

MODIFIKATIONEN DER
CHROMATINKONFORMATION ALS FOLGE UND
URSACHE STRUKTURELLER
CHROMOSOMENVERÄNDERUNGEN

DISSERTATION ZUR ERLANGUNG DES AKADEMISCHEN GRADES DES DOKTORS DER
NATURWISSENSCHAFTEN (Dr. rer. nat.)

EINGEREICHT IM FACHBEREICH BIOLOGIE, CHEMIE, PHARMAZIE
DER FREIEN UNIVERSITÄT BERLIN

VORGELEGT VON

ANNE STEININGER

AUS BERLIN

BERLIN, 2016

Die vorliegende Arbeit wurde von Juli 2008 bis März 2016 in der Arbeitsgruppe
„Molekulare Zytogenetik“ (Abteilung Molekulare Humangenetik) am
Max-Planck-Institut für molekulare Genetik unter der Leitung von PD Dr. Reinhard
Ullmann erstellt.

1. Gutachter: PD Dr. Reinhard Ullmann
Abteilung Molekulare Humangenetik
Max-Planck-Institut für molekulare Genetik
2. Gutachter: Prof. Dr. Rupert Mutzel
Institut für Biologie - Mikrobiologie
Freie Universität Berlin

Datum der mündlichen Prüfung: 15.07.2016

Danksagung

Seit meinem ersten Tag am MPI sind nun schon einige Jahre vergangen, aber mir ist noch sehr gut in Erinnerung geblieben, wie freundlich ich in der AG Ullmann von allen Mitgliedern aufgenommen worden bin. Insbesondere bei meinem Doktorvater, PD Dr. Reinhard Ullmann, möchte ich mich für seinen großen Einsatz für mich, aber auch für alle anderen Arbeitsgruppenmitglieder, bedanken. Ob man mit Fragen oder Problemen kam, immer nahm er sich dafür Zeit und gab Anregungen und Vorschläge. Mir ist bewusst, dass dies in diesem Umfang nicht selbstverständlich ist und möchte daher ihm gegenüber ausdrücklich meinen Dank aussprechen.

Meinen Mit-Doktorandinnen Grit Ebert und Vivien Boldt möchte ich für ihre besondere Hilfsbereitschaft, den Gedankensaustausch und die Koffeinzufuhr danken. Egal ob wir gemeinsam Zellkulturmarathons veranstaltet oder gegen Wasserfälle im Labor gekämpft haben, unser Zusammenhalt dabei hat mir in den letzten Jahren sehr geholfen und Kraft gegeben. Auch die Zusammenarbeit mit Katrin Tebel ist ein wichtiger Bestandteil in dieser Zeit gewesen, ohne die einige *in silico* Ansätze in dieser Arbeit nicht möglich gewesen wären. Für das ausgezeichnete Arbeitsklima waren aber noch weitere Personen am MPI verantwortlich, denen ebenfalls mein Dank gilt: Artur Muradyan, Ines Müller, Melanie Heßler, Alischo Ahmad, Vanessa Suckow und Sophia Schade zeichnen sich auch durch ihre Hilfsbereitschaft und Freundlichkeit aus. Des Weiteren möchte ich mich bei unseren Kooperationspartnern, insbesondere Dr. Markus Möbs und PD Dr. Chalid Assaf für die konstruktive und angenehme Zusammenarbeit bedanken, außerdem bei Prof. Dr. Rupert Mutzel für die Begutachtung meiner Dissertation.

Mein ganz besonderer Dank gilt meinen Eltern, für ihre unermüdliche Unterstützung und Ermutigung, ihr Interesse an dieser Arbeit und das Lesen der schriftlichen Form. Danke, dass ihr immer für mich da seid! Einen großen Anteil an dieser Arbeit hat Udo Georgi. Angefangen mit dem moralischen Zuspruch den du mir gegeben hast, über wissenschaftliche Ratschläge, Austausch von Laptopdisplays bis hin zu den Wochenenden und Abenden die du mit dem Korrekturlesen dieser Arbeit verbracht hast: Danke für alles!

Inhaltsverzeichnis

Zusammenfassung	1
Summary	3
1. Einleitung	5
1.1. Chromosomen: Struktur und Organisation	5
1.1.1. Die Organisation des Interphasechromosoms	5
1.1.2. <i>Chromosome Conformation Capture</i> zur Bestimmung der Chromosomenarchitektur	7
1.1.3. Chromatindomänen des Interphasechromosoms	10
1.2. Strukturelle Veränderungen von Chromosomen	15
1.2.1. Molekulare Mechanismen zur Entstehung struktureller Chromosomenveränderungen	16
1.2.2. Genomische Risikofaktoren struktureller Chromosomenveränderungen	19
1.2.3. Evolutionäre Chromosomenveränderungen am Beispiel des Chromosom 7 während der Primatenevolution	22
1.2.4. Strukturelle Chromosomenveränderungen als Ursache von <i>Genomic Disorders</i> am Beispiel der WBS-kritischen Region	24
1.2.5. Strukturelle Chromosomenveränderungen in Tumoren am Beispiel der kutanen T-Zell-Lymphome	26
1.3. Zielstellung	29
2. Material und Methoden	31
2.1. Material	31
2.1.1. Biologisches Material	31
2.1.2. Chemikalien und Medien	34
2.1.3. Puffer und Lösungen	36

2.1.4.	Kits (Reagentsysteme)	36
2.1.5.	Mikroarrays	37
2.1.6.	Geräte	37
2.1.7.	Verbrauchsmaterialien	38
2.1.8.	Software	39
2.1.9.	Öffentliche Daten und Datenbanken	39
2.2.	Methoden	42
2.2.1.	Auswertung der intrachromosomalen Interaktionen des IMR90 Chromosom 7	42
2.2.2.	Zellkultur	48
2.2.3.	Chromatinimmunpräzipitation	49
2.2.4.	Analyse der früh-apoptotischen DNA-Degradation	50
2.2.5.	Mikroarrayhybridisierung der H4K8ac-ChIP- und DNA-Degradationsexperimente	51
2.2.6.	RNA-Expressionsprofiling	52
2.2.7.	Feinkartierung evolutionärer Bruchpunkte und <i>in silico</i> Revertierung der Inversionen	52
2.2.8.	Analysen zur Syntenie des humanen Chromosom 7 und zur Anreicherung von SDs, <i>Alu</i> -Repeats und G4-Motiven	53
2.2.9.	Weitere Analysen im Zusammenhang der IMR90-Chromatinorganisation des Chromosom 7 bzw. der WBS-kritischen Region	53
2.2.10.	Vorbereitung der Sézary-Tumorzellen der Patienten	55
2.2.11.	Array CGH der Sézary-Proben	56
2.2.12.	FISH	59
2.2.13.	Immunhistochemie	59
2.2.14.	Transfektion, RNA-Isolierung und <i>real-time</i> PCR der Seax-Zellen	60
2.2.15.	EMSA	60
2.2.16.	Proliferationsanalysen	61
2.2.17.	Genexpressionsarray der Zelllinie Seax	61
2.2.18.	Hi-C der Zelllinie Seax	63
2.2.19.	Prozessierung und Kontrolle des Seax-Hi-C-Datensatzes	67
2.2.20.	Aufschlüsselung der Zusammenarbeit	72

3. Ergebnisse	75
3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7	75
3.1.1. Anreicherung von SDs in intrachromosomal interagierenden Regionen des Chromosom 7	79
3.1.2. Evolutionäre Rearrangements und intrachromosomale Interaktionen des Chromosom 7	83
3.1.3. Pathogene Chromosomenveränderungen und intrachromosomale Interaktionen des Chromosom 7	86
3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms	93
3.2.1. DNA-Kopienzahlveränderungen von Sézary-Tumorzellen	93
3.2.2. Die 19p13.3-Deletion der Sézary-Tumorzellen resultiert in einem Verlust von E2A	97
3.2.3. Der Verlust von E2A und die Auswirkungen auf Zell-Proliferation und Genexpression	102
3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax	107
3.3.1. Intra- und interchromosomale Interaktionen der Seax-Zelllinie weisen auf Rearrangementbruchpunkte hin	107
3.3.2. Kartierung der Seax-Translokationen durch interchromosomale Interaktionen	112
3.3.3. Vergleich der identifizierten Fusionsregionen mit publizierten Rearrangements der Zelllinie Seax	121
3.3.4. Genom- und Chromatineigenschaften der Translokationsbruchpunkte	126
3.3.5. Die 19p13.3-Deletion und chromosomale Interaktionen	134
4. Diskussion	137
4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen	137
4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen .	145
5. Literatur	157
6. Verzeichnis der aus dieser Arbeit erfolgten Publikationen	185
7. Anhang	187

A. Curriculum Vitae	211
B. Selbständigkeitserklärung	213

Abbildungsverzeichnis

1.1. Schema eines Hi-C-Experimentes	9
1.2. Die Organisationsebenen der Chromosomen	13
1.3. Die Organisation von TADs in A- und B-Kompartimenten	14
1.4. Mechanismen zur Entstehung struktureller Chromosomenveränderungen . .	18
1.5. Inversionen des Chromosom 7 während der Primatenevolution	23
1.6. Die <i>Genomic Disorders</i> der 7q11.23-Region	25
1.7. T-Zell-Migration neoplastischer T-Zellen in die Epidermis	28
2.1. Relative Häufigkeiten der Interaktionsanzahl pro Binpaar für das Chromo- som 7	43
2.2. Prozessierung der normalisierten IMR90 Hi-C-Daten	45
2.3. Überblick der Prozessierungs- und Normalisierungsschritte des Seax-Hi-C- Experimentes mit HOMER	68
2.4. Kontrollen des Seax-Hi-C-Experimentes	69
3.1. Reproduzierbarkeit des triangulären Interaktionsmusters des Chromosom 7 bei der Anwendung unterschiedlicher Schwellenwerte zur Generierung der Interaktionsbündel	77
3.2. Intrachromosomales Interaktionsmuster des humanen Chromosom 7 über- lappt mit Regionen offenen, transkriptionell aktiven Chromatins und SD- Clustern	80
3.3. Anreicherung von SD-Sequenzen in Regionen der intrachromosomalen In- teraktionsbündel des Chromosom 7	81
3.4. Zu Cj Chromosom 2 syntenische Regionen des humanen Chromosom 7 zeigen eine räumliche Nähe in der Kernorganisation und eine Anreicherung an SD-, <i>Alu</i> - und G4-Motiv-Sequenzen	84

3.5. Schema zur Separation der zu Cj Chromosom 2 syntenischen Region durch die peri- und die nachfolgende parazentrische Inversion des Chromosom 7 während der Primatenevolution	86
3.6. Änderung der Chromatinorganisation an der Grenze der WBS-kritischen Region	88
3.7. SD-Cluster der WBS-kritischen Region inserierten während der Evolution in den homologen Grenzen der topologischen Domänen	92
3.8. Umfangreiche DNA-Kopienzahlveränderungen in Tumorzellproben von 20 Patienten mit Sézary Syndrom	94
3.9. Chromosomale Loci der häufigsten kleinsten gemeinsam veränderten Regionen der Deletionen bzw. Duplikationen und deren Zuordnung zu 20 Sézary-Patienten	96
3.10. Häufigkeit und chromosomale Positionen der 19p13.3-Deletion in Tumorzellen der 20 Sézary-Patienten	98
3.11. <i>E2A</i> -Verlust der Sézary-Zelllinie Seax	101
3.12. Konstrukte zur Rekonstitution der <i>E2A</i> -Aktivität in der <i>E2A</i> -defizienten Zelllinie Seax	102
3.13. Die allgemeine Hochregulation von Genen nach <i>E2A</i> -Rekonstitution in Seax-Zellen	104
3.14. STRING-Netzwerk der nach <i>E2A</i> -Rekonstitution signifikant deregulierten Genprodukte	106
3.15. Vergleich der Interaktionsmatrizen: Chromosomen 6 und 17	108
3.16. Interaktionsmatrix des gesamten Seax-Genoms	110
3.17. Orientierung der chromosomalen Fusionsregionen in den Interaktionsmatrizen	111
3.18. Identifizierung der Fusionsregionen	113
3.19. Chromosomale Veränderungen der Zelllinie Seax: Chromosomen 1 bis 12	114
3.20. Chromosomale Veränderungen der Zelllinie Seax: Chromosomen 13 bis X	115
3.21. Bruchpunktregionen des Rearrangements zwischen den Chromosomen chr5, chr8 und chr10 zeigen eine Rearrangementkette	118
3.22. Genomweite signifikante intra- und interchromosomale Interaktionen der Zelllinie Seax	119
3.23. Beispiele für Rearrangementketten des Seax-Genoms	120
3.24. Das derivative Chromosom aus Segmenten der Chromosomen chr3, chr6, chr9 und chr17	122

3.25. Repräsentation des derivaten Chromosoms über chromosomale Interaktionen	124
3.26. Reziproke Rearrangements zwischen den Chromosomen chr6 und chr17	125
3.27. Chromatineigenschaften der Translokationsbruchpunkte	129
3.28. Genomeigenschaften der Translokationsbruchpunkte	131
3.29. Bruchpunktregionen der Rearrangements im Vergleich zu SD-, <i>self chain</i> -Paaren, ausgesuchten GO-Annotationen und dem DNA-Kopienzahlstatus in Seax	133
3.30. TADs der 19p13.3-Deletion in Seax und weiteren Zelllinien	136
4.1. Schema der evolutionären Inversionen des Chromosom 7 und der Folgen für die Chromosomenorganisation	140
4.2. Reziproke Translokation zwischen <i>AIG1</i> und <i>GOSR1</i>	148
S1. Intrachromosomales Interaktionsmuster des humanen Chromosom 7 über- lappt mit Regionen früh-apoptotisch degradiertes DNA	189
S2. Intrachromosomales Interaktionsmuster des humanen Chromosom 7 ohne Bündelung der Bindaten	190
S3. LAD-Grenzen kolokalisieren mit Änderungen des Profils der apoptotischen DNA-Degradation und der Histon H4K8-Acetylierung	191
S4. Bruchpunkte der perizentrischen und parazentrischen Inversion des huma- nen Chromosom 7 während der Hominidenevolution	192
S5. Zusätzliche Daten zur Chromatinorganisation der WBS-Region	193
S6. DNA-Degradationsprofile korrelieren mit H4K8-Acetylierungsprofilen in- nerhalb und zwischen den Zelllinien IMR90 und IMR91L	194
S7. DNA-Kopienzahlveränderungen der Sézary-Zelllinie Seax	196
S8. Bestätigung der aCGH-Analyse der Sézary-Tumorzellproben mittels FISH- Analysen	198
S9. Nachweis des E2A-Verlustes auf RNA- und Proteinebene	203
S10. Erhöhte Proteinexpression und E-Box-Bindeaktivität von E2A nach der Rekonstitution in Seax-Zellen	204
S11. Änderung der Zellproliferation bei Rekonstitution von <i>E2A</i> in Seax-Zellen	205
S12. Verifizierung der Genexpressionsarray-Analyse anhand von Beispielgenen	206
S13. Chromosomenkompartimentierung der Chromosomen 6 und 17	207
S14. Vergrößerte Ansicht der Bruchpunktregionen der Rearrangements zwischen den Chromosomen chr5, chr8 und chr10	208

S15. Normalisierte Interaktionsmatrizen der chromosomalen Interaktionen des Chromosom 6 mit den Chromosomen chr2, chr3 und chr17	209
S16. Sequenzeigenschaften der Translokationsbruchpunkte	210

Tabellenverzeichnis

3.1. Übersicht der gefilterten SD-Datensätze für die Randomisierung der SD-Positionen	81
3.2. Bekannte Krebs-assoziierte Gene innerhalb der kleinsten gemeinsam veränderten Regionen der untersuchten DNA-Kopienzahlveränderungen	97
3.3. Chromosomale Bruchpunkte der 19p13.3-Deletionen, welche die kleinste gemeinsam deletierte Region umspannen	99
3.4. Nach <i>E2A</i> -Rekonstitution differentiell exprimierte Gene zeigen eine Anreicherung von Promotoren mit E2A-Bindemotiven	105
3.5. GO-Analyse der 222 signifikant deregulierten Gene: GO Molekulare Funktion	105
3.6. GO-Analyse der 222 signifikant deregulierten Gene: GO Biologische Funktion	105
3.7. Bruchpunktregionen und potentielle Fusionsgene	116
S1. Änderung der Binpaaranzahl nach Anwendung der Filterkriterien	187
S2. Änderung der Anzahl an interagierenden Bins und resultierenden Interaktionsbündel nach Anwendung der Filterkategorien und allen untersuchten Schwellenwerten	188
S3. Anteil der DNA-Kopienzahlveränderungen am Gesamtgenom von Tumorzellproben der 20 Sézary-Patienten	195
S4. Bestätigung der Deletion des <i>E2A</i> -Locus in Sézary-Tumorzellproben mittels FISH-Analysen	197
S5. Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	198
S5. Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	198
S5. Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	199

S5.	Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	200
S5.	Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	201
S5.	Signifikant deregulierte Gene nach <i>E2A</i> -Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie	202

Abkürzungsverzeichnis

3C	Chromosome Conformation Capture
3D	Dreidimensional
7-AAD	7-Aminoactinomycin D
aCGH	Array-based Comparative Genomic Hybridization
adj.	adjustiert
APC	Allophycocyanin
BAC	Bacterial Artificial Chromosome
bHLH	Basische Helix-Loop-Helix
BrdU	Bromdesoxyuridin
ca.	circa
CBS	Circular Binary Segmentation
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag Sequencing
ChIP	Chromatin-Immunpräzipitation
Ci	Curie
Cj	<i>Callithrix jacchus</i>
CML	Chronische Myeloische Leukämie
CNV	Copy Number Variation
cpm	Counts per minute
CRE	Cis-regulatorisches Element
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTCF	CCCTC-Binding Factor
CTCL	Cutaneous T Cell Lymphoma
DM	Double Minute
ECM	Extracellular Matrix
EMEM	Eagle's Minimum Essential Medium
EMSA	Electrophoretic Mobility Shift Assay
FACS	Fluorescence Activated Cell Sorting
FCS	Fetal Calf Serum
FDR	False Discovery Rate
FISH	Fluoreszenz <i>in situ</i> Hybridisierung
GEO	Gene Expression Omnibus
GO	Gene Ontology
HRR	Homologe Rekombinationsreparatur

Hs	Homo sapiens
HSR	Homogeneously Staining Region
IA	Interaktion
IL-2	Interleukin-2
LAD	Lamin-assoziierte Domäne
LCR	Low Copy Repeat
LOCK	Large Organized Chromatin K (lysine) modification
Mb	Megabasenpaare
MeDIP	Methylated DNA Immunoprecipitation
Mf	Mycosis fungoides
MHC	Major Histocompatibility Complex
mHG	minimaler hypergeometrischer Wert
NAHR	Nicht-allelische Homologe Rekombination
NGS	Next-Generation Sequencing
NHEJ	Nonhomologous End Joining
NS	nicht signifikant
PBMC	Peripheral Blood Mononuclear Cell
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
RIDGE	Region of Increased Gene Expression
SD	Segmentale Duplikation
Seq	Sequenzierung
SeS	Sézary Syndrom
SNV	Single Nucleotide Variant
SINE	Short Interspersed Nuclear Element
SRA	Sequence Read Archive
SRO	Smallest Region Of Overlap
SUN	Singly Unique Nucleotides
TAD	Topologically Associating Domain
TCF3	Transcription Factor 3
TCR	T-Cell Rearrangement
tRNA	Transfer-RNA
TSS	Transcription Start Site
u.a.	unter anderen
WBS	Williams-Beuren-Syndrom
z.B.	zum Beispiel

Zusammenfassung

Strukturelle Chromosomenveränderungen sind eine treibende Kraft der Evolution, aber auch Ursache verschiedenster Erkrankungen. Sowohl bei evolutionären als auch bei krankheitsassoziierten Chromosomenrearrangements fällt auf, dass viele der ursächlichen Chromosomenbrüche nicht zufällig im Genom verteilt sind. Vielmehr entstehen durch das Zusammenspiel zwischen unterschiedlicher Chromosomenvulnerabilität und selektivem Druck regionale Präferenzen. Evolutionär zeigen sich diese durch wiederholtes Auftreten derselben Chromosomenbruchpunktregionen in unterschiedlichen Spezies oder durch die Anhäufung segmentaler Duplikationen (SDs) in bestimmten chromosomalen Regionen. In Tumoren bedingt diese nicht zufällige Verteilung der Chromosomenbruchpunkte das Auftreten rekurrenter Aberrationen, welche aufgrund ihrer Häufigkeit und Spezifität zunehmende Relevanz als Startpunkt zur Kandidatengensuche sowie als therapeutische und diagnostische Marker haben. In der Literatur finden sich Hinweise, dass auch die räumliche Organisation der Chromosomen im Zellkern die Häufigkeit der Entstehung von Chromosomenveränderungen beeinflusst. Die gegenseitige Abhängigkeit von dreidimensionaler Chromatinorganisation, Genregulation und Genomintegrität konnte jedoch aufgrund technischer Limitationen bislang nur eingeschränkt untersucht werden. Erst die Einführung von auf Hochdurchsatzsequenzierung beruhenden Methoden eröffnete hier neue Perspektiven. In der vorliegenden Arbeit wurden mit diesen neuen Methoden generierte Daten genutzt, um mittels umfangreicher bioinformatischer Analysen am Beispiel des humanen Chromosom 7 nachzuweisen, dass trotz zweier Inversionen im Zuge der Primatenevolution die dadurch dislozierten chromosomalen Abschnitte ihre räumliche Nähe dennoch bewahren. Die betroffenen Segmente zeigen dabei eine signifikante Anreicherung von segmentalen Duplikationen. Dies könnte darauf hindeuten, dass die räumliche Nähe die Propagation von SDs begünstigt oder SDs sogar zur Aufrechterhaltung dieser räumlichen Nähe beitragen. SDs des Chromosom 7 sind auch an der Entstehung von Chromosomenveränderungen beteiligt, die Ursache des Williams-Beuren-Syndroms sind. Die Analyse der Chromatineigenschaften dieser Region weist darauf hin, dass hier eine topologische Chromatindomäne

mit spezieller Chromatinorganisation vorliegt, an deren Grenze sich SD-Sequenzen während der Evolution inseriert haben. Auch bei den somatischen Chromosomenaberrationen in Tumorzellen gibt es bereits Hinweise auf einen Zusammenhang zwischen chromosomalen Veränderungen, Genregulation und der 3D-Organisation des Zellkerns. Zusätzlich zu der Analyse eines tumorassozierten Kandidatengens, wird am Beispiel eines kutanen T-Zell-Lymphoms mit sehr komplexem Karyotyp untersucht, inwieweit sich chromosomale Rearrangements im Interaktionsmuster widerspiegeln. Mittels molekularbiologischer und bioinformatischer Analysen konnte in dieser Arbeit gezeigt werden, dass sich auf Basis chromosomaler Interaktionsdaten genomweite Rearrangements und deren resultierende Fusionsgene feinkartieren lassen.

Summary

Structural chromosomal aberrations are drivers of evolution, but are also the cause of numerous diseases. The breakpoints of these chromosomal rearrangements are often not randomly distributed. Instead they show regional preferences, which result from a combination of chromosomal vulnerability and selective pressure. In terms of evolution this phenomenon is reflected by the accumulation of segmental duplications (SDs) within distinct segments in the course of evolution and the reuse of evolutionary breakpoint regions. In tumors the non-random distribution of chromosomal breakpoints leads to characteristic patterns of recurrent chromosomal rearrangements. Given their frequency and specificity these aberrations are increasingly used as diagnostic and therapeutic markers as well as starting points for the identification of candidate genes that drive tumorigenesis. During recent years it became apparent that the distribution of chromosomal aberrations is also influenced by the 3D chromatin organisation. Yet, the analysis of the interplay between higher order chromatin organisation, gene regulation and genome integrity was hampered by the lack of appropriate techniques. The advent of methods based on massive parallel sequencing has opened new possibilities to study genome-wide chromosomal interactions with unprecedented resolution. In this thesis, datasets generated by means of these new techniques have been employed to study the mutual influence of evolutionary rearrangements and higher order chromatin conformation. On the example of two inversions of chromosome 7 that have occurred during primate evolution it could be demonstrated by comprehensive bioinformatic analysis that despite these evolutionary rearrangements, the formerly adjacent segments have retained their spatial proximity within the nucleus. The affected segments are characterized by an enrichment of segmental duplications. This observation suggests that spatial proximity could impact the propagation of SDs or alternatively, that SDs have a role in establishing higher order chromatin organisation. Rearrangements within one cluster of SDs located on chromosome 7 are the underlying cause of a genomic disorder named Williams-Beuren syndrome. An analysis of the chromatin organisation within this region in comparison to adjacent regions has revealed that

this interval makes up a distinct chromatin domain and that SDs have preferentially inserted at the borders of this domain. Meanwhile, there is also accumulating evidence that the distribution of somatic chromosomal aberrations in tumors is associated with higher order chromatin organisation. On the example of a cutaneous T-cell lymphoma with a complex karyotype and additional to the analysis of a tumor-associated candidate gene, it has been demonstrated that the presence of chromosomal aberrations can be inferred from alterations within the chromosomal interaction patterns. Moreover, it has been possible to use these interaction data to fine map translocations and identify fusion genes.

1. Einleitung

1.1. Chromosomen: Struktur und Organisation

In den 70iger Jahren des 19. Jahrhunderts untersuchte Walther Flemming den Mechanismus der Zellteilung und identifizierte im Zuge seiner Experimente einen farbstoffaffinen Komplex im Nukleus, das Chromatin. Er erkannte, dass sich das Chromatin während der Zellteilung zu kompakten Strukturen umwandelt, die später von Heinrich Wilhelm Waldeyer als Chromosomen bezeichnet wurden [314]. Tatsächlich liegt diese Struktur der Chromosomen nur in einer kurzen Phase der Zellteilung vor, der Metaphase. Während der Interphase weicht diese stark kondensierte Struktur jedoch einer bei lichtmikroskopischer Betrachtung diffusen Erscheinungsform der Chromosomen. Dass jedoch auch die Interphasechromosomen bestimmten Organisationsprinzipien folgen, erkannte Carl Rabl bereits 1885 und postulierte eine territoriale Organisation der Chromosomen im Nukleus [330]. Lange Zeit blieben die vorgeschlagenen Modelle hypothetisch und konnten mangels geeigneter Techniken nicht experimentell untermauert werden. Erst mit den methodischen Fortschritten im Bereich der Biochemie, Mikroskopie und schließlich der Hochdurchsatzsequenzierung konnten die Erkenntnisse zu dem heutigen Bild der Chromosomenstruktur und -organisation zusammengesetzt werden.

1.1.1. Die Organisation des Interphasechromosoms

Chromosomen setzen sich aus einem Komplex negativ geladener DNA und hauptsächlich basischen Proteinen zusammen, wie z.B. den Transkriptionsfaktoren und Histonen. Oktamere der Histondimere H2A-H2B und H3-H4 formen dabei in Kombination mit der DNA die Nukleosomen. Dabei sind um ein Nukleosom 147 bp der DNA in einer rechtsgängigen superhelikalen Drehung aufgewunden [206] [340]. Benachbarte Nukleosomen werden

1.1. Chromosomen: Struktur und Organisation

durch eine 10 bis 80 bp lange Linker-DNA untereinander verbunden, welche mit einem weiteren Histon, dem Histon H1, assoziiert ist. Diese Ebene der Chromosomenorganisation eines DNA-Protein-Komplexes wurde entsprechend dem vorliegenden Durchmesser als 10-nm-Struktur bezeichnet (*beads-on-a-string* [444]). Abhängig von der intramolekularen Aneinanderlagerung der 10-nm-Chromatinfaser wurde eine regelmäßig gefaltete 30-nm-Struktur als übergeordnete Organisation postuliert, wie z. B. das Solenoid- [118] [343] oder *Zigzag*-Modell [445]. Auch wenn die 30-nm-Struktur in einigen wenigen Zelltypen bestätigt wurde (z. B. in Hühnererythrozyten oder in den Spermatozoiden von Seesternen [443]), konnte diese Chromosomenkonformation selbst mit hochauflösender Kryoelektronenmikroskopie in den meisten untersuchten Zellen nicht *in vivo* belegt werden [105] [298].

Durch Fluoreszenz-in-situ-Hybridisierung-Studien (FISH) konnte nachgewiesen werden, dass die Chromatinfasern nicht zufällig im Nukleus verteilt vorliegen. Vielmehr nehmen die einzelnen Interphasechromosomen innerhalb des Zellkerns distinkte Chromosomenterritorien ein [330] [227] [319] [81] [229], wobei die radiale Position der Chromosomen im Nukleus einem evolutionär konservierten Muster folgt [401]. Große Chromosomen mit einem allgemein hohen Gengehalt liegen dabei präferentiell im Zentrum des Nukleus, während kleine, genarme Chromosomen vermehrt an der Zellkernperipherie positioniert sind [40] [284] [83] [363] [401]. Die Anordnung der 2-4 μm großen Chromosomenterritorien wird innerhalb der G₁-Phase etabliert [53] [75]. Die Positionen können jedoch Änderungen unterworfen werden z. B. während der Tumorgenese [127], Zelldifferenzierung [252], Phytohämagglutinin-Aktivierung bei Lymphozyten [176] oder Seneszenz [42]. Die Chromosomenterritorien werden während der Prophase aufgegeben, daher unterscheiden sich die nach der Zellteilung resultierenden Tochterzellen in der relativen Anordnung der Territorien von der Vorläuferzelle [75]. Auch innerhalb der Chromosomenterritorien zeigen sich Organisationsprinzipien. So bilden z. B. die einzelnen Chromosomenarme und das Zentromer Subkompartimente aus [434]. Auch bei den Chromosomenterritorien sind transkriptionell inaktive Bereiche der Chromosomen eher im Zentrum des Territoriums, aktive Bereiche eher an der Peripherie lokalisiert [161]. Die aktiven Regionen der Chromosomenperipherie können hierbei über Schleifenbildung benachbarte Chromosomen kontaktieren [56]. Diese Kontaktregionen zeigen oftmals eine Anreicherung von Transkriptionsfaktoren und der Transkriptionsmaschinerie und gelten als subnukleare Kompartimente der Genexpression [341]. Die involvierten chromosomalen Regionen weisen im Vergleich zu anderen chromosomalen Bereichen erhöhte Anfälligkeit für strukturelle Veränderungen auf [304].

Während mittels FISH die Existenz der Chromosomenterritorien zwar nachgewiesen wer-

den konnte, blieben die Annahmen über die innere Architektur dieser Territorien lange Zeit spekulativ. Diskutiert wurden unterschiedliche Strukturmodelle von regelmäßig angeordneten bzw. irregulär gefalteten Chromatinfasern [246]. Erst durch die Einführung einer neuen, auf intrachromosomalen Interaktionen beruhenden Methode in Kombination mit Next-Generation-Sequenzierung (NGS), konnte ein Modell bestätigt werden [229].

1.1.2. *Chromosome Conformation Capture* zur Bestimmung der Chromosomenarchitektur

Mikroskopie-basierte Methoden zur Detektion der chromosomalen Architektur erfassen nicht die chromosomale Gesamtstruktur, sondern sind pro Experiment auf wenige Loci beschränkt (z. B. 3D-FISH) bzw. ohne Hinweis auf die zugrundeliegenden Sequenzen (z. B. Elektronenmikroskopie) [30] [215]. Diese Limitationen konnten erst mit der Entwicklung neuer Methoden überwunden werden, die eine sequenzbasierte und genomweite Analyse der Chromosomenorganisation ermöglichen. Durch Methoden auf Basis des *Chromosome Conformation Capture* (3C) kann die DNA-Sequenz selbst für die Analyse der Chromosomenorganisation herangezogen werden und die Ergebnisse der oben genannten Methoden auf dieser Ebene ergänzen. 3C-basierte Methoden erfolgen über ein gemeinsames Grundprinzip: Detektiert wird die räumliche Nähe zwischen verschiedenen chromosomalen Loci über die Häufigkeit ihrer physischen Interaktionen miteinander [90]. Dabei werden DNA-Abschnitte, die sich in räumlicher Nähe (von ca. 10-100 nm [89]) befinden, über die Zugabe von Formaldehyd untereinander quervernetzt (Abb. 1.1). Anschließend werden die Chromosomen fragmentiert und die resultierenden Fragmentenden durch Ligasen miteinander ligiert. Komplexe mit Fragmentenden von verschiedenen chromosomalen Loci bilden hierbei chimäre Ligationsfragmente aus, welche über unterschiedliche Methoden identifiziert werden können. Die Häufigkeit der Kombination von bestimmten DNA-Sequenzen innerhalb der Ligationsfragmente lässt Rückschlüsse auf eine räumliche Nähe ihrer chromosomalen Loci zu. Inzwischen gibt es eine Vielzahl an Variationen dieser Methode, welche sich nicht nur in der Art der Detektion der Ligationsprodukte, sondern auch in ihren Anwendungsmöglichkeiten unterscheiden [25].

