST1H3F, FOXI1, CYP8B1, CT47A4, PSCA, KCN7K, LAX1, WDR76, CLDN9, KIF4A, HES2, HIST1H2BE, OTOG, RP11-758M4.1, XIRP2, C2CD4D, TTK, TMEM184A, CFC1, CRB2, FAM153A, HIST1H2BL, RP11-362K2.2, KLK1, MAFA, GABRR1, RDM1, LGR6, BFSP2, HIST2H4B, DBF4, ATP6V1C2, P2RX1, EPSSL2, GGTL2C, HIST2H4A, HPGD, DCST1, DIAPH3, SLC1A6, CD1B, MYOG, NLRP8, SOST, C9orf171, DLX3, DTL, RAD21L1, STAP1, AQP12B, POTEE, TPD52L1, HPN, DNAH3, C1orf222, USP17L10, FAM9A, XAGE1E, SGK3, GPA33, XCL1, FXYD1, CCDC42B, SPERT, TTC24, TRIM59, MYO1H, FAM111B, RP11-383H13.1, CDC6, TTC6, SIGLEC5, MAP1LC3C, SSX3, NKAIN4, CLECL1, OR9A2, ATP1A4, FOXP4, GABRA1, GRAPL, E2F8, VSTM1, SLC4A1, SEC61G, GPR114, TTL6, VTCN1, PGK2, USP17L21, CD1D, TBX21, CLCNKB, SEC14L6, TSPG1, HMMR, CCDC87, ANKRD33, CLEC17A, MT1F, USP29, UTS2B, AC007040.11, CPSF4L, TULP1, PRSS54, UBE2C, YR1, GAPDHS, USP17L20, USP44, FAM83E, TSPAN19, hsa-mir-150, C1orf53, IFNLR1, FSHR, ASCL1, CFC1B, TTC40, RPL22L1, USP17L12, PRKCG, GRB7, BTBD17, HIST1H2BH, TEX15, ABCA4, A4GNT, RSPH10B2, AK8, TOP2A, NWD1, TCP10L, PNLD1, SH3GL3, USP17L19, KRBOX1, FAM72A, SLC35G5, LRTM2, XDH, GPR31, C1orf180, TSPEAR, USP43, C1orf116, GAST, TK1, KIAA1257, CXorf57, NAALADL1, CLEC4F, MUC5B, FAM209A, HOXA6, ABCG5, TPTE, P2RX2, EGFL6, C15orf32, CCNN1B, C1orf100, COL24A1, DNAH12, NXT2, C9orf116, HIST1H4L, ATP5J2-PTCD1, GCSAML, KLHL1, NCAPH, SULT1C3, EFCAB10, TDRD5, PRSS57, RGS6, GMNN, NDC80, RP11-12J10.3, C14orf64, STMND1, OR1K1, USP17L22, BMP2, CBLN1, RIBC2, GNG13, CCDC151, SNX31, USP17L15, BUB1B, USP17L18, AC012123.1, USP17L11, HIST2H2AC, CD5L, TNNC1, CCNB2, VWA3A, TMEM236, OR1Q1, USP17L23, C10orf107, SYCE2, CCL14, ST8SIA5, RSPH10B, CAPN11, S100A3, USP17L17, LIN28A, HIST1H2BF, KIAA0101, RMI2, MCM2, UBE2T, RHPN1, MMP13, HIST2H3A, KIF6, CABP1, HIST2H3C, AC006372.1, MPP7, AL355531.2, MATK, HIST1H2BI, EME1, OR6V1, AP003733.1, KIF11, ERP27, HEMGN, LEP, HIST1H2AK, AGR2, C10orf53, ZNF563, TRPV4, NOS1, CCDC150, NOTUM, CTD-2021H9.3, CR2, C7orf63, USP17L5, NEK2, PKHD1L1, LRRRC31, NEUROD4, HIST3H2A, KRT12, ATP8B3, TTL13, C17orf47, DEFB124, AC008267.1, AGMO, LPO, SH2D3A, CCNA2, C14orf39, NAT1, C17orf50, ANKRD65, KIF4B, EPN3, FAM83F, CENPU, MYO7B, PADI4, GLP1R, CENPH, WDR72, TNFSF4, INHBE, SLC2A2, TIFA, USP17L25, KCNS2, PPP1R36, USP17L29, STRA8, COL22A1, LY6H, HIST1H2BG, CHRD, ATP6V1B1, ANKRD22, ETV2, MS4A1, AC092687.4, HNF1A, TRMT10C, USP17L24, SCNA4, TFPC2L1, CCR6, ICOS, RP11-385D13.1, UXT, MAP3K19, RIN1, USP17L26, NPSR1, OR52N5, G6PC2, GYPB, CBLN4, KCNT1, LRRRC43, SUCNR1, MELK, GYLTL1B, CNR2, SPATA12, POU6F2, MAD2L1, SCIN, TBX6, FGF19, SLC26A7, TULP2, CLEC5A, TREML2, CTD-221E18.1, FOXL1, SPIB, SKA1, USP17L27, MYLK2, CNPY1, MET, SGOL1, MYH7, HIST1H3G, CENPF, SP7, USP17L28, HOXA3, PPM1J, FAM160A1, POLE2, NUSAP1, C9orf170, ZNF812, SLC9A3, SLC6A11, USP17L30, MS4A4E, MYOCD, SLC4A11, GINS2, ZYG11A, SLC2A12, EMR1, KIF2C, CALHM3, CTD-3203P2.2, SLC35G1, TYMS, RAB19, HIST1H3A, AIM1L, BRCA1, CNGB3, IGFLR1, BRIP1, HERC5, LSM3, MKRN3, GNGT1, RASL10A, C9orf50, NPFRR1, RNASE10, SP140, CCDC129, GLYATL2, ABCG8, KAZALD1, C1orf61, ZWILCH, CELSR1, ICAM2, EVPL, HSH2D, ZNF98, SERPINA11, TMEM182, PDZD9, FCRLA, TRPC7, HIST1H2BD, GARI, CDCA2, ATAD3C, TST, CKS1B, ATAD2, MYOZ1, RAD51C, HES5, SLC16A11, EPX, TRPV6, WFDC2, MARVELD3, OVCH1, RRR, KCNG2, FEN1, FAM72C, ORC1, FOLH1, STK33, PYGM, HIST1H4I, CEACAM21, GRAP, RHBG, NR2F1, RBL1, CCDC11, ABC2, RIMBP3C, DCST2, TUBA8, WNT8B, PAX6, GMPD4, AL953854.2, RIBC1, TSPAN33, FCRL3, ANKRD18B, RAD51AP1, SDSL, PARP15, STPG1, EXO1, GJD4, TMEM31, EFCAB6, LY9, POLA1, GCM2, GABBR2, KIF15, CTA-299D3.8, PPP4R4, MYO5B, WDR96, ACCSL, RIMBP3B, ZNF165, SFTPD, NCAPG, HIST1H1C, RAG1, POC1A, PKD2L2, PIWIL1, MYLPF, PPIAL4D, SH2D5, KLHL10, AVPR2, CACNA1I, CENPN, RASGEF1A, DNA2, SHCBP1, RIMBP3, MIS18A, RGS2, GPR133, ACOT4, SGOL2, CCNB1, CHAC2, KCNJ6, TLCD1, KDM4E, IRX3, HRSF12, LYPLA1, DMC1, SLC12A8, ACOT12, PLA1A, TMAI6, HIST1H2AE, CDC20, LCT, CTLA4, GRM1, RP11-451M19.3, PRODH, TMA7, RTDR1, LCK, GTF2A2, GPR26, ALS2CR11, CD7, RHOH, SLC38A5, AL031663.2, IQGAP3, HIST1H2AC, MYO16, GCK, CDC25C, C1orf15, FANCL, EPHA1, GSTO2, HIST1H2BO, ARHGFE39, KIF24, FOXD4L2, ANGPT1, ROPN1, ESCO2, IL24, FOXD4L4, RNF39, XCL2, CDC45, UPK1B, C2orf61, NR6A1, CRHBP, GPAT2, ADCY8, NHLRC4, MND1, HTR6, GBP5, MRPL11, LAMB3, MDF1, NUP62CL, KCNHC, ATP13A4, STRIP2, TEX14, FANCD2, TRIP13, ANKRD30BL, ZNF670, PDE1B, TXK, SPATA21, NME1, CCDC108, EVC2, SH3TC2, IGLL1, PRR11, TRAABD2A, ACN9, PKP2, CDK1, FAIM3, TEK2, ZNF367, GPRC5C, PBK, CDK2, IBA57-AS1, AURKA, DSCAM, HSPB9, SAMD3, ZNF385C, HIST1H3E, NOL10, DEPDC1B, TRMT10A, JPH1, FAM180B, PPP1R3G, IQCA1, DHFR, DDX1, LIN28B, CDCA5, HSD11B2, CCDC141, C5, UBASH3A, TPRG1, GPR156, TLR10, ADCY10, CCR2, RPS17L, CCR7, LIN9, CDKL4, C12orf50, RRM1, C21orf90, AUNIP, GPR171, AC017081.1, ZNF595, SEMA3E, MTFR2, BAIAP3, MRPS36, AP000974.1, ORC6, CABP7, FBXO15, SAAL1, PTPRQ, CD28, HIST2H2BE, KBTBD12, CCR4, UNC5CL, DLEC1, IL17B, NEIL3, SLC22A15, MYOZ2, SKAP1, SPATA24, LRRRC36, F10, HSF2BP, GRIK1, NUF2, C11orf16, FAM72B, SYNPR, CDCA7L, GSG2, SPTBN5, ABC8, ZNF799, TPX2, HPCA, EPHX3, HRC, LAMC2, FKSG61, PPM1N, HMGB2, TXNDC9, CCHCR1, CKS2, C4orf46, PTRHD1, COX4I2, NUP37, SLC44A3, CARD11, CERKL, CLSPN, TAF9B, ECT2, AC008132.13, UBXN10, TNFRSF11A, BAI1, TMEM243, HJURP, LTB4R, COL11A2, FKSG62, BRINP3, POP1, NUDT7, CHCHD7, NCF1, PPM1D, TMC7, RPS13, ARHGFE19, SPN, CENPA, IPO4, AURKB, LRRN3, NDUFB6, WDR62, POTEF, PTPN22, C7orf25, KREMEN2, PRKCG, ANLN, SLC6A12, TRH, CEP70, NME1-NME2, GABRA3, TMEM74, MRPL3, PCDH8, ZC2HC1C, C17orf104, SLAMF6, SLFN12L, PFKFB1, DNAH17, KIFC1, RASEF, CHAF1B, LRRN4CL, CD5, RPL17, AL645608.1, DLGAP5, SLC22A18, NINJ2, RPL34, LRMP, E2F2, FBXO5, GPRIN2, MCM6, CCT8, SLIRP, ELP4, CXCR6, PPA1, VAV3, PLK1, C19orf67, CD274, TIMM17B, LRRRC46, ZNF639, SLC29A2, RPL39, RP11-146D12.2, PPP1R14A, TBC1D10C, C9orf43, CDKN2A, SF3B14, ZNF575, KCTD19, SELL, SOWAHA, RNASEH2A, PKHD1, MRPL1, C3orf67, STPG2, C9orf96, FAM27E1, CPSF3, RAC2, TMEM97, SLC7A3, DNAH14, CNH3, LYRM1, TRAPP2P1, CRYZ, BUB1, AP001055.1, C4orf21, ZNF684, CCNE1, EMILIN3, ANKRD37, TG, SEMA3A, FKBPL, CMBL, ZWINT, DSN1, RPL38, GRM2, GNB1L, ZNF724P, CCDC62, OIP5, ZNF280C, ZNF284, NPM3, PRIM1, TUBD1, FKSG68, B3GALT1, IGSF22, RCOR2, TAX1BP3, CKAP2L, GVQW1, HSD17B7, EFCAB11, CEP135, RFC3, POLA2, SDK2, SRD5A3, RACGAP1, ANKRD32, GPR113, AC019171.1, CENPI, DPP4, KY, BZW2, ADM5, RAB41, H2AFZ, WDR34, KRTPAP2, CDT1, MAB21L2, INTS2, ACP6, ALPK1, PHF6, TOP1MT, SYTL1, ZNF230, MYBL2, CDC25A, SLC39A4, E2F3, ESPL1, GRIN2C, GINS1, ZNF280B, LRRRC39, ISCA2, CDH23, SPRNT, KNSTRN, VRK1, SP140L, MRPS9, AKAP5, ERCC6L, SNX32, BORA, ZNF420, MIR3654, CRHR1, TMEM56, SMKRI1, AC027307.3, RPA3, NXPH3, DNAH10, C16orf54, XRCC6BP1, C11orf48, BRCA2, GPR12, KBTBD8, HIST1H4B, XRCC4, MKI67, HAPLN3, NME9, CORIN, FAM151B, RHOF2B, PANK1, CD38, PCTP, PTGER4, IMMP1L, KPNA2, ABHD16A, AC069547.1, IKBIP, BTBD11, MIPEP, CYP3A5, GSTCD, ICT1, PTPRCAP, PDCD2L, SCML2, CDC7, RPS25, HIST1H3I, MAATS1, ITGB3BP, MSH2, KCNP4, FBXL13, CNTD2, SAPCD2, DNASE1L2,