Die ursprüngliche 3C-Methode wurde zur Identifizierung der Konformation eines He-fechromosoms entwickelt [90]. Dabei werden die an den chimären Ligationsfragmenten beteiligten DNA-Sequenzen über die Polymerase-Kettenreaktion (*polymerase chain reac-*

1.1. Chromosomen: Struktur und Organisation

tion; PCR) mittels locuspezifischer Primer detektiert. Nachteil dieser Identifizierungsmethode ist, dass die Interaktionen von nur wenigen chromosomalen Loci gleichzeitig verglichen werden können. Weiterhin erschweren Fragmente, die im Zuge des Experimentes nicht ligiert wurden, die Auswertung der Analysen. Die 3C-Methode wurde daher weiterentwickelt und u.a. mit Sequenzierungsmethoden ergänzt. Mittels der 3C-Variante Hi-C ist es daher möglich, genomweite chromosomale Interaktionen zu detektieren und Rückschlüsse auf ein allgemeines Organisationsprinzip der Chromosomen zu ziehen [229]. Diese Erweiterung der 3C-Methodik zeichnet sich durch zwei Modifikationen aus. Zum einen werden die tatsächlich ligierten Fragmente durch eine Biotinselektion angereichert, was den Einfluss von unspezifischen Fragmenten auf die Ergebnisse vermindert. Zum anderen werden die Ligationsfragmente über die *paired-end*-Sequenzierung ausgelesen (siehe Abb. 1.1), was die Erstellung genomweiter Interaktionsprofile ermöglicht.

Für das Experiment werden, wie auch bei anderen 3C-basierten Methoden, zunächst räumlich nahe Chromatinstränge mittels Formaldehyd in ihrer Position fixiert und das Genom nach der Zelllyse mit Typ-II-Restriktionsendonukleasen fragmentiert (z. B. *HindIII*) (Abb. 1.1). An die resultierenden 5'-Überhänge werden biotinmarkierte Nukleotide angehängt. Die anschließende Ligation findet dabei unter stark verdünnten Bedingungen statt, um eine intramolekulare Ligation gegenüber einer intermolekularen zu begünstigen. Die ligierten DNA-Fragmente werden einer Aufhebung der Quervernetzung unterzogen, aufgereinigt und präzipitiert. In dem nachfolgenden Schritt werden die biotinmarkierten Nukleotide der unligierten Enden mittels der Exonukleaseaktivität der T4-DNA-Polymerase entfernt. Da eine Sequenzierung bestimmte Fragmentgrößen bevorzugt, werden die Ligationsprodukte geschert und die resultierenden Fragmente entsprechend ihrer Länge selektioniert (abhängig von der Sequenzierungsmethode, z. B. zwischen 150-300 bp). Eine Anreicherung von biotinylierten Ligationsprodukten kann über die Bindung an Streptavidin-Beads realisiert und somit können unmarkierte Ligationsprodukte verworfen werden. Die verbleibenden DNA-Fragmente werden dann mit Adaptoren für die Sequenzierung versehen, amplifiziert und sequenziert [229] [31] [30] [215]. Bei der *paired-end*-Sequenzierung werden die Ligationsprodukte ausgehend von beiden 5'-Enden für wenige Basenpaare sequenziert. Somit wird jedes chimäre Ligationsprodukt durch zwei kurze Sequenzen, einem Readpaar, repräsentiert. Diese ausgelesenen Sequenzen können nun im Referenzgenom verortet und dadurch ihre Position im Genom bestimmt werden [130]. Häufige Readpaare weisen hierbei auf chromosomale Loci mit geringen räumlichen Distanzen innerhalb des Nukleus hin. Jedoch werden die Rohdaten zusätzlich einer Reihe von Prozessierungs- und Normalisierungsschritten unterzogen. Dabei sollen u. a. Reads, die unspezifische Ligati-

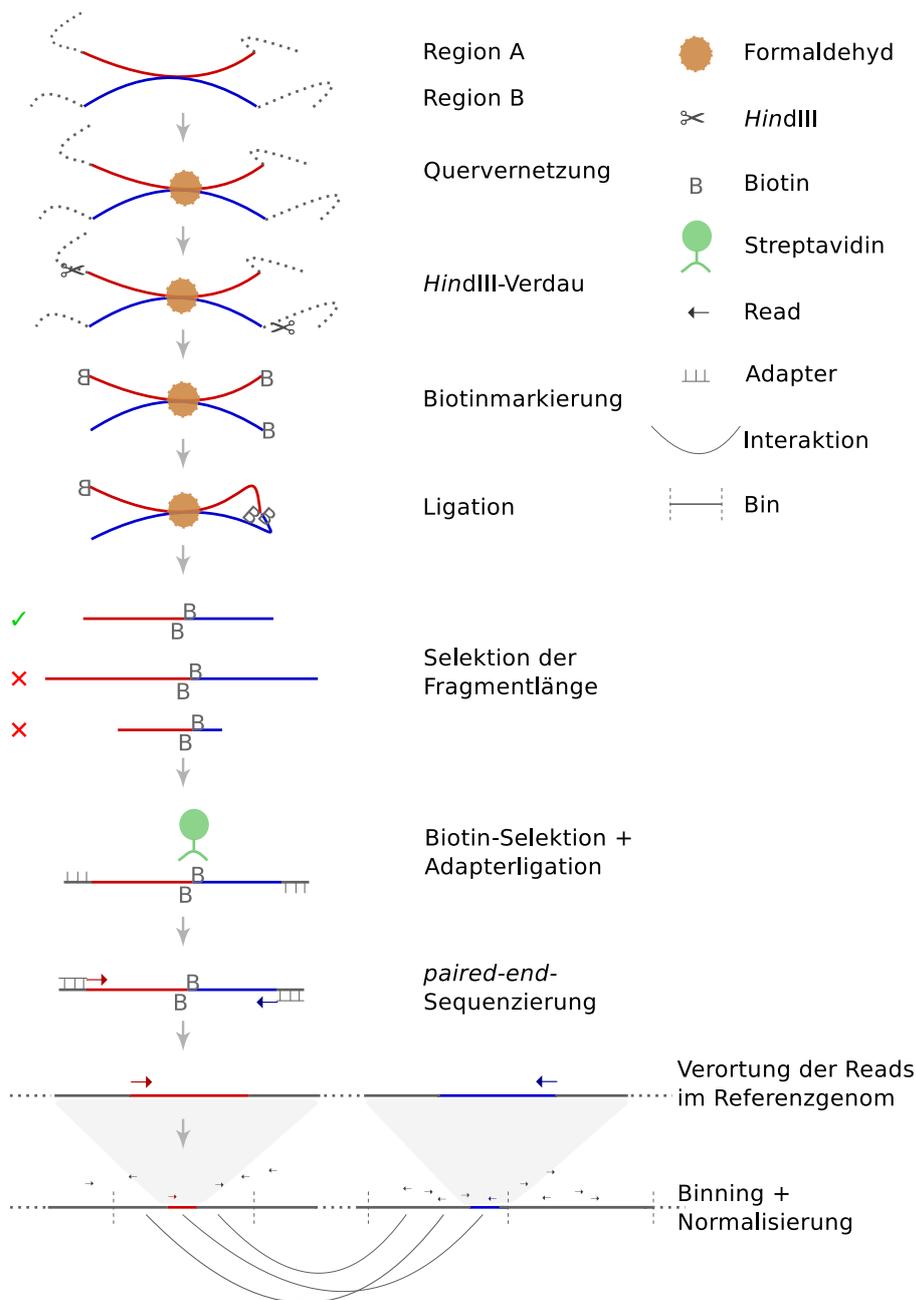


Abbildung 1.1.: Schema eines Hi-C-Experimentes.

Interagierende chromosomale Regionen werden quervernetzt, fragmentiert und an den Fragmentenden mit Biotin markiert. Die anschließende Ligation resultiert u. a. in chimäre Ligationsfragmente. Diese werden entsprechend ihrer Länge gefiltert und über die Biotinselektion von nicht-ligierten Fragmenten getrennt. Nach dem Zufügen von Adaptersequenzen werden die Fragmente *paired-end*-sequenziert. Die resultierenden Reads werden *in silico* im Genom verortet. Die Auswertung der Daten erfolgt nach der Zusammenfassung der Reads in Bins und deren Normalisierung. Modifiziert nach [30].

1.1. Chromosomen: Struktur und Organisation

onsprodukte repräsentieren, verworfen werden.

Durch die Analyse des gesamten Genoms sind theoretisch 10^{12} Kombinationen an *HindIII*-Fragmenten möglich, die in chimären Ligationsfragmenten resultieren. Die häufigsten Ligationsfragmente entstehen jedoch über Selbstligationen einzelner *HindIII*-Fragmente oder aufgrund der direkten Nachbarschaft im linearen Genom, sodass z. B. interchromosomal positionierte Readpaare verhältnismäßig selten auftreten. Somit muss die Komplexität der Ligationsfragmentbibliothek groß genug sein, um auch diese Interaktionen zu repräsentieren. Für eine statistische Repräsentation der interagierenden Regionen werden die Reads in Bins zusammengefasst (Abb. 1.1), welche die Kontaktwahrscheinlichkeiten zwischen chromosomalen Regionen einer definierten Größe wiedergeben [442] [30] [215]. Jedes Bin wird dabei anhand von Backgroundmodellen entsprechend variierender regionaler Sequenzierungsbedingungen normalisiert. So können sich verschiedene Regionen des Genoms aufgrund von z. B. der Anzahl der Restriktionsfragmente oder des GC-Gehalts unterscheiden [442]. Je nach Normalisierungsmethode wird für verschiedene mögliche systematische und biologische Fehlerquellen korrigiert. Das resultierende Verhältnis aus beobachteter Interaktionsanzahl pro Bin (entspricht der Readpaaranzahl) zu dem erwarteten Wert gibt die normalisierte Kontaktwahrscheinlichkeit zwischen einem Binpaar wieder [172] [448]. Eine Darstellung der Kontaktwahrscheinlichkeiten zwischen den Binpaaren erfolgt über eine Kontaktmatrix, wobei jedem möglichen Binpaar die Kontakthäufigkeiten in Form der normalisierten Readanzahlen zugeordnet und als Heatmap visualisiert wird [333].

Mittlerweile wird Hi-C und dessen Varianten genutzt, um Änderungen der Genomorganisation während der Zelldifferenzierung, des Zellzyklus oder der Seneszenz zu detektieren, sowie diese im Rahmen pathophysiologischer Prozesse und der Evolution zu analysieren [263] [179] [57] [291] [91] [426] [25]. Weitere Anwendungen gibt es bei der Haplotypisierung [363] und *de novo* Assemblierung einzelner Genome [47] [190] bzw. von Metagenomen [48]. Eine weitere Anwendungsmöglichkeit bietet die Detektion von chromosomalen Rearrangements [108] [47] [125] [25].

1.1.3. Chromatindomänen des Interphasechromosoms

Lange Zeit war unklar, wie sich die DNA innerhalb der Chromosomenterritorien organisiert. Mehrere konkurrierende Polymer-Modelle zur Organisation der Chromatinfasern wurden vorgeschlagen. Im *equilibrium globule*-Modell liegt das Chromatin diffus, in sich

verknäult und ohne Kompartimentierung im Nukleus verteilt vor [427] [259]. Hingegen postuliert das *fractal globule*-Modell [139] [140] eine selbstorganisierte hierarchische Zusammenlagerung globulärer Strukturen des Chromatins. Wie bei dem mathematischen Fraktalmodell, ähnelt sich diese Organisation auf allen Ebenen selbst. Die globulären Strukturen setzen sich dabei auf der Ebene von Kilobasen bis zu gesamten Chromosomen aus weiteren globulären Strukturen zusammen [229] [276]. Durch diese Organisation von flexiblen globulären Strukturen wäre es dem Chromatin möglich, sich über Brownsche Molekularbewegungen innerhalb des Nukleus zu bewegen. Vorteile einer solchen hierarchischen Organisation wäre eine Zugänglichkeit des Chromatins, die bei dem *equilibrium globule*-Modell nicht gegeben ist [229] [276] [163].

Mithilfe der 3C-basierten Methoden konnte das Modell des *equilibrium globule* bestätigt werden [229]. Die Hi-C-Analysen zeigen, dass sich die Kontaktwahrscheinlichkeiten zwischen zwei chromosomalen Regionen mit Zunahme der linearen genomischen Distanz verringern. Betrachtet man die durchschnittlichen Kontaktwahrscheinlichkeiten zwischen zwei Loci als Funktion der linearen Distanz, so folgt laut [229] das Interaktionsprofil für die Distanzen von 0,5 Mb bis 7 Mb dabei der Potenzgesetz-Verteilung von

$$P(s) \sim s^{-1} \quad (1.1)$$

(mit P - Kontaktwahrscheinlichkeit, s - lineare genomische Distanz zwischen zwei Loci). Eine derartige Verteilung wird am besten durch das Polymermodell des *fractal globule* repräsentiert. Auch die Organisation von Chromosomen in voneinander abgegrenzten Territorien lassen sich mit den distanzabhängigen Kontaktwahrscheinlichkeiten bestätigen [229]. Im Einklang mit den molekular-zytogenetischen Befunden zu den Chromosomenterritorien zeigen Hi-C Interaktionsprofile, dass selbst bei linearen Distanzen von über 200 Mb Regionen innerhalb eines Chromosoms höhere Kontaktwahrscheinlichkeiten aufweisen als interchromosomal positionierte Loci [229]. Das Prinzip der topologisch abgegrenzten Domäne zieht sich dabei über mehrere Ebenen der Chromosomenorganisation.

Chromatin ist in aktiven und inaktiven Regionen organisiert

Im Allgemeinen sind genreiche, aktive chromosomale Bereiche bevorzugt im Zentrum des Zellkerns lokalisiert, während genarme, inaktive Regionen insbesondere an der Zellkernperipherie vorliegen. Dabei sind ca. 30 % des humanen Genoms mit der Lamina der inneren

1.1. Chromosomen: Struktur und Organisation

Kernmembran assoziiert (Lamina-assoziierte Domänen; LADs), einem filamentösen Geflecht aus A- und B-Laminen. Die 0,1 bis 10 Mb großen LADs repräsentieren vorwiegend inaktive, spät-replizierende Bereiche des Genoms und korrespondieren u. a. mit heterochromatischen Regionen der H3K9me3-Hypermethylierung (*large organized chromatin K modifications*; LOCKs) [143] [436] [317] [199] [75]. Die LADs können entsprechend ihrer Konservierung in konstitutive und fakultative LADs eingeteilt werden. Hierbei sind konstitutive LADs stark konserviert, während fakultative LADs zwischen verschiedenen Zelltypen und Spezies variieren [272] [317]. Aber auch bei Laminopathien [263] oder Tumorzellen [265] kann es zu einer Änderung der Domänengrenzen kommen.

Die Positionen der Chromosomen sind u. a. über die Verankerungen mit den LADs im Nukleus fixiert. Die Flexibilität der Chromatinfasern erlaubt jedoch, eingeschränkt durch diese topologischen Zwänge, eine Bewegung chromosomaler Regionen im Raum [53]. Einerseits kann der Repositionierung von aktiven Genen an die Kernmembran eine Verminderung der Genaktivität folgen [334]. Andererseits können sich aktivierte Regionen aus der LAD lösen und in zentraleren Bereichen des Zellkerns positioniert werden [317]. Dabei wurde jedoch gezeigt, dass die Loslösung aktiver Bereiche von der Lamina nicht zwingend zu deren Aktivierung führt [211].

Eine invertierte Organisation der eu- und heterochromatischen Bereiche findet sich in den Nuklei der Photorezeptoren nachtaktiver Tiere. Hier sind heterochromatische Regionen vorwiegend im Zentrum des Zellkerns positioniert, während euchromatische, aktive Bereiche an der Peripherie vorliegen. Aufgrund der höheren Dichte des Heterochromatins wird vermutet, dass dies einer verbesserten Lichtbündelung dient [380].

Auch innerhalb der Chromosomenterritorien können Bereiche mit einem ähnlichen Aktivierungsstatus in Subkompartimenten aggregieren, die mehrere Mb umfassen. So rekrutieren sich die LADs hauptsächlich aus Regionen der B-Kompartimente, welche dementsprechend eher genarme und transkriptionell inaktive Bereiche abdecken. Die offenen A-Kompartimente sind dagegen mit genreichen und aktiven Abschnitten der DNA assoziiert. A-Kompartimente repräsentieren somit euchromatische, B-Kompartimente heterochromatische Regionen des Genoms [229]. Dabei können jedoch noch weitere Unterteilungen in zwei A- und vier B-Kompartimenttypen vorgenommen werden [333]. Regionen des gleichen transkriptionellen Status interagieren dabei präferentiell untereinander und bilden die subnukleare Umgebung unterschiedlicher Genaktivität [229].

Topologisch-assoziierte Domänen bilden die funktionellen Einheiten des Genoms

Die A- und B-Kompartimente setzen sich aus voneinander abgegrenzten Chromatindomänen ähnlicher Aktivität zusammen, den topologisch-assoziierten Domänen (TADs) (Abb. 1.2 B und C) [92] [299]. Diese durchschnittlich 800 kb großen Domänen korrespondieren mit linearen Einheiten koexprimierter Gene, was für eine Genregulation über gemeinsame cis-regulatorische Elemente (CREs) spricht [52] [383] [217] [75] [88]. Gekennzeichnet sind

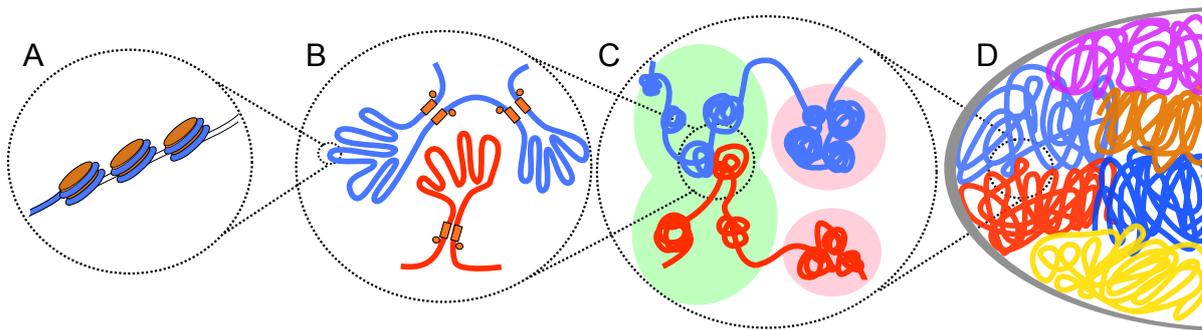


Abbildung 1.2.: Die Organisationsebenen der Chromosomen.

A 10-nm-Chromatinfaser: Komplexe aus DNA und Histonoktameren bilden die Nucleosomen **B** Topologisch-assoziierte Domänen (TADs): Über Grenzelemente topologisch separierte Domänen von unter 1 Mb **C** A- und B-Kompartimentierung: TADs mit aktivem Status arrangieren sich in A-, TADs mit inaktivem Status in B-Kompartimenten **D** Chromosomenterritorien: Chromosomen nehmen im Nukleus voneinander räumlich abgegrenzte Bereiche ein. Modifiziert nach [163] und [114].

diese Domänen durch ihre Präferenz, chromosomale Interaktionen mit Regionen innerhalb der eigenen Domäne auszubilden [92]. Dabei können über Distanzen von mehreren 100 kb lange genomische Interaktionen etabliert werden, die über eine Schleifenbildung Elemente wie Enhancer und Promotoren in räumliche Nähe bringen. Der resultierende Kontakt dieser Sequenzen ermöglicht, z. B. die Transkription in ihren Zielgenen zu initiieren (Abb. 1.3). Die Interaktionen, die innerhalb der Domäne ausgebildet werden können, sind sehr variabel und zelltypspezifisch [289]. Durch die Variabilität der TAD-Interaktionen kann sich beispielsweise der Chromatinstatus einer TAD im Laufe der Zelldifferenzierung ändern und die TAD das übergeordnete Kompartiment wechseln [92]. Dabei konnte gezeigt werden, dass auch die Bindung zelltypspezifischer Transkriptionsfaktoren an der Ausbildung chromosomaler Interaktionen beteiligt ist und damit Einfluss auf die Kernarchitektur nimmt [148]. Bei der Ausbildung der Interaktionen innerhalb der TAD geht man von ineinander verschachtelten Schleifen von etwa 100 kb aus, deren Formierung über Proteine

1.1. Chromosomen: Struktur und Organisation

wie *CCCTC-Binding Factor* (CTCF) und Cohesin unterstützt werden. Da das Bindungsmotiv von CTCF nicht palindromisch ist, kommt dessen Verteilung und Orientierung eine große Bedeutung bei der dreidimensionalen Organisation der DNA zu [333] [88].

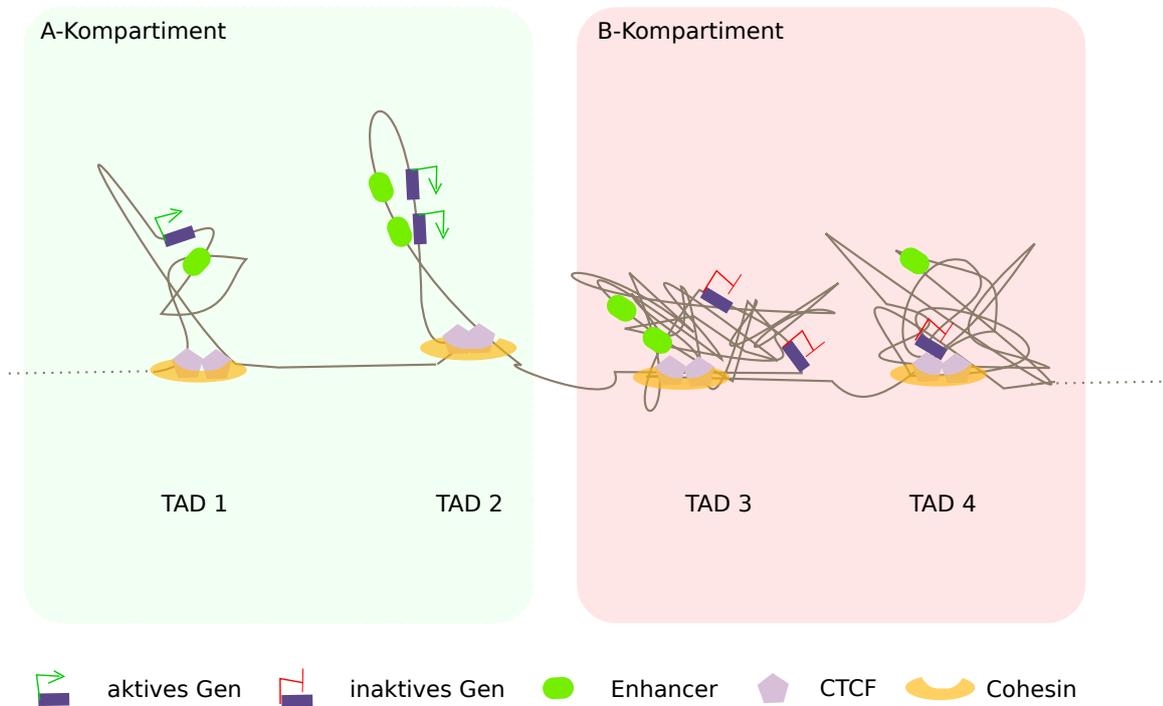


Abbildung 1.3.: Die Organisation von TADs in A- und B-Kompartimenten.

TADs mit einem aktiven Chromatinstatus organisieren sich in offenen A-Kompartimenten, inaktive TADs in kompakten B-Kompartimenten. Durch eine Schleifenbildung innerhalb der aktiven TAD werden Enhancerelemente und Promotoren in eine räumliche Nähe gebracht. Die Ausbildung der Schleifen wird dabei durch die Bindung der Chromatinproteine CTCF und Cohesin an der Basis unterstützt. Modifiziert nach [322] und [75].

Im Gegensatz zu den Interaktionen innerhalb einer TAD sind die relativen Positionen der TADs im Genom evolutionär konserviert und zelltypunabhängig [92] [55] [426]. Eine experimentelle Entfernung der TAD-Grenzen führt zu Fusionen und neu formierten Enhancer-Promotor-Interaktionen zwischen Loci benachbarter TADs. Dies hat Auswirkungen auf die Aktivität der assoziierten Gene und wurde bereits als Ursache genetischer Erkrankungen beschrieben [299] [240].

Wie die TAD-Grenzen aufrechterhalten werden ist noch ungeklärt. Obwohl eine Anreicherung verschiedener Sequenzelemente innerhalb der Grenzregionen vorliegt, konnten noch keine allgemeingültigen Grenzelemente (Insulatoren) identifiziert werden. Zu den potentiellen Grenzelementen gehören u. a. *Housekeeping*- und Transfer-RNA (tRNA)-

Gene, repetitive *short interspersed nuclear elements* (SINEs) und Bindungsmotive für die Chromatinproteine CTCF und Cohesin [92] [299]. Typischerweise finden sich auch ein aktiver Chromatinstatus und transkriptionelle Aktivität an den Grenzen der TADs [415]. Auch wenn sich die Domänen in ihren Charakteristika unterscheiden, so finden sich TAD-ähnliche Strukturen auch in Organismen wie *D. melanogaster*, *C. elegans*, *S. pombe* und *A. thaliana* [91] [430]. TADs stellen somit grundlegende funktionelle Einheiten der Kernarchitektur dar, deren Ausprägung Einfluss auf die weitere Kompartimentierung und Selbstorganisation des Genoms hat [25] [39].

Die 3D-Organisation der Chromosomen ist essentiell für die Genregulation und Genomstabilität. Störungen in der Aufrechterhaltung der Konformation sind mit strukturellen chromosomalen Veränderungen und auch Erkrankungen assoziiert [278]. Welche Faktoren bei der Aufrechterhaltung der Organisation neben den oben beschriebenen Beispielen eine Rolle spielen ist jedoch noch unklar. In Übereinstimmung mit dem Modell der Selbstorganisation der DNA wurden hierfür u. a. repetitive Sequenzen vorgeschlagen, die sich aufgrund ihrer Sequenzsimilarität aneinanderlagern können und innerhalb der Nukleusorganisation räumlich nah positioniert sind. Die Repeatpaarungen könnten so eine sequenzbasierte und zielgerichtete Organisation der Chromosomen bewirken [403] [366] [212]. Damit rücken die Sequenzeigenschaften der DNA wieder in den Vordergrund der Untersuchungen der Chromosomenarchitektur, wobei auch die Beteiligung weiterer noch nicht identifizierter Sequenzen möglich ist.

1.2. Strukturelle Veränderungen von Chromosomen

Die DNA-Sequenz zwischen zwei Individuen gleicher Art variiert in etwa 0,1 % der Basenpaare voneinander [116]. Interindividuelle Unterschiede beschränken sich jedoch nicht nur auf Abweichungen in der Basenpaarabfolge durch z. B. Einzelnukleotid-Varianten (*single nucleotide variants*; SNVs), sondern umfassen auch strukturelle Veränderungen größerer chromosomaler Regionen. Dazu gehören DNA-Kopienzahlveränderungen, wie Deletionen, Duplikationen (*copy number variations*; CNVs), sowie chromosomale Rearrangements, welche Änderungen der Sequenzorientierung (Inversionen) und Position (Translokationen) einbeziehen [327] [112]. Schätzungen zufolge tragen strukturelle Chromosomenveränderungen zu 0,5-1 % der genetischen Variation bei und haben damit den größten Anteil an der genomischen Variabilität (bezogen auf die Gesamtanzahl sich unterscheidender Basenpaare) [435]. Die möglichen Konsequenzen dieser chromosomalen Veränderungen sind

1.2. Strukturelle Veränderungen von Chromosomen

vielfältig und abhängig vom zellulären Kontext. Zumeist sind die Auswirkungen der strukturellen Chromosomenveränderungen phänotypisch neutral, sie können jedoch auch die molekulare Ursache von evolutionärer Diversität und verschiedenen Erkrankungen sein [60] [385] [167]. Unterscheiden kann man zwischen chromosomalen Veränderungen der Keimbahn und somatischer Zellen. Strukturelle Chromosomenveränderungen der Keimbahnzellen können *de novo* entstehen und vererbt werden. Wenn durch deren Auftreten die Genaktivität verändert wird, z. B. durch die Veränderung der Gendosis, kann dies Konsequenzen für den Phänotyp haben. So haben z. B. viele humane Erkrankungen rekurrente Keimbahn CNVs zur Grundlage [412]. Die Chromosomenveränderungen können über resultierende interindividuelle Unterschiede im Phänotyp auch im Zusammenhang mit der natürlichen Selektion stehen [77]. So beruht die genetische Resistenz gegenüber Malaria u. a. auf einer Deletion die gleichzeitig für eine milde Form der Thalassämie prädisponiert [77]. Somatische Chromosomenveränderungen hingegen werden innerhalb eines Lebens akkumuliert und spielen eine zentrale Rolle bei der Tumorgenese [394] [151]. Die Auswirkungen durch das Auftreten von strukturellen Chromosomenveränderungen können direkt sein, z. B. über die Bildung von Fusionsgenen, die Änderung ihrer DNA-Kopienzahl oder indirekt über Änderungen der Chromatinorganisation. Dabei können beispielsweise TAD-Grenzen verschoben und regulatorische Sequenzen disloziert werden oder aktive Chromatinsegmente in inaktive Regionen und *vice versa* transferiert werden (Positionseffekte) [299] [240] [435] [153].

1.2.1. Molekulare Mechanismen zur Entstehung struktureller Chromosomenveränderungen

Chromosomale Instabilität kann unterschiedliche Ursachen haben. Das gehäufte Auftreten von DNA-Läsionen kann u. a. durch umweltbedingte DNA-schädigende Agenzien (z. B. ionisierende Strahlung) oder Fehlfunktionen physiologischer Rekombinationsprozesse (z. B. RAG-induzierte Doppelstrangbrüche während der V(D)J-Rekombination) verursacht werden [275]. Die resultierenden Strangbrüche sind nicht zufällig im Genom verteilt, sondern zeigen eine Assoziation mit bestimmten Sequenz- und Chromatineigenschaften chromosomaler Regionen [404]. Eine Anhäufung von Strangbrüchen innerhalb des Genoms gefährdet die genomische Integrität, was sich lethal auf die Zelle auswirken kann. Da jede Zelle täglich Zehntausende dieser DNA-Läsionen erleidet, dienen verschiedene zelleigene Mechanismen der Reparatur der DNA. Dabei unterscheiden sie sich u. a. in

ihren beteiligten Proteinkomplexen, der Abhängigkeit von DNA-Homologien und ihrem Wirkungszeitpunkt während des Zellzyklus [435].

Bei der Reparatur der DNA-Doppelstrangbrüche liegt, neben weiteren Alternativmechanismen (zusammengefasst in [64]), insbesondere die nichthomologe End-zu-End-Verknüpfung (*non-homologous end joining*; NHEJ) und homologe Rekombinationsreparatur (HRR) vor (zusammengefasst in [307]). Die HRR wird dabei über die Rekombination zwischen homologen Sequenzen realisiert und ist daher vermehrt in der späten S-Phase bis G₂-Phase des Zellzyklus aktiv, wenn die Schwesterchromatiden nächstgelegen sind [400]. Diese Reparaturmechanismen sind nicht immer perfekt. Eine fehlgeleitete Rekombination von benachbarten nicht-allelischen Segmenten mit hohen Sequenzsimilaritäten (nicht-allele homologe Rekombination; NAHR) kann zur Ausbildung von DNA-Gewinnen oder DNA-Verlusten führen (Abb. 1.4 A) [435] [65]. In verschiedenen Individuen ist dabei häufig das gleiche chromosomale Intervall betroffen. Solche rekurrenten Chromosomenveränderungen treten besonders bei *Genomic Disorders* auf [243]. Allgemein ist NAHR eher Ursache von Rearrangements der Genome der Keimbahnzellen und weniger von Genomen somatischer Zellen [412].