		<p>PABPC4L, POLN, UQCRQ, PRR5L, C10orf25, SHQ1, RPS4X, TCF19, DEPDC1, NGDN, TMEM139, ZFAND1, DENND2D, MYLIP, EAPP, CLEC18A, SMC2, APIP, KIF23, THOC3, HAUS1, PSMG2, EXD3, TMEM38B, BAIAP2, RFC4, PTPMT1, GRM4, TSEN54, ACKR4, NDUFS8, SLC25A19, NPFB9, VDAC3, F12, CLDN4, MFNG, WDR53, RANBP17, PCNA, MRPL50, ILF2, PLK4, CCL5, RFWD3, NUP107, RPL10A, HMGB3, FSTL4, HIST1H1E, RPS6, UCHL3, FOXM1, C19orf48, ACP1, E2F1, AC073333.1, GFRA2, CHAF1A, STIL, RPS14, ZNF850, KCNH3, DNAH6, TTC33, ARL13B, KLHL15, UBA2, MRPL40, CEP152, POLE3, LNPI, NCAPG2, IGF2BP1, FGGY, ZNF519, IGBP1, CHEK1, EIF3M, FOXN2, C12orf61, POLD1, SUV39H1, SPAG16, SUV39H2, KLHL32, CCNF, CCDC18, LIPT2, FAM64A, MRPL9, CAMKMT, CCDC15, TMEM232, FAM154B, PRC1, C1orf74, LRRN2, TMEM99, SUPT4H1, CLDN5, FAM19A2, C3orf33, ARHGAP11B, SPATA18, HINT1, MCTS1, ZNF157, GRM6, GGCT, PPLI3, TAF5, MCM5, SLC39A11, AC004381.6, C16orf91, TSHR, DTWD2, WNT3, RPS21, CDKN2C, XRCC2, ITPRIPL1, CENPO, ZBTB12, SLC25A14, SLMO1, ZNF19, GCA, SLC16A1, FBXO48, C19orf73, FES, RBBP8, COQ6, ZNF569, VBP1, MCM4, CSRP2, BLM, CCSER1, PIGP, CHCHD3, RPL29, CCDC157, MAP6D1, UTP18, TBC1D31, METTL21A, FAM115C, ORC5, TMC6, ITGAL, ASRGL1, PDCL3, CHGA, MRPL52, ANAPC10, COPS4, ACTN2, TP53TG3D, MALSU1, TONSL, HMG2, DUSP15, RPS26, ZNF738, DKC1, THAP10, HIST1H2BN, CDCA4, E2F7, GMCL1, GGACTION, SARDH, IL17RB, CLP1, SPC24, YOD1, SLC9A7, RFC5, ATP5H, IKZF1, HLTFF, C7orf31, IFT57, NUDT5, MRPS7, PACRGL, COPB1, PRMT3, GALK1, MARCH1, GJB7, GINS3, NFYB, AC009060.1, RFC2, MRPL44, SMG8, ZNF90, MRPL21, HTR2B, MYNN, NOL11, HOXD9, BCL11A, E2F6, ZNF501, RWDD3, PDHX, ZNF530, ZDHH13, BTBD10, HAUS8, RPS24, PCDH17, SPECC1, SKP2, CCDC117, TOPBP1, YY2, DPY19L2, TRIM39, GEMIN6, AXIN2, CTD-2192J16.20, CDCA7, VWCE, PARVG, MRPL48, LSM2, RARS, STRBP, TICRR, COX16, DBR1, MSH6, RPS6KA1, LMNB1, MCM3, ASIC4, NIPAL3, PPP1CA, PDE3B, APEX1, CENPK, AL645730.2, FADD, TACC3, FANCA, TRMT112, KIAA1524, LAGE3, PSTK, PSCR1, BRMS1, ANAPC11, RNASEH2C, GAL3ST3, MBIP, KB-1507C5.2, CCT3, DHX40, ZNF627, PTPN6, ICAM3, POLE4, BOLA3, PTPLB, TMEM216, OPR11, WDR3, MTHFD2, DTNBP1, WEE1, CASC5, PDK3, ZNF620, RASGRP2, CSE1L, GEN1, MTHFS, JAK3, IPO11, RTTN, SLC38A9, BCAS4, KIAA0020, MRPS17, ZNF675, LRRC45, TMBIM4, WDFY4, C21orf59, C17orf80, GTF3A, ARHGAP11A, ATP2A3, MRPS11, PPP2R3B, ZNF169, AP4E1, STAG2, NUDT1, RPS15A, SSRP1, ACAA2, TMEM223, MALT1, NONO, FAM65B, TP53RK, LRBA, CAMKV, ITGB7, RANBP1, ZDHH12, E2F5, NSF, BNIP1, MRPL27, FTSJ3, HELLS, COQ2, ZNF773, BLOC1S6, PDP2, MPV17L, TAF12, LBR, CCD47, ZIK1, NT5DC3, RNASEH1, SFXN2, KIF18B, POLQ, NFXL1, TTF2, LIG1, SMURF2, JAKMIP3, NUP35, SPIN2B, SEH1L, TMEM101, C18orf21, SLC25A13, APTX, FAM105A, DDB2, EIF1AD, C14orf93, RAD51, KNTC1, NPL, NSL1, SKA2, NABP2, ZNF2, IPO5, GTF2H2, CEP192, BEST1, DYNC2L1, FIGLN1, DUT, FIBP, PET112, METTL2A, PDHA1, STYXL1, TOP3B, MRPS33, TMEM108, BRN3BP, DCLRE1C, SLC35B1, COMMD10, PSMC3, CCP110, DPF3, CLPTM1L, FADS1, MPST, ZC4H2, INCENP, FAM114A2, NDUFS3, SEC11A, RPS7, C17orf75, ZSCAN1, TMEM231, PPIH, CCZ1B, AKIP1, SAE1, RAB40B, TTF1, CCM2L, PSM13, FAM133B, DAXX, MRPS6, IGHMBP2, HSPA14, MRPS22, CBX1, UQCRB, LRRTM1, AHCY, ERCC4, VARS, C17orf53, WDR74, MRPL16, FANCG, NKAP, AZI1, COX18, FBXW7, RAVER2, SYT3, GTPBP10, TMEM194A, ALG1, RRP1, PIK3C2A, DUS2, METTL9, SENP1, PASK, RPL31, TMEM86A, ZNF77, ATP5J2, CDCA8, ZNF273, POLD2, MRPL46, PIH1D1, ZNF394, NUCB2, SERPINE3, RUVBL1, SMYD3, TDPI, MRRF, SCNM1, CHTF8, TMEM116, EIF4A3, ARNTL, RBM38, SPDL1, AGBL5, TXNL4B, TSTA3, FBXO3, NUP160, SPTY2D1, CIT, PIK3R4, QTRTD1, MTHFD1, C3orf70, TRMT11, RAD9A, TAF1B, FAM117A, MFAP3, FOCAD, MTRF1, NELFE, MCU, TRIM4, CEP85, CBX2, ENGASE, MACROD1, SETD4, PLEKHH1, TUT1, RAB11FIP1, LIN54, VPS25, N6AMT2, ACTR3B, C18orf25, INTS12, ZNF605, MKS1, CBR1, RNF34, GJC1, ASPSCR1, NUP98, CCZ1, NUDT4, GLE1, TIMELESS, SLC37A3, RWDD2B, EIF4E, LMBR1, NUP93, IFT43, ZNF619, ADAT1, RPS6KB2, MED12L, NBPFF10, PCCA, PRCD, GEMIN5, TCEB1, MYRF, EDRF1, TRIM45, NFYA, GPATCH1, TMEM209, ARFIP2, HOMEZ, IGF2BP3, TTC9C, IKZF2, TECP, RECQL4, ZNF212, TEX10, SLC7A6OS, DLG3, ATLL2</p>
NB_s4M ⁻	Down	<p>PLEKHG5, PLBD2, SREBF1, PLXDC2, MINK1, TNS1, TOM1L2, ARHGAP26, ZBTB4, GPR153, MAP7D1, LZTS1, GRAMD1B, FAM124A, POLR2A, BCL9L, TRIM67, AXL, B4GALT1, FBXO44, RAB15, DONSON, RAIL, TSPAN11, KLF6, DOC2B, IGSF9B, SYT2, ARHGAP23, PLAGL1, BHLHE40, GAB2, CDC42EP1, SPRN, IFITM3, MCAM, PHLDB1, PER1, ZEB2, ATP1B3, C17orf107, MAP7, NXPH1, KDM6B, SERPING1, IRS2, LDLR, LMNA, THY1, IGFBP5, ZFP36L1, DNER, DCC, PCDH9, SOX10, LRRC7, CTGF, NFIC, ELL2, KIAA1683, ST8SIA2, IGFBP4, PDE10A, MYADM, ABR, CRYAB, COL1A1, SCARB1, CHD5, SEMA3C, MYL9, IL1R1, EMP1, CRISPLD2, CREB5, DUOX2, PVRL1, NOV, EYA4, PLXNA4, EPHX1, FOSL2, C1orf233, PAPP, CHL1, NR4A2, C11orf96, NR4A3, SERPINE1, ZFP36, EPB41L3, PI15, CEPPD, PIRT, SV2C, NR4A1, HES4, KLRK1, RASD1, DRD2, CEPPB, HSPB6, SFRP2, HMOX1, ODAM, CYP21A2, AP000783.1, VGF, IL6, GSTA1, MT2A, APOD, CYP11A1, STAR, CYP17A1, CYP11B1, HSD3B2, MTRNR2L2, UTS2, VIP</p>
NB_s4S	UP	<p>KRT14, KRT1, KRT5, KRT10, DSC3, DSG1, FLG, FLG2, KRT6B, PERP, TYRP1, KRT6A, DSP, KRT25, SFN, COL17A1, EPPK1, HBG2, PKP1, SEC61G</p>
NB_s4S_4	UP	<p>KRT14, FLG, KRT10, KRT1, HSD3B2, KRT5, DSG1, CYP17A1, DSC3, DSP, FLG2, SFN, KRTDAP, TYRP1, KRT6B, PERP, KRT6A, APOD, KRT25, CXCL14, CLDN1, PLP1, HBG1, HBG2, DCN, TRIM29, OGN, DRD2, GPC3, PIRT, HIST1H1A, MATN2, FGL2, IGF2, COL3A1, NOV, PLXNA4, TCF7L2, ASPRV1, PVRL1, SLC18A1, ANGPTL7, PMP22, AHNAK, EPB41L3, GSN, COL1A2, CHD5, LRGLK, MPZ, NXPH1, TGFBI, GAS7, OLFML2A, DCC, DNER, FAXDC2, MXRA5, DOC2B, CD9, KLHDC8A, ADCY1, SYT2, PLAGL1, ATP2B4, CGNL1, CREB5, ST8SIA2, MAP7, LAMA4, KIAA1683, ZEB2, CAPN5, ROBO1, PHLDB1, KIAA1522, ALDH3A2, PLXDC2, ZBTB4, TOM1L2, CRT2, ARHGAP23, MAPK4, MINK1, GALNT16, TNS1, KSR1, DPH1</p>
NB_s4S_4	Down	<p>PRR11, FANCI, HJURP, CDCA5, CDCA2, KIF4A, CDT1, BRIP1, DSCC1, E2F3, RRM2, ALK, ENDOG</p>