Ein rekombinationsunabhängiger Reparaturmechanismus ist NHEJ, welcher hauptsächlich in der G₁-Phase und frühen S-Phase des Zellzyklus vorliegt [400]. Im Rahmen des NHEJ werden die chromosomalen Bruchstellen mit Hilfe des Ku-Proteinkomplexes und anderer Proteine wieder ligiert. Dabei können z. B. DNA-Sequenzen verloren oder zusätzliche Segmente aufgenommen werden [228] (Abb. 1.4 A). NHEJ gilt auch als Hauptursache chromosomaler Translokationen [142] und ist insbesondere für die komplexen chromosomalen Veränderungen der Tumorgenome mitverantwortlich [391] [203]. Neben einfachen Änderungen der DNA-Kopienzahl kann es über mehrere Replikationszyklen hinweg zur Amplifikation genetischen Materials kommen, welches sich als *homogeneously staining regions* (HSRs) oder *double minutes* (DMs) manifestieren kann. Während HSRs intrachromosomal amplifizierte Segmente repräsentieren, bilden DMs autonom replizierende extrachromosomale DNA-Fragmente aus [392]. Die Neuorganisation kann auch verschiedene Chromosomen involvieren. So resultiert in vielen Tumorgenomen die Segmentierung mehrerer Chromosomen und die anschließende Religation in derivaten Chromosomen, welche bis hin zu mehreren hundert Mb großen Neochromosomen anwachsen können [133] (Abb. 1.4 B). Betrifft die Reorganisation das gesamte Genom und ist stabil spricht man von einem genomischem Chaos, was eine wichtige Rolle in der Evolution von Tumoren spielt [234] [454]. Infolge der komplexen Rearrangements können Kombinationen von strukturellen Chromosomenveränderungen auftreten, die der Zelle einen selektiven Vorteil inner-

1.2. Strukturelle Veränderungen von Chromosomen

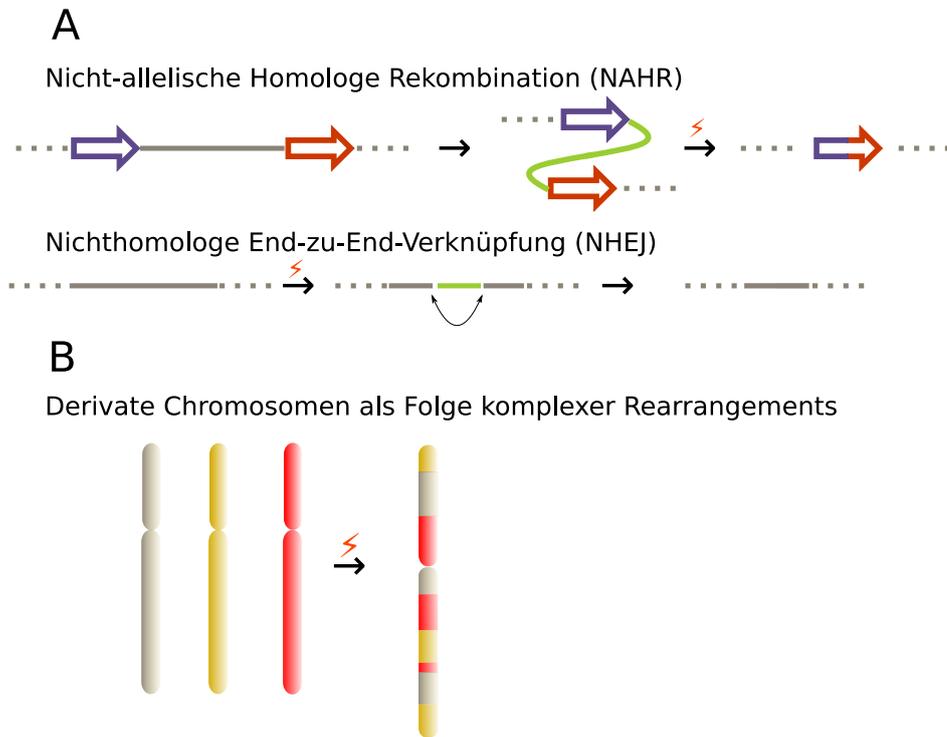


Abbildung 1.4.: Mechanismen zur Entstehung struktureller Chromosomenveränderungen.

A-B Molekulare Mechanismen chromosomaler Rearrangements **A** NAHR: Die nicht-allelele homologe Rekombination über Sequenzen mit hohen Sequenzsimilaritäten wie z. B. segmentalen Duplikationen führt zu CNVs, Inversionen und Translokationen. NHEJ: Die DNA-Reparatur wird über Proteinkomplexe ermöglicht, welche doppelsträngige DNA-Läsionen ohne die Voraussetzung langer Sequenzhomologien religiert. Bei mehreren DNA-Läsionen kann dies zu CNVs oder Translokationen führen. **B** Derivates Chromosom: Komplexe Rearrangements können z. B. unter Beteiligung mehrerer Chromosomen zu komplex rearrangierten Chromosomen führen. Farbige Pfeile: Segmente hoher Sequenzsimilarität; grüne Linie: von CNVs betroffenes Intervall; roter Blitz: Einführung von Doppelstrangbrüchen. Schema modifiziert nach [435] [64] [398].

halb der Zellpopulation bringen. Dabei wurde lange angenommen, dass Tumorgenome die chromosomalen Veränderungen graduell akkumulieren. Es konnten jedoch auch Prozesse identifiziert werden, welche sprunghafte Veränderungen des Genoms involvieren (z. B. Chromothripsis und Chromoplexy) [454] [411] [13].

1.2.2. Genomische Risikofaktoren struktureller Chromosomenveränderungen

Die Bruchpunkte chromosomaler Veränderungen zeigen oftmals eine regionale Häufung, die auf Regionen erhöhter Chromosomenvulnerabilität hinweisen. Diese genomischen Hotspots findet man sowohl bei vererbten als auch somatischen Veränderungen, selbst bei den evolutionären Chromosomenrearrangements sind die Bruchpunkte nicht zufällig verteilt. Chromosomale Regionen mit einer Anreicherung von Bruchpunkten weisen hierbei spezielle sequenz- und chromatinbasierte Eigenschaften auf. Welche Faktoren dabei eine Rolle spielen und wie diese sich gegenseitig beeinflussen ist noch nicht genau geklärt, jedoch konnten einige Eigenschaften von Rearrangement-anfälligen chromosomalen Regionen bestimmt werden [3] [18] [201] [348] [32] [439].

Sequenzeigenschaften wie ein hoher GC-Gehalt, non-B-DNA-Konformation wie Quadruplexstrukturen und repetitive Sequenzen wie z. B. *Alu*-Repeats erhöhen die Wahrscheinlichkeit eines Rearrangements. Weiterhin erhöhen die Sequenzsimilaritäten der SDs die Gefahr eines NAHR-Ereignisses [248] [432]. Neben den Sequenzeigenschaften spielt auch der Chromatinstatus eine Rolle. So ist die transkriptionelle Aktivität mit einem offenen, dekondensierten Chromatinstatus verbunden, was die Wahrscheinlichkeit eines Doppelstrangbruchs erhöht [46]. In diesem Zusammenhang wird zunehmend auch der Einfluss der dreidimensionalen Chromatinorganisation für das Auftreten rekurrenter Veränderungen diskutiert. Die räumliche Konzentration von DNA-Sequenzen unterschiedlicher Chromosomen, wie sie in *Transcription Factories* beobachtet werden können, könnte zum Beispiel Ursache für zelltypspezifische Translokationen sein. In Übereinstimmung mit dieser Annahme kann man rekurrente Translokationspartner in dem entsprechenden Zelltyp auch ohne Rearrangement häufig in einer räumlicher Nähe vorfinden [346] [310] [361] [304] [214] [293] [264] [41] [214] [72] [201] [128] [108] [310] [455] [297].

Genomische Risikofaktoren: Quadruplexstrukturen

Quadruplexstrukturen formieren sich aus dem G4-Sequenzmotiv:



Diese Motive finden sich gehäuft an Telomeren, Promotoren und Exon/Intron-Grenzen und haben das Potential, wenige Nukleotid-lange intra- und intermolekulare Schleifen zu formen. Dabei bilden sich über die beteiligten Guanin-Nukleotide Hoogsteen-Basenpaarungen aus, die zu einer zyklisch planaren Anordnung von vier Guanin-Nukleotiden innerhalb der Schleife führen. Mehrere dieser planaren Guanin-Anordnungen sind nacheinander in der Schleife gruppiert und stabilisieren diese. Durch die Ausbildung dieser Sekundärstruktur kann die Replikationsmaschinerie des Genoms beeinträchtigt werden und DNA-Läsionen zur Folge haben. Ihre häufige Lage in funktionell essentiellen Regionen gibt Quadruplex-Strukturen eine Rolle in z. B. der Transkription, Replikationsinitiierung und Chromosomenstabilität [248] [339].

Genomische Risikofaktoren: Segmentale Duplikationen

Viele der den *Genomic Disorders* zugrundeliegenden rekurrenten Chromosomenveränderungen entstehen durch die illegitime Rekombination der segmentalen Duplikationen (SDs, auch *Low Copy Repeats* genannt; LCRs). Dabei handelt es sich um duplizierte Sequenzblöcke von über 1 kb bis mehreren 100 kb, die Sequenzsimilaritäten von über 90 % aufweisen [175] [15] [19]. SDs decken ca. 5 % des humanen Genoms ab. Die Verteilung der SDs im Genom ist dabei nicht gleichmäßig, sondern zeigt eine Anreicherung von SDs in perizentromerischen und subtelomerischen Regionen. Jedoch können sie auch in interstitiellen chromosomalen Regionen akkumulieren und große SD-Cluster ausbilden [15] [16] [351]. Einen besonders hohen Anteil an solchen intrachromosomalen und interstitiellen SDs hat das Chromosom 7 [157] [356] [71]. Die der Entstehung und Verteilung der SD-Cluster zugrundeliegenden Mechanismen sind noch nicht vollständig aufgeklärt. Die Insertionen von SDs können über NHEJ oder der homologen Rekombination erfolgen [233] [102] [145] [160] [17]. Der Insertionsmechanismus scheint hierbei abhängig von der chromosomalen Region und dem evolutionären Alter der translozierenden SD zu sein. Ältere SDs in z. B. subtelomerischen Regionen sind vermutlich zu einem großen Teil über *Alu*-Sequenzen als homologe Substrate und NAHR generiert worden. Jüngere SDs hingegen formierten sich wahrscheinlich häufiger aufgrund von NHEJ-Ereignissen [197]. Man nimmt an, dass ei-

ne Insertion durch eine non-B DNA-Konformation der Zielsequenz begünstigt wird, z. B. durch *Alu*-Sequenzen oder G4-Motive. Hier kann eine SD als *core duplicon* inserieren, was die Instabilität der Region erhöht und so die Insertion weiterer SDs und deren nicht allelische Rekombination fördert. Schließlich kann es zur Bildung komplex organisierter Cluster bestehend aus mehrfach rearrangierten SD-Blöcken kommen (*duplication shadowing*) [70] [181] [254].

Die Segmente mit einem hohen Anteil an SDs sind Regionen hoher chromosomaler und evolutionärer Instabilität [288] [336] [351] [367] [414]. Aufgrund der hohen Sequenzsimilaritäten, auch zwischen nicht-allelischen SDs, bilden SDs Substrate für die homologe Rekombination und verringern somit die genomische Integrität. Deutlich wird dies durch die Anreicherung von SD-Clustern an evolutionären Bruchpunktregionen und das SD-vermittelte häufige Auftreten von CNVs in unabhängigen Individuen [10] [258] [4] [253] [385]. Obwohl SDs damit eine Gefahr für die Integrität des Genoms darstellen, kam es während der Hominidenevolution zu einer Anreicherung dieser Sequenzen im humanen Genom [256]. Dieser Umstand spricht dafür, dass SDs einen selektiven Vorteil vermitteln, der zu ihrer Fixierung im Genom geführt hat. Eine Möglichkeit ist, dass die Genduplikationen der SDs eine Quelle für neue Gene darstellen und auf diese Weise einen wichtigen Faktor der Primatenevolution darstellen. Über die Rekombination zwischen SDs können Sequenzen dabei unidirektional von einer SD zur anderen transferiert werden (zusammengefasst in [63]). Nachfolgend kann das duplizierte Gen durch adaptive Selektion eine neue bzw. modifizierte Funktion annehmen (*neofunctionalization* bzw. *subfunctionalization*; zusammengefasst in [247]), wobei die parentale SD jedoch die ursprüngliche Funktion beibehält [237]. Dieser Mechanismus hat im Laufe der humanen Evolution verschiedene Genfamilien, wie z. B. die Chemokinrezeptorgene [420] oder die Hämoglobingene [375], hervorgebracht. Die Häufung von SDs an der Grenze zwischen eu- und heterochromatischen Regionen, insbesondere in den perizentromerischen Bereichen, könnte auf eine weitere Funktion der SDs hinweisen, die für deren Fixierung im Genom verantwortlich sein könnte. So wird diskutiert, ob die perizentromerischen SDs die inhibierenden regulatorischen Einflüsse des Heterochromatins auf die Transkription der euchromatischen Gene blockieren [102] [136].

SDs spielen eine wichtige Rolle in der Ausbildung evolutionärer und pathogener Rearrangements. Daher kann die Aufklärung der SD-Verteilung und deren mögliche Beteiligung an der Genomarchitektur auch Aufschluss über das Auftreten von chromosomalen Veränderungen, evolutionären Mechanismen und der Entstehung von Erkrankungen geben.

1.2.3. Evolutionäre Chromosomenveränderungen am Beispiel des Chromosom 7 während der Primatenevolution

Bei strukturellen Chromosomenveränderungen, welche sich innerhalb einer Spezies etabliert haben, wird davon ausgegangen, dass essentielle Genkombinationen innerhalb der Rearrangements beibehalten und nur Neuorganisationen mit selektivem Vorteil (oder mindestens ohne Nachteil) in der Population fixiert werden [220]. Inwieweit chromosomale Rearrangements dabei zur Speziation beitragen ist jedoch unbekannt [193].

Strukturelle Chromosomenveränderungen haben einen großen Anteil an der genomischen Unterschiedlichkeit zwischen den Primatenspezies [236]. Unter anderem wurde das Chromosom 7 im gemeinsamen Vorfahr von Mensch und afrikanischem Affen durch eine perizentrische Inversion verändert (Abb. 1.5), gefolgt von einer parazentrischen Inversion vor der Aufspaltung in Mensch und Schimpanse. Die genauen Auswirkungen dieser Rearrangements und die zugrundeliegenden Mechanismen dahinter sind noch ungeklärt. Jedoch spricht einiges dafür, dass die spezielle Genomarchitektur, die sich während der Homini-denevolution etabliert hat, einen großen Anteil an der Ausbildung der Rearrangements hat. Insbesondere könnte die beachtliche Akkumulation von SDs während der Homini-denevolution die Ausbildung struktureller Chromosomenveränderungen über NAHR begünstigt haben [255]. Dafür spricht auch die häufige Assoziation der evolutionären Synteniebruchpunkte mit SDs [351] [356] [9] [456] [193]. Allerdings spricht die zeitliche Abfolge von SD-Akkumulation und dem Auftreten der evolutionären Rearrangements in einigen Fällen dafür, dass SD-Cluster nicht Ursache, sondern Konsequenz des chromosomalen Rearrangements waren [18] [355] [136].