NB_s4S_4M ⁺	UP	KRT14, CYP17A1, HSD3B2, FLG, KRT5, DSG1, CALML5, DSC1, KRT1, FLG2, DSC3, CALML3, KRT10, KRTDAP, MEI4, KRT17, DSP, SFN, CALHM3, KRT6B, KRT34, S100A7, KRT6A, TNP1, TACSTD2, HTR1B, KRT15, CLDN1, NCMAP, NOV, ANKFN1, DMKN, PERP, COL17A1, INSRR, KRT71, GCGR, KRT25, NTRK1, TYRP1, CD58, KRT16, APOD, ADRB2, PLP1, PI16, S100B, WDR16, TMEM255A, DRD2, SERPINA12, MEOX2, KIRREL3, XCR1, TRIM29, ARSF, SV2B, FGL2, ERBB3, SLC24A2, POF1B, CDH1, XKR4, PTPRZ1, CFD, SOX6, AC074389.6, COL28A1, MPO, CACNG3, NTN1, CHD5, TP63, RHBDL3, C7, ATP1A2, PLXNA4, PTGDS, ABCA9, DCN, CCBE1, AMIGO2, DIRAS3, PKP1, SNTG2, EPHA5, PRX, PLA2G2D, EPHX3, SLC22A3, RNF43, KRT19, RGMA, KLF5, SOX10, RGS9, ALOX15B, PMP22, ASPRV1, CAMK2A, RP11-386G21.1, SOWAHA, FCGBP, TNXB, KCNC3, FAM19A1, ITGB4, C17orf107, MATN2, C2CD4B, CLEC10A, IGF2, RAMP3, NLRP3, ALS2CL, CLDN4, TCF7L2, ADAMTSS, LRRTM4, TNF, FAT2, RCAN3, FMO2, EPB41L4A, RAB3B, CLEC3B, C1orf134, TGFBR3, NT5C1A, GSN, GAS7, PLA2G4D, PROKR1, C1QTNF2, CDCP1, ITGA8, CRIP3, EGFR, LRGUK, SERPINA1, PIRT, TRPM3, CD52, FAM131C, DBH, PRKCB, MYO15A, PPP1R1A, NEBL, CGNL1, TMEM132C, TNFRSF25, FAM129A, LCK, TIMD4, CACNA2D3, SGK2, SPN, MDGA1, ATP2B4, CD9, GPR17, CREG2, F11R, OSBPL3, C1S, DSC2, DNER, EDN1, TMEFF2, MEOX1, WNT4, EBF1, KIF17, FREM1, SLC18A1, KCNK6, CAPG, NXPH3, CCM2L, ITGA7, ABCC3, NGFR, PIK3CD, RASSF4, EMP2, CD300LF, SLC18A2, AHR, C10orf11, VSTM2A, SCPEP1, PLCH2, ZFPM2, PVRL1, MAP3K6, MPZ, CYP46A1, TRIM55, MYCL, ZCCHC24, S100A6, KLHDC8A, AK5, ROBO2, ADC, PEAR1, PLA2R1, GPNMB, ATP2A3, AHNAK, UNC5D, PLEKHG5, HPCAL4, GATM, C1R, IFFO2, ANGPL7, CRYAB, IL7, BMP8B, ARHGAP15, SCN7A, COL9A2, FAXDC2, PTH1R, CCDC36, KCN110, FOXO1, PTPN7, PLAT, PKD1L1, ZBTB16, EPHB6, PLAGL1, LRRC71, PTPN14, ACHE, ELFN1, ZBTB42, PRR5, CNTNAP3, OLFML2A, LTBP4, DDC, ZHX2, LSP1, LPIN3, BCL6, TCAP, ADCY1, NPC2, CITED4, ANXA2, HES1, SEMA3G, ANO4, SYT2, FBLN5, KIAA0513, IGIP, ZBTB8B, NXPH1, TFE3, EPHA4, DARC, GBP2, CCNJL, ICAM1, KLHL34, SELPLG, VGLL3, KCNC4, ELN, SLC03A1, CRISPLD2, LFNG, CTSK, HLA-F, TLN2, HMGCL, CAMTA1, CD84, VAMP8, AIM1, CACNA1D, CCDC109B, SLC46A3, B3GNT7, FAM115C, ALPK3, DOCK1, SPSB1, FNDC5, KIAA1522, NOTCH2, ITGB5, ERBB2, SOX15, APOL6, PTPRH, KRBA2, SLC9A3R1, MYO10, SLAMF8, PIDI, GPR123, VSTM4, HK3, CCDC24, HRH1, PSD4, APBA1, GIMAP8, COL7A1, CDH3, C1orf173, EPB41L3, IL32, SDC3, CBX7, ABLIM3, ARHGEF10L, KIAA0319, MYO7A, CYBRD1, CAPN5, PBXIP1, SORBS2, CD97, LAMA4, S100A16, CYP561, COL6A6, ZNF185, DOCK3, RPH3AL, CPQ, SEPT4, IRF8, TNFAIP3, TMEM234, LCP1, CTSS, ZBTB7C, CTSH, DKK3, CES4A, LRIG1, HCL51, ALOX12, CADM3, DENND2D, KNDC1, DAB2, TMEM63A, MXRA5, GPPD5, ST3GAL1, NDRG2, HLA-E, ZNF438, RREB1, NRP1, ITPKB, DOC2B, UNC13D, C1QL1, FGD6, UTRN, TOM1L2, HID1, SYT5, SYT7, BCAR3, TMEM173, NR3C1, CELF2, TMBIM1, EPB41L4B, CTNBP1, OSBPL10, SLC37A1, GPPEP1, PSD2, TFE3, HIVEP3, LHFPL2, CPEB4, S1PR3, CD163L1, CD4, SDK1, ZBTB20, DNALI1, ZBTB4, MOB3C, ALDH3A2, KLHL36, DPP6, TSHZ2, CROT, CASP9, PRKAG2, ITGA3, ABLIM2, AGBL4, TMEM130, AC135178.1, KANK2, ABCA1, MAFB, CAMK2B, CCND2, ABCG4, KCNAB2, CORO2B, NLRP1, RPS6KA1, KLF11, FCHSD1, ARHGAP23, MYO1C, FRMPD3, CPNE2, ZNF423, STON2, ADAP2, ITGB2, PILRA, ZEB2, RAP1GAP2, PLD2, KCTD21, FAM214B, ZBTB48, LRRC48, SYNE3, AHDC1, JAG1, TMEM150C, RAB15, EFNB2, MKNK1, IRF1, MPEG1, MERTK, TENC1, MAST4, CMKLR1, STAT3, TVP23C, PLCL1, PPIE, BCAP29, CSF1R, CLCN6, OPTN, ECE1, PLXDC2, C11orf63, DUOX1, ACE, TNS1, TRIM62, H6PD, GPR153, CRTC2, MAPK4, FMNL1, TTC7A, NBPF1, KIAA0319L, DOCK9, NECAP2, RORA, MAP7D1, KIF1C, VPS13D, STAT2, ZNF800, PHLDB1, GPN2, TPRG1L, SLC02B1, CROCC, RUND3B, RIN2, ATXN1, HECTD3, CLTCL1, SIRPA, NDEL1, NRBP2, EVC, MINK1, SLC9A1, ANKFY1, RAPGEFL1, INPP5K, SMG6, GSTK1, ACADVL, TCP11L2, MTF1, ZFYVE1, FLCN, MYO1D, CTC1, PLEKHM1, DD2, NAGK, ITS2, INPP5B, WIPF2, CALCOCO1, ADAM19, PITPNA, SZT2, SMIM12, KSR1, VPS53
NB_s4S_4M ⁺	Down	ABCC1, CAD, PPRC1, IGF2BP3, HNRNPA1, MTHFD1L, QSOX2, TTLL12, XPO5, ELOVL6, FAM110A, ST6GAL1, RPL7, HOXD3, HNRNPD, PRMT3, TACC3, KCNT2, GTF3C2, RPS7, TCF3, SFXN4, RNF150, FANCA, ITPRIP1, ERCC6L, CDCA7, HMGAI1, RRP1B, AHCY, FBL, NAT8L, SMC4, GPR125, DDX39A, NOB1, E2F1, TRAP1, PKMYT1, PHF21B, FANCI, PLK1, CDCA5, CASC5, VARS, ZNF581, CAMKV, GINS2, FBXO5, DHFR, HMGB3, HSPD1, NPM1, LSM4, SLC7A5, MYBL2, SLC05A1, NETO2, KIF14, CDCA2, ARHGAP11A, BUB1, MPV17L, CHEK1, HJURP, BZW2, TNFRSF10B, TICRR, ECT2, PUS7, NCAPG, ZNF695, UBE2S, FOXM1, ASPM, GTSE1, SH3GL3, KIF18A, TMEM97, PEX5L, SLC1A5, RCOR2, PAICS, BUB1B, MKI67, KIF4A, MTHFD2, MAD2L1, CCNB1, CBS, PSAT1, KIF15, LDHAL6A, EXO1, YPEL5, FRS2, TPX2, CEP44, CEP55, UBE2C, CPLX1, DSCC1, DEPDC1B, EHD3, CUX2, CBLN1, PRSS12, CDT1, E2F3, CENPF, LRRN1, DSCAM, LBH, HIST2H3D, LCLAT1, H2AFX, GAL, HIST1H2AB, NCR3LG1, HOXD9, TRIB2, CLIP4, GABRA3, HIST1H2AL, MDM2, YEATS4, NCAN, ADCY8, TWIST1, LPPR5, LPIN1, ENDOG, NBAS, RET, SGCZ, FAM19A5, WDR43, SLC30A3, DUSP4, ISLR2, GALNT14, ALK, LIX1, VGF, ANO3, CPSF6, CCT2, MYCN, CRABP1, DDX1, PENK, HOXD10, FAM49A
NB_s4S_4M ⁻	UP	KRT10, KRT14, KRT1, FLG, KRT5, DSG1, DSP, DSC3, SFN, PLP1, COL3A1, MATN2, DRD2, TGFBI, PLXNA4, TCF7L2, EPB41L3, DCC, PVRL1, CREB5, FAXDC2, GAB2, ABR, PHLDB1, KSR1

Tabelle D.3.: Ermittelte Gensignaturen aus den stadienbasierten Vergleichen zwischen den klinischen Stadien der Neuroblastomproben. Die Anordnung der Gene beruht auf den logFC-Werten der Vergleiche der differentiellen Expressionsanalyse zwischen den Stadien. Für die Signaturen mit den Namen NB_s4S und NB_s4S_4M⁻ konnten keine Down-Sets ermittelt werden.

D.2. Analyse der Signaturen

Im digitalen Anhang dieser Arbeit sind die vollständigen Analyseergebnisse der GSEA enthalten, deren Darstellung innerhalb der Arbeit aufgrund der Datenmenge vermieden wurde. Im Verzeichnis `Transkriptomanalyse/Gensignatur/GSEA_Ranglisten/` befinden sich die erzeugten Ranglisten der einzelnen Signaturen, während im Verzeichnis `Transkriptomanalyse/Gensignatur/GSEA_Analyse/` die Ergebnisse der Analyse gespeichert sind.

E. Öffentliche Daten

Aus Platzgründen wurden die Expressionsdaten der öffentlichen Daten an dieser Stelle nicht aufgelistet. Der digitale Anhang zu dieser Arbeit enthält im Verzeichnis `Transkriptomanalyse/Public/Expression/` die Daten zur Veränderung der Genexpression aus den Vergleichen zwischen den Neuroblastomstadien für den R2-Seq und den R2-Array Datensatz. Die Ranglisten und die Ergebnisse aus der Validierung der in dieser Arbeit ermittelten Gensignaturen sind in den Verzeichnissen `Transkriptomanalyse/Public/Ranglisten/` und `Transkriptomanalyse/Public/Validierung/` gespeichert.

F. *microRNA-Targets*

F.1. Informationen zu den *microRNA*-Bindestellen

Die Anzahl der *microRNA*-Bindestellen (MREs) für die annotierten Gene der öffentlich verfügbaren Datenbank TargetScan in Version 7.1 kann aus dem digitalen Anhang der Datei `hg19_mRNA_longest_3UTR_lengths.csv` im Verzeichnis `Transkriptomanalyse/Targets/` entnommen werden. Die Tabelle beinhaltet die Namen und IDs der einzelnen Gene sowie die Länge des längsten 3'UTRs der zugehörigen Transkripte. Zudem sind die Anzahl der vorhergesagten *microRNA*-Bindestellen und die längennormalisierten Anzahl der Bindestellen angegeben.

F.2. Tabellen der *microRNA*-Targetanalyse

Auf eine tabellarische Auflistung aller ermittelten *microRNA*-Target Beziehungen der einzelnen Vergleiche der differentiellen Expression dieser Arbeit wurde an dieser Stelle aus Platzgründen

verzichtet. Die Tabellen der microRNA-Target Beziehungen der einzelnen Vergleiche können für die hoch- und runterregulierten microRNAs sowie deren putative Targets aus den Dateien im Verzeichnis Transkriptomanalyse/Targets/ entnommen werden. Die einzelnen Tabellen beinhalten die Namen der Targets und microRNAs, die Informationen aus der differentiellen Expressionsanalyse für die Targets als auch die genutzten Informationen aus der Überlebenszeitanalyse der klinischen Daten, der microRNA-Bindestellen und dem Score zur Gesamtbewertung der jeweiligen Beziehung.