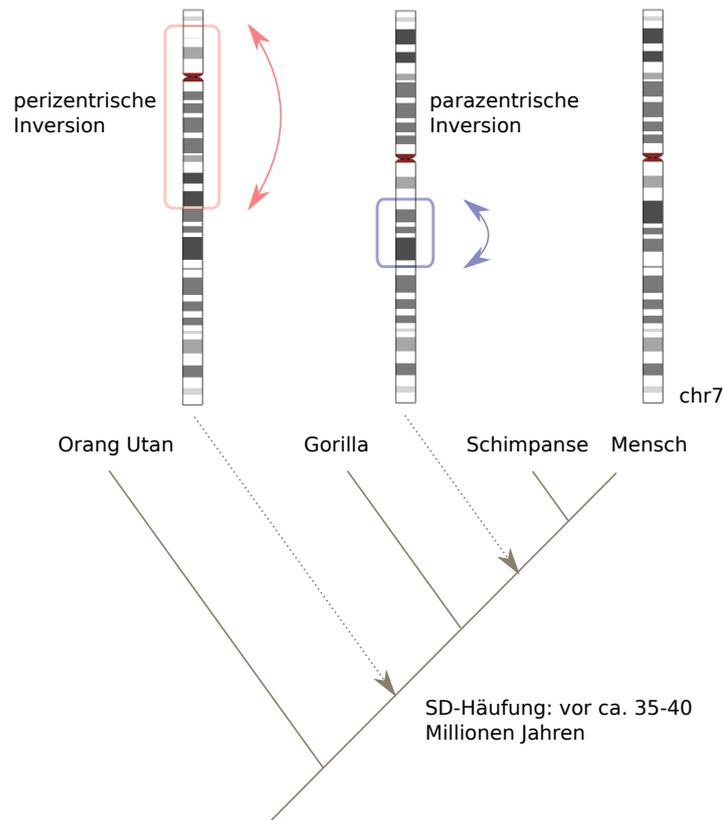


Abbildung 1.5.: Inversionen des Chromosom 7 während der Primatenevolution.

Eine perizentrische Inversion des gemeinsamen Vorfahren der afrikanischen Affen, gefolgt von einer parazentrischen Inversion vor der Aufspaltung in Mensch und Schimpanse, führten zu der heutigen Organisation des humanen Chromosom 7. Möglicherweise ist die Ausbildung von strukturellen Chromosomenveränderungen im Genom der Hominiden durch eine Anhäufung von SDs vor ca. 35 bis 40 Millionen Jahren begünstigt worden. Die hypothetische Reversion der Idiogramme des Chromosom 7 berücksichtigt ausschließlich die peri- und parazentrischen Inversionen.

1.2.4. Strukturelle Chromosomenveränderungen als Ursache von *Genomic Disorders* am Beispiel der WBS-kritischen Region

Erkrankungen, die aufgrund von rekurrenten Rearrangements des Genoms erfolgen werden als *Genomic Disorders* bezeichnet [241] [242]. Häufige Ursache von *Genomic Disorders* ist die SD-vermittelte NAHR, in deren Verlauf das von den SDs flankierte Intervall entweder verloren oder dupliziert wird [384]. Charakteristisch ist dabei, dass die Verluste bzw. Gewinne des gleichen chromosomalen Materials oft in unabhängigen Patienten auftreten [241] [242]. Ein typisches Beispiel der *Genomic Disorders* ist das Williams-Beuren-Syndrom (WBS; OMIM #194050), eine Entwicklungsstörung zu deren klinischem Erscheinungsbild u. a. kardiovaskuläre Erkrankungen und variierende intellektuelle Behinderungen gehören (zusammengefasst in [321] [270]). Die Prävalenz des WBS beträgt 1:7.500 [395], wobei WBS zumeist *de novo* auftritt [29].

In 89 % der Fälle wird WBS durch eine rekurrente ca. 1,5 Mb große hemizygote Deletion innerhalb der 7q11.23 Chromosomenbande verursacht [113] [268] [316] [271] [315] [158] [29] [417]. Diese sogenannte WBS-kritische Region (*WBS critical region*, WBSCR) umfasst ca. 20 bis 30 Gene [271] (Abb. 1.6 A). Ursache für die Rekurrenz dieser Deletionsregion sind drei SD-Cluster, die innerhalb der WBSCR lokalisiert sind und zumeist über NAHR die DNA-Kopienzahländerung bedingen [316] [344] (Abb. 1.6 B).

Die SD-Cluster setzen sich jeweils aus den modularen SD-Blöcken A, B und C zusammen, die in den verschiedenen Clustern unterschiedlich arrangiert sind [417] [315] [29]. Aufgrund der hohen Sequenzähnlichkeit kann es während der Zellteilung zu einer Rekombination (intra- und interchromosomal) zwischen den SD-Blöcken kommen [98]. Durch NAHR wird zusätzlich zu der Deletion eine reziproke Duplikation der WBSCR generiert, welche das 7q11.23-Duplikationssyndrom verursacht (OMIM #609757, [381]). Die Prävalenz liegt hier mit 1:13.000-20.000 jedoch niedriger als beim WBS [418]. Distal der WBSCR lokalisiert eine Region, die mit der \sim 1,2 Mb großen rekurrenten Deletion des distalen 7q11.23-Deletionssyndroms (OMIM #613729, [332]) assoziiert ist. Auch die Region des distalen 7q11.23-Deletionssyndroms wird von segmentalen Duplikationen flankiert (Abb. 1.6 A).

Aufgrund der Häufung von Polymorphismen und der Prädisposition für genetische Erkrankungen geht man von einer hohen genomischen Dynamik für die 7q11.23-Region aus [305] [29] [6]. Ob es zusätzlich zu den SDs noch weitere Faktoren gibt ist jedoch noch unbekannt.

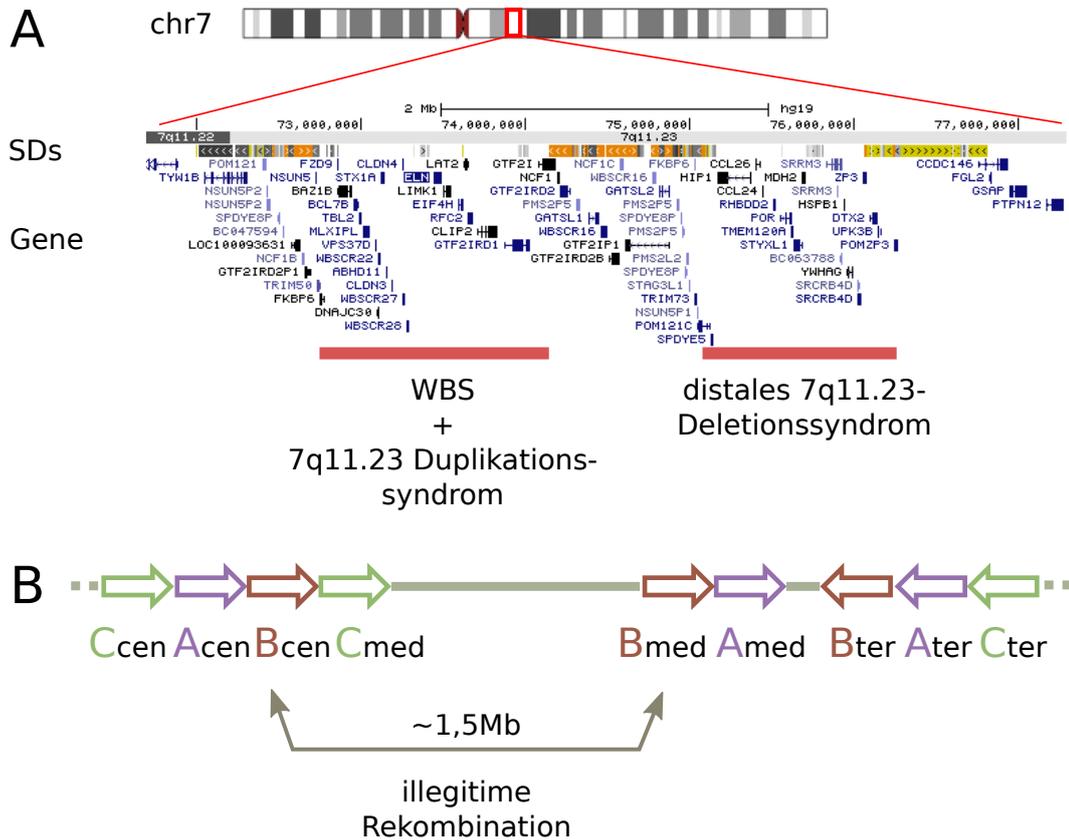


Abbildung 1.6.: Die *Genomic Disorders* der 7q11.23-Region.

A UCSC Genome Browser Ansicht der 7q11.23-Region in hg19. Innerhalb der 7q11.23-Region sind die genomischen Intervalle des WBS, 7q11.23-Duplikationssyndroms und des distalen 7q11.23-Deletionssyndroms lokalisiert. Flankiert werden diese von SD-Blöcken, welche über NAHR die strukturellen Veränderungen generieren. **B** Schema der Orientierung und Anordnung der WBS-flankierenden Segmentalen Duplikationen. Die flankierenden SDs sind modular aus den Blöcken A, B und C aufgebaut, welche sich in Anordnung und Orientierung unterscheiden können. Die häufigste Deletion des WBS von ca. 1,5 Mb wird durch NAHR des zentromerischen B-Blockes mit dem medialen B-Block verursacht. B modifiziert nach [6] und [271].

1.2.5. Strukturelle Chromosomenveränderungen in Tumoren am Beispiel der kutanen T-Zell-Lymphome

Die Anhäufung chromosomaler Veränderungen ist ein typisches Merkmal der Tumorentwicklung [151]. Die Komplexität der Veränderungen in Tumorgenomen variiert dabei stark in Abhängigkeit der Tumorentität [390]. Dabei korreliert zumeist ein komplexer Karyotyp mit einer negativen Prognose. Bei vielen Tumorarten fällt auf, dass die Akkumulation solcher Veränderungen innerhalb des Genoms nicht zufällig passiert. Stattdessen beobachtet man häufig bevorzugte Kombinationen chromosomaler Veränderungen, die zu einem tumorspezifischen Muster führen [150] [267] [337]. Obwohl diese Muster chromosomaler Veränderungen mittlerweile diagnostische und therapeutische Bedeutung haben, ist deren Ursache nur unzureichend verstanden.

Verschiedene Mechanismen könnten zur Ausprägung dieser tumorspezifischen Muster führen. Zum einen können die Aberrationsmuster der Tumorgenome Folge der Unterschiede in der Chromosomenvulnerabilität verschiedener Regionen sein, die von Sequenz- bzw. Chromatineigenschaften beeinflusst wird. Zum anderen besteht die Möglichkeit einer zunächst stochastischen Verteilung von Chromosomenbrüchen im Genom, welche erst durch den selektiven Druck tumorspezifische Muster ausbilden. So stehen Regionen mit tumorfördernden Onkogenen, tumorreprimierenden Tumorsuppressorgenen und Fusionsgenen aberranter Funktion unter ganz besonders starkem selektivem Druck und können, je nach Effekt auf die Tumorzelle, beibehalten (positive Selektion) oder verworfen (negative Selektion) werden [223]. Diese können dann als rekurrente chromosomale Veränderungen identifiziert werden [280] [198] [378] [394].

Ein Beispiel für Tumoren mit spezifischen Mustern komplexer Chromosomenveränderungen sind die kutanen T-Zell-Lymphome (*cutaneous T cell lymphoma*; **CTCL**). Diese sind eine heterogene Gruppe von hämatologischen Neoplasien, welche durch die Akkumulation neoplastischer haut-residenter T-Zellen verursacht werden [345]. Anders als primär nodale Lymphome, treten die Tumorzellen zunächst hauptsächlich in der Haut der Patienten auf und breiten sich erst sekundär auf extrakutane Organe aus [244] [441] [169]. Die häufigsten Varianten des CTCL sind Mycosis fungoides (**Mf**) und das Sézary Syndrom (**SeS**). Beide Tumorentitäten weisen ähnliche Charakteristika auf wie z. B. das Auftreten neoplastischer Zellen mit hyperchromatischen und atypisch geformten (cerebriformen) Nuklei. Jedoch unterscheiden sie sich in ihrem klinischen Bild und der Prognose für die Patienten [244]

[441]. Zum Beispiel ist die indolente CTCL-Variante Mf mit einer im Allgemeinen guten Prognose assoziiert. Jedoch können sich aus den anfänglichen Symptomen von trockenen und schuppigen Rötungen in der Haut (Patches), über das Stadium der Plaques, kutane Tumore innerhalb der Läsionen entwickeln. Die neoplastischen Zellen des Mf sind hauptsächlich epidermotrop, d. h. innerhalb der Epidermis lokalisiert. Im Gegensatz zu MF zeigt das SeS eine leukämische Ausschwemmung der neoplastischen T-Zellen. Dadurch kommt es zu vermehrtem Auftreten von malignen Zellen im Blut der Patienten, welche sekundär Lymphknoten und viszerale Organe befallen können. Außerdem kommt es zu einem starken Juckreiz und einer Rötung nahezu der gesamten Körperoberfläche (Erythrodermie) [440]. Patienten mit der Diagnose SeS haben eine schlechtere Prognose als MF-Patienten und weisen eine mediane 5-Jahres-Überlebensrate von 24 % auf [441]. Das SeS ist mit einer Inzidenz von 0,01 pro 100.000 Personen jedoch weniger häufig als Mf [173]. Die neoplastischen Lymphozyten des CTCL stammen von ausgereiften T-Zellen, die normalerweise in der Haut zirkulieren und der Immunabwehr dienen. Die Aktivierung dieses T-Zell-Subtyps läuft als Folge pathogener Stimulierung ab. Pathogene Antigene in der Haut aktivieren in der Epidermis lokalisierte dendritische Zellen (u. a. Langerhans-Zellen), welche die Antigene phagozytieren, in Lymphknoten einwandern und über den *Major-Histocompatibility-Complex*-(MHC)-Klasse-II-Komplex Antigenpeptide den naiven $CD4^+$ T-Zellen (T_N) präsentieren. Diese werden über die resultierende TCR-Aktivierung zur Proliferation angeregt und in T-Effektorzellen differenziert. Anschließend verlassen die ausdifferenzierten T-Effektorzellen die Lymphknoten und wandern über dermale postkapilläre Venolen in die Hautschichten ein, um dort antipathogen zu wirken (Abb. 1.7). Bei CTCL proliferieren monoklonale T-Zellen dieses Typs und akkumulieren als Mikroabzesse in der Haut bzw. zirkulieren im Blut der Patienten [12] [452]. Wie bei vielen anderen Neoplasien ist auch in CTCL-Genomen die Anhäufung von chromosomalen Aberrationen mit einer schlechten Prognose assoziiert, wobei spätere CTCL-Stadien mehr Aberrationen aufweisen als frühe Stadien [407] [27] [120]. Alle Chromosomen können dabei von chromosomalen Veränderungen betroffen sein. Gehäuft finden sich aber Verluste von Regionen der Chromosomen 1p, 10q und 17p und ein DNA-Zugewinn von 8q, 10q und 17q [423] [231] [196].

In der Literatur werden eine Vielzahl von Genen und microRNAs [22] [331] [257] [352] [121] diskutiert, welche für die Tumorphathogenese verantwortlich sein können [369]. Häufig sind diese Teil gemeinsamer Signalwege. So wurden Kandidatengene ermittelt, welche die Apoptose regulieren oder eine Apoptoseresistenz vermitteln (z. B. *FAS* [446], *PTEN* [82]), die TCR-Aktivierung und Proliferation der T-Zellen beeinflussen (*CD28* [74], *PLCG1*

1.2. Strukturelle Veränderungen von Chromosomen

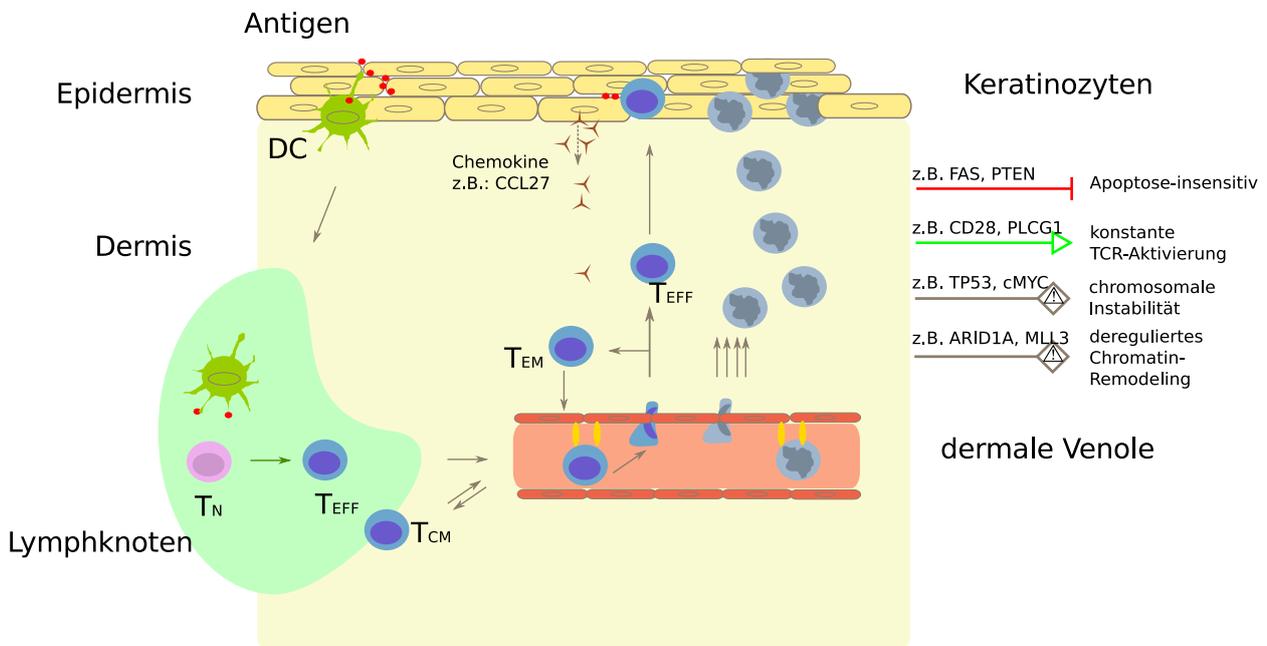


Abbildung 1.7.: T-Zell-Migration neoplastischer T-Zellen in die Epidermis.