F.3. Bestenliste der microRNA-Targetanalyse

Vergleich	Target ↓	microRNA ↑	Score _{pdB}	Target ↑	microRNA ↓	Score _{pdB}
1/2/3 – 4/4S	RASD1	miR-375	0,267	FRS2	miR-183-5p	0,288
	FAM124A	miR-1246	0,149	FRS2	miR-182-5p	0,288
	FAM124A	miR-375	0,078	PHF6	miR-182-5p	0,236
				PHF6	miR-183-5p	0,236
				SEMA3A	miR-183-5p	0,213
				LIN28B	miR-493-5p	0,198
				LIN28B	miR-183-5p	0,198
				LIN28B	miR-146a-5p	0,198
				WDR43	miR-493-5p	0,189
				LPIN1	miR-182-5p	0,186
1 – MYCN	ZFPM2	miR-222-3p	0,396	EZH2	miR-26a-5p	0,494
	ZFPM2	miR-106b-5p	0,389	IDH2	miR-183-5p	0,428
	ZFPM2	miR-221-3p	0,389	EZH2	miR-26b-5p	0,388
	ZFPM2	miR-19a-3p	0,389	CREBZF	miR-26b-5p	0,378
	TGFBR2	miR-20a-5p	0,372	ADM	miR-26b-5p	0,377
	LMNA	miR-9-5p	0,370	H2AFX	miR-24-3p	0,376
	TGFBR2	miR-19b-3p	0,365	LBR	miR-148a-3p	0,374
	KLF2	miR-92a-3p	0,363	DNMT1	miR-148a-3p	0,372
	ZFPM2	miR-20a-5p	0,360	RPS6KA6	miR-26b-5p	0,367
	ZFPM2	miR-106a-5p	0,360	HMGA1	miR-26a-5p	0,352
1 – 4	LMNA	miR-9-5p	0,363	EZH2	miR-26a-5p	0,494
	ZBTB4	miR-19b-3p	0,353	H2AFX	miR-24-3p	0,376
	ATXN1	miR-19a-3p	0,350	LBR	miR-148a-3p	0,374
	ATXN1	miR-19b-3p	0,350	DNMT1	miR-148a-3p	0,372
	ANKRD52	miR-93-5p	0,347	CLP1	miR-98-5p	0,368
	THBS1	miR-19a-3p	0,344	E2F3	miR-125b-5p	0,358
	MINK1	miR-20a-5p	0,342	CCDC47	miR-196a-5p	0,358
	MINK1	miR-106a-5p	0,342	TIMM17B	miR-98-5p	0,349
	TOM1L2	miR-19b-3p	0,326	CDC25A	miR-98-5p	0,345
	ANKRD52	miR-20a-5p	0,312	RAVER2	miR-99a-5p	0,343
1 – 4 ^{+M}	ZFPM2	miR-222-3p	0,396	ADM	miR-26b-5p	0,377
	ZFPM2	miR-19a-3p	0,389	LBR	miR-148a-3p	0,374
	ZFPM2	miR-106b-5p	0,389	DNMT1	miR-148a-3p	0,372
	LMNA	miR-9-5p	0,370	RPS6KA6	miR-26b-5p	0,367
	ZFPM2	miR-20a-5p	0,360	E2F3	miR-125b-5p	0,358
	ZFPM2	miR-106a-5p	0,360	KDM3A	miR-30e-5p	0,353
	ZBTB4	miR-19b-3p	0,353	HMGA1	miR-26a-5p	0,352
	ZFPM2	miR-19b-3p	0,353	HNRNPA1	miR-149-5p	0,348
	ATXN1	miR-19a-3p	0,350	ADM	miR-26a-5p	0,342
	ATXN1	miR-19b-3p	0,350	ABCB7	miR-148a-3p	0,341
1 – 4 ^{-M}	ZBTB4	miR-19b-3p	0,353	PCDH8	miR-182-5p	0,420
	ANKRD52	miR-93-5p	0,347	FBXW7	miR-27a-3p	0,412
	TOM1L2	miR-19b-3p	0,326	FGFR3	miR-99a-5p	0,401
	SERPINE1	miR-301a-3p	0,301	CLP1	miR-98-5p	0,368
	LDLR	miR-93-5p	0,298	TIMM17B	miR-98-5p	0,349
	LDLR	miR-19b-3p	0,298	CDC25A	miR-98-5p	0,345
	MAP7	miR-93-5p	0,295	YOD1	miR-98-5p	0,344
	CHD5	miR-301a-3p	0,294	RAVER2	miR-99a-5p	0,343
	ZBTB4	miR-93-5p	0,290	MET	miR-23b-3p	0,341
	ZEB2	miR-25-3p	0,287	TAF5	miR-98-5p	0,336
1 – 4S				KRT1	miR-146a-5p	0,238
				COL17A1	miR-183-5p	0,183
				PKP1	miR-125a-5p	0,181
				PKP1	miR-182-5p	0,181

				PKP1	miR-183-5p	0,181
				KRT10	miR-146a-5p	0,177
				KRT6B	miR-146a-5p	0,177
				DSC3	miR-125a-5p	0,145
				FLG	miR-125a-5p	0,142
				KRT6B	miR-183-5p	0,142
4 – 4S	E2F3	miR-204-5p	0,145	MINK1	miR-20a-5p	0,342
	CDT1	miR-204-5p	0,113	MINK1	miR-106a-5p	0,342
	RRM2	miR-204-5p	0,112	COL1A2	miR-92a-3p	0,318
	PRR11	miR-204-5p	0,111	MINK1	miR-17-5p	0,307
	CDCA2	miR-204-5p	0,078	CHD5	miR-20a-5p	0,301
	HJURP	miR-744-5p	0,078	MAP7	miR-106a-5p	0,295
				ZBTB4	miR-17-5p	0,290
				ZBTB4	miR-106a-5p	0,290
				ZBTB4	miR-20a-5p	0,290
				ZEB2	miR-92b-3p	0,287
4⁺M – 4S	HNRNPA1	miR-204-5p	0,313	ZFPM2	miR-222-3p	0,396
	HNRNPA1	miR-129-5p	0,313	ZFPM2	miR-221-3p	0,396
	IGF2BP3	miR-196a-5p	0,296	NDEL1	miR-130b-3p	0,375
	ELOVL6	miR-204-5p	0,280	ZBTB4	miR-130b-3p	0,360
	FAM19A5	miR-196a-5p	0,259	ZFPM2	miR-20a-5p	0,360
	FAM19A5	miR-196b-5p	0,259	ZFPM2	miR-106a-5p	0,360
	LRRN1	miR-181c-5p	0,256	ZFPM2	miR-130b-3p	0,360
	FRS2	miR-204-5p	0,252	ZEB2	miR-221-3p	0,358
	FRS2	miR-129-5p	0,252	ATXN1	miR-221-3p	0,357
	E2F3	miR-148a-3p	0,245	ZFPM2	miR-454-3p	0,353
4⁻M – 4S				KRT1	miR-92b-3p	0,238
				PVRL1	miR-92b-3p	0,220
				MATN2	miR-92b-3p	0,203
				MATN2	miR-1246	0,203
				ABR	miR-92b-3p	0,195
				DCC	miR-92b-3p	0,187
				DSG1	miR-92b-3p	0,177
				CREB5	miR-1246	0,152
				CREB5	miR-3609	0,152
				MATN2	miR-92b-3p	0,132