Schema der T-Zell-Migration in die Epidermis nach Antigenstimulierung. Antigene aktivieren dendritische Zellen. Diese wandern in die Lymphknoten wo sie naiven T-Zellen Antigenpeptide präsentieren. Über den JAK/STAT-Signalweg wird die Proliferation und Differenzierung der T-Zellen in Effektorzellen (z. B. T_H1 und T_H2) eingeleitet. Aktivierte T-Zellen können über Adhäsionsproteine auf der T-Zelloberfläche und Endothelzellen (Rezeptoren: in Abb. als gelbe Stäbchen) die dermalen Venolen überwinden und entlang von Chemokingradienten in die Epidermis migrieren, wo sie die Immunantwort unterstützen. Der Großteil der Effektorzellen geht anschließend in die Apoptose. Einige verbleiben als Gedächtniszellen vornehmlich in der Haut (T_{EM}) bzw. im lymphatischen System (T_{CM}). In CTCL erfolgt nach der T-Zell-Rezeptoraktivierung eine Proliferation von monoklonalen T-Zellen die vermehrt in die Haut einwandern und dort akkumulieren bzw. im lymphatischen System zirkulieren. Molekulare Ursachen hierfür sind die aberrante Expression tumorrelevanter Gene, die u. a. zu einer konstitutiven *T-Cell-Rearrangement*-(**TCR**)-Aktivierung und einer akquirierten Apoptoseresistenz führen.

[419]), die DNA-Schadensantwort und damit auch die chromosomale Stabilität beeinflussen (*TP53* [216], *MYC* [423]) oder für das Chromatinremodelling verantwortlich sind (*ARID1A* [196] [74], *MLL3* [416]).

Ein Modell für das SeS ist die Zelllinie Seax, welche aus malignen T-Zellen des peripheren Blutes einer 66 Jahre alten SeS-Patientin etabliert wurde [183]. Zellen die mehrere Monate in Kultur gehalten wurden wiesen einen annähernd tetraploiden Karyotyp, jedoch keine sézarytypische cerebriforme Struktur des Nukleus auf [183]. Auf molekularer Ebene teilt die Zelllinie jedoch viele Gemeinsamkeiten mit primären Sézary-Zellen. Wie in den malignen Zellen der Patienten zeigt auch die Seax-Zelllinie Apoptoseresistenz, deren molekularbiologische Ursachen u. a. die konstitutive Aktivität der Transkriptionsregulatoren NF- κ B [382], Notch1 [122], Transkriptionsfaktoren des JAK/STAT-Signalweges [325] und Deletionen von z. B. *TNFAIP3* [382] sind. Durch die Untersuchung dieser Zelllinie ist es daher möglich, einen Einblick in das Genom der Sézary-Tumorzellen zu bekommen.

1.3. Zielstellung

Strukturelle Chromosomenveränderungen konnten in bisherigen Studien wenig und nur mit geringer Auflösung im Zusammenhang mit der 3D-Organisation des Interphasekerns betrachtet werden. Folglich ist kaum bekannt, inwieweit das Auftreten von chromosomalen Aberrationen durch die Zellkernarchitektur beeinflusst wird und *vice versa*. Diese Arbeit hat daher zum Ziel, strukturelle Chromosomenveränderungen im Kontext der dreidimensionalen Chromatinorganisation zu analysieren, um so den Zusammenhang zwischen Zellkernarchitektur und chromosomalen Veränderungen besser beurteilen zu können. Am Beispiel des humanen Chromosom 7 soll über die Integration eigener Experimentdaten zur Chromatinorganisation und öffentlicher Daten zu intrachromosomalen Interaktionen, die Verteilung der SDs und evolutionärer Rearrangements in Bezug zur Zellkernarchitektur gesetzt werden. Besonderer Schwerpunkt soll dabei auf SD-Cluster der 7q11.23-Region gelegt werden, welche durch illegitime Rekombination, wiederkehrend in unabhängigen Individuen, zu DNA-Verlusten und in weiterer Folge zum Williams-Beuren-Syndrom führen können.

Chromosomale Veränderungen sind auch ein charakteristisches Merkmal der Tumorentwicklung. Die Auswirkungen der komplexen Chromosomenveränderungen auf die räum-

1.3. Zielstellung

liche Organisation des Zellkerns sind aufgrund bisheriger technischer Limitationen allerdings kaum untersucht. In dieser Arbeit soll deshalb in einer Zelllinie eines kutanen T-Zell-Lymphoms mit einer auf Hochdurchsatzsequenzierung basierenden Methode, die dreidimensionale Chromatinorganisation des Tumorgenoms analysiert werden. Dabei soll überprüft werden, in welchem Ausmaß sich die chromosomalen Aberrationsmuster in der Zellkernarchitektur widerspiegeln, und ob die chromosomalen Interaktionsmuster Rückschlüsse auf das Vorliegen von Chromosomenveränderungen erlauben. Zudem soll untersucht werden, inwieweit Chromatineigenschaften und Gengehalt chromosomaler Veränderungen Anhaltspunkte geben können, ob die tumorspezifischen Veränderungsmuster mehr Folge des selektiven Druckes oder regional erhöhter Chromosomenvulnerabilität sind.

2. Material und Methoden

2.1. Material

2.1.1. Biologisches Material

Zelllinien

Zelllinie	Hersteller/Referenz
H9	[323]
HH	[387]
Hey	[44]
Hut78	[134]
IMR90	[295]
IMR91L	[294]
Jurkat	[358]
KE-37	[94]
Mewo	ATCC American Type Culture Collection, USA
Molt-14	[95]
Myla	[184]
Seax	[183]
SH-SY5Y	[34]
SK-BR-3	ATCC American Type Culture Collection, USA

Patienten mit Sézary Syndrom

Die hier untersuchten 20 Patienten mit Sézary Syndrom, wurden entsprechend der Richtlinien des *World Health Organization-European Organization for Research and Treatment of Cancer classification for cutaneous lymphomas* [441] diagnostiziert und charakterisiert

2.1. Material

(siehe Patientendaten, S. [33](#)). Die Verwendung des Tumormaterials wurde von der Ethikkommission der Charité Berlin genehmigt und entsprechend der Deklaration von Helsinki gehandhabt.

Patientendaten: Tab. erstellt von Dr. M. Möbs und PD Dr. C. Assaf und entnommen aus [389]

Patient Nr.	Geschlecht	identischer TCR-R in Blut und Haut*	Zeitpunkt der Blutabnahme (MM/JJ)	Alter	Parameter des peripheren Blutes					Behandlungen vor Blutentnahme
					WBC (Anzahl/ml)	Sézary Zellen (in %)	CD4 ⁺ /CD3 ⁺ (in %)	CD4 ⁺ /CD8 ⁺ Ver-hältnis	Vβ ⁺ /CD4 ⁺ (in %)	
1	W	+	03/99	76	6.4	0	92	21	89	UVB, Pr
2	W	+	12/02	67	20.9	80	99	99	n.d.	Ch+P,IFN
3	W	n. d.	08/99	95	20.9	64	>99	49	>99	-
4	M	+	01/05 ^a	73	22.2	60	>99	665	99	UV,ECP
			02/05	73	22.7	76	96	712	n.d.	UV,ECP
5	W	n. d.	02/00	58	28.7	59	>99	94	98	-
6	M	+	10/09	59	31.1	33	99	136	n.d.	IFN,ECP
7	M	+	01/07 ^a	59	12.6	31	>99	63	97	IFN+ECP,Bex+SAHA, Ch+Pr
			03/07	59	23.0	33	99	321	n.d.	IFN+ECP,Bex+SAHA,Ch+Pr,αCD52
8	M	+	10/09	89	7.5	0	91	n.d.	74	PUVA
9	M	+	08/02	67	17.2	26	91	98	81	CyA,PUVA,IFN,ECP
10	M	+	02/03	74	13.2	25	97	29	93	PUVA,IFN
11	W	n.d.	08/00	86	6.9	21	97	24	96	PUVA Cyc+Pr,Cyc+Pr+V,Ch+Pr
			11/00 ^a	86	4.5	n.d.	94	n.d.	n.d.	PUVA Cyc+Pr,Cyc+Pr+V, Ch+Pr
12	M	n.d.	04/02	52	8.0	16	>99	n.d.	n.d.	PUVA,ECP,IFN,Dx, MTX
13	W	+	08/02	79	17.2	28	94	49	68	PUVA,Ch+Pr
14	W	n.d.	08/99	65	21.3	72	99	90	n.d.	ECP
15	W	+	08/07 ^a	64	21.3	n.d.	95	n.d.	n.d.	IFN,ECP,Ch+P,PUVA,MTX,Cyc,Bex+SAHA
			03/08	64	30.1	28	n.d.	n.d.	n.d.	IFN,ECP,Ch+P,PUVA,MTX Cyc,Bex+SAHA
16	W	+	08/08	79	9.2	8	92	20	84	ECP,IFN
			01/09 ^a	79	7.6	n.d.	91	15	n.d.	ECP,IFN
17	W	+	09/08	72	8.1	25	99	197	94	ECP,PUVA,IFN
18	M	+	03/97	62	43.0	44	n.d.	n.d.	n.d.	PUVA
19	M	+	01/10	49	11.9	n.d.	97	48	96	-
20	W	+	03/10	62	11.8	n.d.	96	31.9	89	ECP

TCR-R, T Zell Rezeptor Rearrangement; WBC, *white blood cell* (Leukozyten); Ac, Acitretin; Bex, Bexaroten; Ch, Chlorambucil; CyA, Cyclosporin A; Cyc, Cyclophosphamid, Dx, Doxorubicin; ECP, *extracorporeal photopheresis* (extrakorporale Photopherese); Flu, Fludarabin; MTX, Methotrexat; Pr, Prednison; PUVA, Psoralen + Bestrahlung mit langwelligem UV-Licht (UV-A); SAHA, *suberoylanilide hydroxamic acid* (suberoylanilide hydroxamische Säure); UV, Bestrahlung mit UV-Licht (nicht-spezifizierter Wellenlänge); UVB, Bestrahlung mit kurzwelligem UV-Licht (UV-B); V, Vincristin; αCD52, Alemtuzumab (humanisierter monoklonaler Antikörper gegen CD52); n.d., nicht determiniert; *, bestimmt über TCR-Rearrangement und nachfolgender Fluoreszenz-DNA-Fragment-Analyse, ^a Proben der FISH-Analyse

2.1. Material

Antikörper und Enzyme

Antikörper/ Enzym	Hersteller
β -Agarase	New England Biolabs, Ipswich, USA
Alu I (10 U/ μ l)	Promega, Madison, USA
<i>Hind</i> III	New England Biolabs, Ipswich, USA
α -H3K4me3, pAb-003-050 (polyklonal)	Diagenode, Liège, BEL
α -H4K8ac, pAb-103-050 (polyklonal)	Diagenode, Liège, BEL
α -m ⁵ C, BI-MECY-0100 (monoklonal)	Eurogentec, Köln, D
α -H3K4me3, ab8580 (polyklonal)	Abcam, Cambridge, UK
α -H3K9me2, ab1220 (monoklonal)	Abcam, Cambridge, UK
α -H4ac, 06-598 (polyklonal)	Upstate, Temecula, USA
Proteinase K	Thermo Fisher Scientific, Waltham, USA
RNase A	Thermo Fisher Scientific, Waltham, USA
Rsa I (10 U/ μ l)	Promega, Madison, USA

2.1.2. Chemikalien und Medien

Produkt	Hersteller
7-AAD	BD Biosciences, San Jose, USA
Agarose	Invitrogen, Groningen, NL
Ampullenwasser	Baxter, Santa Clara, USA
Annexin V-APC	BD Biosciences, San Jose, USA
BSA, acetyliert (10 mg/ μ l)	Promega, Madison, USA
Buffer C	Promega, Madison, USA
Casytron	Schärfe System GMBH Reutlingen, D
Chloroform	Merck Millipore, Darmstadt, D
Cot-DNA, human (1 μ g/ μ l)	Roche Applied Science, Penzberg, D
DMSO	Sigma-Aldrich, St. Louis, USA
dNTPs	QIAGEN, Hilden, D
DPBS	Lonza, Basel, CH
EDTA	Merck Millipore, Darmstadt, D
EGTA	Merck Millipore, Darmstadt, D
EMEM	Lonza, Basel, CH
Ethanol	Merck Millipore, Darmstadt, D

Fortsetzung auf der nächsten Seite

Produkt	Hersteller
Ethidiumbromid	Sigma-Aldrich, St. Louis, USA
FCS	Sigma-Aldrich, St. Louis, USA
Formaldehyd, 16 %, methanolfrei	Thermo Fisher Scientific, Waltham, USA
Gene Expression Wash Buffer I	Agilent, Santa Clara, USA
Gene Expression Wash Buffer II	Agilent, Santa Clara, USA
Glycin	Merck Millipore, Darmstadt, D
HEPES	Sigma-Aldrich, St. Louis, USA
Interleukin-2	Sigma-Aldrich, St. Louis, USA
Isopropanol	Merck Millipore, Darmstadt, D
Low-Melt-Agarose	Biozym Scientific, Hessisch Oldendorf, D
Natriumbutyrat	Sigma-Aldrich, St. Louis, USA
Natriumchlorid	Roth, Karlsruhe, D
Natriumpyruvat	Lonza, Basel, CH
Natriumperchlorat	AppliChem GmbH, Darmstadt, D
Oligo aCGH/Chip-on-ChIP Wash Buffer I	Agilent, Santa Clara, USA
Oligo aCGH/Chip-on-ChIP Wash Buffer II	Agilent, Santa Clara, USA
Oligo Gene Expression Wash Buffer I	Agilent, Santa Clara, USA
Oligo Gene Expression Wash Buffer II	Agilent, Santa Clara, USA
Penicillin/Streptomycin	Lonza, Basel, CH
Phenol:Chloroform:Isoamylalkohol (25:24:1)	Sigma-Aldrich, St. Louis, USA
Protease-Inhibitoren	Sigma-Aldrich, St. Louis, USA
RPMI-1640	Lonza, Basel, CH
SDS 10 %	Sigma-Aldrich, St. Louis, USA
Staurosporin	Cell Signaling Technology, Danvers, USA
T4 DNA-Ligase	New England Biolabs, Ipswich, USA
T4 DNA-Polymerase	New England Biolabs, Ipswich, USA
T4 Ligationspuffer	New England Biolabs, Ipswich, USA
TE Puffer (pH 8,0), 1x	Life Technologies, Carlsbad, USA
Triton-X-100	Sigma-Aldrich, St. Louis, USA
Trypsin-EDTA	Lonza, Basel, CH
TRIzol	Life Technologies, Carlsbad, USA
UltraGlutamine I	Lonza, Basel, CH