Tabelle F.4.: Auszug der ermittelten microRNA-Target Beziehungen aus der Targetanalyse. Gelistet sind die maximal zehn besten putativ direkten microRNA:mRNA Bindungen für die hoch- und runterregulierten microRNAs der differentiellen Vergleiche zwischen den Tumorstadien. Bisher nicht mit dem Neuroblastom assoziierte microRNAs oder Targets sind als fett gedruckt gekennzeichnet. Um die Dimension der Tabelle gering zu halten, wurden die Namen der microRNAs in verkürzter Schreibweise, ohne des Präfix „hsa-“, dargestellt.

Eidesstattliche Erklärung / Declaration under Oath

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Dissertation wurde weder in gegenwärtiger bzw. in anderer Form einer anderen Fakultät zur Begutachtung vorgelegt.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content. This thesis has not been submitted for review to another faculty either in the presented or any other form.

Danny Misiak

Halle (Saale), den 26. November 2018

Lebenslauf

Persönliche Angaben

Name	Danny Misiak
akademischer Grad	Diplom-Bioinformatiker
Geburtsdatum	11. April 1983
Geburtsort	Aschersleben
Geschlecht	männlich
Anschrift	Niemeyerstraße 11, 06110 Halle (Saale)
Staatsangehörigkeit	deutsch

Bildungsgang

1989 - 1993	Grundschule Helmut-Welz Aschersleben
1993 - 2002	Gymnasium Ascaneum Achersleben
Mai 2002 - Februar 2003	Zivildienst in den Kreiskliniken Aschersleben-Staßfurt
Oktober 2003 - April 2009	Diplomstudium der Bioinformatik der Martin-Luther-Universität Halle-Wittenberg Schwerpunkt: Bildverarbeitung, Molekulargenetik akademischer Grad: Diplom-Bioinformatiker (Dipl.-Bioinf)
Diplomarbeit	angefertigt in der Arbeitsgruppe von Prof. Dr.-Ing. Stefan Posch am Institut für Informatik
Titel	<i>Automatisierte Detektion fluoreszenzmarkierter Neuronen</i>
seit April 2009	wissenschaftliche Mitarbeiter an der Martin-Luther-Universität Halle-Wittenberg am Institut für Molekulare Medizin, Abteilung Molekulare Zellbiologie
Fachgebiet der Promotion	Bildanalyse und molekulare Zellbiologie

Ich versichere, dass gegenwärtig keine Vorstrafen vorhanden oder Ermittlungsverfahren gegen mich eingeleitet sind.

Danny Misiak

Halle (Saale), den 26. November 2018

Publikationen

- Rashmi Minocha, Varvara Popova, Daria Kopytova, Danny Misiak, Stefan Hüttelmaier, Sofia Georgieva, and Katja Sträßer. Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Research*, 46(18):9749–9763, 2018.
- Simon Müller, Nadine Bley, Markus Glaß, Bianca Busch, Vanessa Rousseau, Danny Misiak, Tommy Fuchs, Marcell Lederer, and Stefan Hüttelmaier. IGF2BP1 enhances an aggressive tumor cell phenotype by impairing miRNA-directed downregulation of oncogenic factors. *Nucleic Acids Research*, 46(12):6285–6303, 2018.
- Fengbiao Zhou et al. AML1-ETO requires enhanced C/D Box snoRNA/RNP formation to induce self-renewal and leukaemia. *Nature Cell Biology*, 19:844–855, 2017.
- Bianca Busch, Nadine Bley, Simon Müller, Markus Glaß, Danny Misiak, Marcell Lederer, Martina Vetter, Hans-Georg Strauß, Christoph Thomssen, and Stefan Hüttelmaier. The oncogenic triangle of HMGA2, LIN28B and IGF2BP1 antagonizes tumor-suppressive actions of the let-7 family. *Nucleic Acids Research*, 44(8):3845–3864, 2016.
- Birgit Möller, Markus Glaß, Danny Misiak, and Stefan Posch. MiToBo - A toolbox for image processing and analysis. *Journal of Open Research Software*, 4(1):e17, 2016.
- Juliane Braun, Danny Misiak, Bianca Busch, Knut Krohn, and Stefan Hüttelmaier. Rapid identification of regulatory microRNAs by miTRAP (miRNA trapping by RNA in vitro affinity purification). *Nucleic Acids Research*, 42(8):e66, 2014.
- Danny Misiak, Stefan Posch, Marcell Lederer, Claudia Reinke, and Stefan Hüttelmaier. Extraction of protein profiles from primary neurons using active contour models and wavelets. *Journal of Neuroscience Methods*, 225:1–12, 2014.
- Birgit Möller and Danny Misiak. SnakeOptimizer - Object segmentation with parametric active contours in ImageJ. In *ImageJ User & Developer Conference*, Mondorf-les-Bains, Luxembourg, 2012.
- Danny Misiak. Automatische Detektion primärer Neuronen in Fluoreszenzmikroskopbildern. *Informatik Spektrum*, 33(3):315–320, 2010.
- Danny Misiak, Stefan Posch, Nadine Stöhr, Stefan Hüttelmaier, and Birgit Möller. Automatic analysis of fluorescence labeled neurites in microscope images. In *Workshop on Applications of Computer Vision (WACV)*, Snowbird, UT, USA, 2009.