2.1. Material

2.1.3. Puffer und Lösungen

Puffer	Zusammensetzung
Annexin-Bindungspuffer	10 mM HEPES, 140 mM NaCl, 5 mM CaCl ₂
DNA-Lysepuffer	0,4 M Tris-HCl pH 8.0, 0,06 M Na-EDTA, 0,15 M NaCl, 1 % SDS
Hi-C Ligationsmix	745 µl 10 % Triton X-100, 745 µl 10× T4 Ligation Buffer, 8 µl 10 mg/ml BSA, 6,04 ml Bidest
2× Bindungs- und Waschpuffer	10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl

2.1.4. Kits (Reagentsysteme)

Produkt	Hersteller
BioPrime Total Genomic Labeling Module	Invitrogen, Groningen, NL
BioPrime Purification Module	Invitrogen, Groningen, NL
DNA Enzymatic Labeling Enzyme Module	Agilent, Santa Clara, USA
Oligo aCGH/ChIP-on-chip Hybridization Kit	Agilent, Santa Clara, USA
Cyanine dUTP Dye Pack	Agilent, Santa Clara, USA
ERCC RNA Spike-In Control Mix	Thermo Fisher Scientific, Waltham, USA
Genomic DNA Screentape	Agilent Technologies, Santa Clara, USA
Transcription Factor ChIP Kit	Diagenode, Liège, BEL
IPure Kit	Diagenode, Liège, BEL
Rneasy Mini Kit	QIAGEN, Hilden, D
RiboMinus Eukaryote Kit for RNA-Seq	Life Technologies, Carlsbad, USA
SOLiD Total RNA-Seq Kit	Life Technologies, Carlsbad, USA
SOLiD 6-Lane FlowChip Kit	Life Technologies, Carlsbad, USA
SOLiD RNA Barcoding Kit	Life Technologies, Carlsbad, USA
One-Color RNA Spike-In Kit	Agilent, Santa Clara, USA
QIAamp DNA Purification Kit	QIAGEN, Hilden, D
Quick Amp Labeling Kit, one-color	Agilent, Santa Clara, USA
GenomePlex Whole Genome Amplification (WGA) Kit	Sigma-Aldrich, St. Louis, USA
Low RNA Input Linear Amplification Kit	Agilent, Santa Clara, USA

2.1.5. Mikroarrays

Mikroarray	GEO-Nummer
021850 SurePrint G3 Human CGH Microarray 2x400K	GPL9777
014850 Whole Human Genome Microarray 4x44K	GPL6480
023939 SurePrint Custom Human Chr19p CGH Microarray 8x60K	GPL10304
037305 SurePrint G3 Human Custom CGH Microarray 4x180K	GPL17964
36k <i>whole human genome tiling path</i> BAC Array	GPL5114

Alle verwendeten Mikroarrays (bis auf 36k *whole human genome tiling path bacterial artificial chromosome (BAC) Array*) wurden vom folgenden Hersteller bezogen: Agilent, Santa Clara, USA. Der 36k *whole human genome tiling path* BAC Array wurde in der Arbeitsgruppe AG Molekulare Zytogenetik unter Leitung von PD Dr. Reinhard Ullmann (MPIMG) in Kooperation mit der Arbeitsgruppe Kreuzberger (MPIMG) erstellt unter Verwendung der BAC-Klone aus [117, 209, 306].

2.1.6. Geräte

Gerät	Name	Hersteller
Filterabzug	captair bio	erlab, Köln, D
DNA Qualitätskontrolle	2200 TapeStation	Agilent, Santa Clara, USA
Elektrophoresekammer	Horizon 11.14	Life Technologies, Carlsbad, USA
PCR-Maschine	Primus HT	MWG AG, Ebersberg, D
Tischzentrifuge	Centrifuge 5415D	Eppendorf, Hamburg, D
UV/Vis Spektrophotometer	NanoDrop ND-1000	Peqlab, Erlangen, D
Pipetten	Pipetman P	Gilson, Middleton, USA
Vortexer	VF2	Janke+Kunkel, Staufen, D
Vakuumpzentrifuge	Concentrator 5301	Eppendorf, Hamburg, D
FACS	FACSAria III	BD Biosciences, San Jose, USA
Thermostat	Thermomixer comfort	Eppendorf, Hamburg, D
Sequenzierer	SOLiD 5500xl Genetic Analyzer	Life Technologies, Carlsbad, USA

Fortsetzung auf der nächsten Seite

2.1. Material

Gerät	Name	Hersteller
Hybridisierungssofen	G2545A	Agilent, Santa Clara, USA
Luftreiniger	Multigas GC	IQ Air, La Mirada, USA
Magnetrührer	IKAMAG REO	Janke+Kunkel, Staufen, D
Magnetrührer/ Heizelement	IKAMAG REC-G	Janke+Kunkel, Staufen, D
DNA Mikroarray Scanner	G2565BA	Agilent, Santa Clara, USA
Schüttler	KL-2	Edmund Bühler, Hechingen, D
Mikrozentrifuge	Centrifuge 5810R	Eppendorf, Hamburg, D
Rotator	Reax 2	Heidolph, Schwabach, D
Bioruptor	UCD-200	Diagenode, Liège, BEL
Pipettierhilfe	Pipetus akku	Hirschmann Laborgeräte, Eberstadt, D
Mikrozentrifuge	Centrifuge 5417R	Eppendorf, Hamburg, D
Minizentrifuge	Rotilabo	Roth, Karlsruhe, D
Pipettierhilfe	Pipetteboy acu	IBS Integra Biosciences, Fernwald, D
Magnetpartikelkonzentrator	Dynal MPC-E	Dynal, Oslo, NO
Zell-Zähler	CASY 1 TT	Schärfe System, Reutlingen, D

2.1.7. Verbrauchsmaterialien

Produkt	Hersteller
Amicon Ultra-0.5 Centrifugal Filter Unit	Merck Millipore, Billerica, USA
Phase Lock Gel Heavy 2 ml Tubes	5 PRIME GmbH, Hilden, D
SureHyb Gasket slides (für 2x Mikroarrays)	Agilent, Santa Clara, USA
SureHyb Gasket slides (für 4x Mikroarrays)	Agilent, Santa Clara, USA
SureHyb Gasket slides (für 8x Mikroarrays)	Agilent, Santa Clara, USA
SureHyb Hybridisierungskammer 62534A	Agilent, Santa Clara, USA
Slidehalter (für Scanner)	Agilent, Santa Clara, USA
Zellkulturflaschen	TPP, Trasadingen, CH
Zellschaber	TPP, Trasadingen, CH
Rnase Zap Wipes	Ambion, Wiesbaden, D

2.1.8. Software

Software	Referenz/Hersteller
Beanplot	[186]
BEDTools	[326]
cluster	[245]
Circos	[210]
FACSDiva	BD Biosciences, San Jose, USA
Feature Extraction 10.5.11	Agilent, Santa Clara, USA
GenePix Pro 5.0	Molecular Devices, Sunnyvale, USA
Genesis	[396]
CGHPRO/ GenomeCAT	[67] (K. Tebel et al., Manuskript eingereicht)
ggplot2	[438]
hicpipe	[448]
HOMER	[156]
Java Treeview	[350]
LiftOver	[195]
LifeScope Genomic Analysis 2.5.1	Life Technologies, Carlsbad, USA
Limma	[431]
MASS	[421]
PANTHER	[274]
R	[328]
Reshape2	[437]
RStudio	[329]
RazerS 3	[433]
TADBit	[365]
VennDiagram	[62]
visPIG	[354]

2.1.9. Öffentliche Daten und Datenbanken

Datensatz/ Datenbank	Referenz
4DGenome	[406]
<i>Alu</i> -Repeats	[377]
Chromatinstatus (ChromHMM, IMR90, GM12878, NHEK, NHLF, HUVEC, HMEC, H1- hESC, K562)	University of California, San Diego: Human Reference Epigeno- me Mapping Project

Fortsetzung auf der nächsten Seite

2.1. Material

Datensatz/ Datenbank	Referenz
CTCF ChIA-PET interactions (MCF-7) (GSM970215)	[226]
CTCF ChIP-seq Signale (GSM935404)	[376]
Database of Genomic Variants (DGV)	[170]
EnrichR	[61]
Ensembl	[123] [451]
ESCAPE	[447]
Evolutionäre Bruchpunkte	[225]
<i>Fragile sites</i>	[131]
GATHER	[59]
<i>Genomic Disorders</i> (DECIPHER)	[119]
Gene Ontology (GO) Database	[154]
GEO	[101]
H3K9me3 (GSM469974)	University
H3K27me3 (GSM469968)	of
H4K20me1 (GSM521915)	California,
H3K4me2 (GSM521900)	San
H3K4me3 (GSM469970)	Diego:
H3K18ac (GSM521884)	Human
H3K14ac (GSM521883)	Reference
H3K4me1 (GSM521897)	Epigenome
H3K27ac (GSM469966)	Mapping
H3K36me3 (GSM521890)	Project
Hi-C GM12878	“
Hi-C IMR90 Replikat 1 in 20kb-Bins und normalisiert	“
Hi-C IMR90 Replikat 1 FASTQ-Rohdaten (GSM862724)	[363]
<i>Human Chained Self Alignments</i>	[92]
<i>HindIII</i> -Schnittstellen	[92]
Human Epigenome Browser	[73, 194, 362]
KEGG	[342]
Lamin B1 Interaktionen	[458]
Lamin-assoziierte Domänen (LADs)	[187] [188]
Multi-Genome Synteny Viewer (mGSV)	[143]
NCG 5.0 Cancer Genes	[143]
	[338]
	[5]

Fortsetzung auf der nächsten Seite

Datensatz/ Datenbank	Referenz
non-B Database (Mirror Repeats, zDNA, G4-Motive)	[54]
RefSeq	[301]
<i>Replication timing</i>	[152] [405]
Segmentale Duplikationen	[15] [19]
SRA	[224]
STRING	[399]
Subkompartimente (GM12878)	[333]
Sub-TADs (GM12878, HeLa, HMEC, HUVEC, IMR90, K562, KBM7, NHEK)	[333]
Topologische Domänen (IMR90)	[92]
Topologische Domänen der Maus	[92]
TRANSFAC	[262]
UCSC Genome Browser	[195] [273]
UCSC Table Browser	[191]

2.2. Methoden

2.2.1. Auswertung der intrachromosomalen Interaktionen des IMR90 Chromosom 7

Prozessierung der IMR90 Hi-C-Daten zur Analyse intrachromosomaler Interaktionen des Chromosom 7

Für die Analyse der 3D-Konformation des Chromosom 7 wurden öffentlich verfügbare Hi-C-Daten der Fibroblastenzelllinie IMR90 (hg18) untersucht [92]. Die bereits nach [448] normalisierten Daten des Replikats 1 liegen als 20-kb-Bins vor, denen in Interaktionsmatrizen ihre Kontaktwahrscheinlichkeiten (Interaktionen (IAs)/Binpaar) zu allen anderen Bins des Chromosom 7 zugeordnet sind. Als Präprozessierung wurden diese Binpaare mehreren Filterschritten unterzogen, um Bins mit unzuverlässigen Kontaktwahrscheinlichkeiten zu minimieren. Anschließend wurden die verbleibenden interagierenden Bins zu einem Interaktionsmuster zusammengefasst. Um den Einfluss dieser Filterschritte auf die Reproduzierbarkeit des Interaktionsmusters zu testen, wurden Datensätze mit unterschiedlichen Schwellenwerten für die nachfolgenden Filterkategorien generiert und die resultierenden Interaktionsmuster miteinander verglichen. Dabei wurden folgende Filterkategorien angewendet (Abb. 2.2 A):

- 1) Interaktionsanzahl der Binpaare
- 2) Interaktionsdistanz der Binpaare und
- 3) Überlappung der Binpaare mit SD-Sequenzen.

Filterkategorie 1: Interaktionsanzahl pro Binpaar

Aus der Interaktionsanzahl zwischen den Binpaaren lässt sich auf die räumliche Nähe dieser chromosomalen Regionen untereinander schlussfolgern. Zufallsinteraktionen können die Werte jedoch verfälschen. Eine niedrige Interaktionsanzahl zwischen Binpaaren kann auf solche Zufallsinteraktionen hinweisen, daher wurden die betroffenen Binpaare von der Analyse ausgeschlossen. Für die Festlegung des Schwellenwertes einer geringen Kontaktwahrscheinlichkeit wurden die relativen Häufigkeiten der Interaktionsanzahl pro Binpaar ermittelt (Abb. 2.1). Dabei nimmt die Interaktionsanzahl der Binpaare des Chro-

mosom 7 Werte von 0 bis über 600 IAs/Binpaar ein. Werden die Binpaare ohne Interaktionen (IAs/Binpaar = 0) aus dem Datensatz entfernt, fallen 97 % der verbleibenden Werte auf unter 15 IAs/Bin. Diese Binpaare mit geringen Kontaktwahrscheinlichkeiten, aber dominierender Anzahl gegenüber Binpaaren mit hohen Kontaktwahrscheinlichkeiten, wurden für die folgenden Analysen entfernt. Um die Reproduzierbarkeit dieses Filterschrittes zu testen, wurden auch Datensätze für Binpaare mit den Schwellenwerten von > 10 IAs/Binpaar bzw. 10-50 IAs/Binpaar generiert und die Interaktionsmuster untereinander verglichen (siehe Abb. 3.1 A-F).

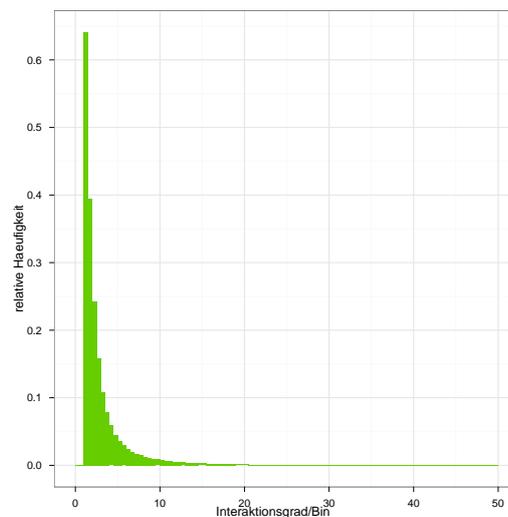


Abbildung 2.1.: Relative Häufigkeiten der Interaktionsanzahl pro Binpaar für das Chromosom 7.

Binpaare ohne Interaktionen wurden für diese Darstellung ausgeschlossen. Ausschnitt für die Werte 1 bis 50 IAs/Binpaar.

Filterkategorie 2: Interaktionsdistanz zwischen Binpaaren

Die Kontaktwahrscheinlichkeiten zwischen chromosomalen Regionen nehmen mit zunehmender linearer Distanz ab [229]. Dementsprechend besteht die Mehrheit der hohen Kontaktwahrscheinlichkeiten von Bins aufgrund ihrer linear benachbarten Position zueinander, was die Identifizierung von Interaktionen zwischen weiter entfernten Regionen erschweren kann. Daher wurden im zweiten Filterschritt Binpaare verworfen, deren Loci eine geringere genomische Distanz als 25 Mb aufweisen. Die verbleibenden Binpaare repräsentieren damit über lange Distanzen interagierende Regionen. Zusammen mit der Filterkategorie 1 werden damit 99,8 % der Binpaare verworfen (Tab. S1). Als Test für

2.2. Methoden

die Reproduzierbarkeit dieses Filterschrittes wurden zusätzlich Datensätze mit einer Interaktionsdistanz von über 10 Mb generiert und deren resultierende Interaktionsmuster verglichen (siehe Abb. 3.1 A-F).

Filterkategorie 3: Überlappung der Bins mit SDs

Unzuverlässige Binpaare können auch durch das fehlerhafte Alignment der Reads an das Referenzgenom auftreten. Reads, welche mit SD-Sequenzen [15] [19] überlappen, können aufgrund der großen Sequenzähnlichkeit zwischen den SDs untereinander nur schwer ihrer tatsächlichen Lokalisation zugeordnet werden. Damit steigt die Gefahr einer fehlerhaften Annotation. Als Folge kann den betreffenden Bins eine verfälschte Anzahl von Interaktionen zugeordnet werden. Um die Auswirkungen dieser Fehlerquelle auf das Interaktionsmuster zu testen, wurden die Hi-C-Datensätze zusätzlich auf Überlappung mit SD-Sequenzen überprüft und je nach Art der Überlappung in drei weitere Kategorien unterteilt (Abb. 2.2 A und 3.1 I): Die Binpaare überlappen 3a) mit allen SDs, auch mit parallel zu dem Binpaar verlaufende SD-Paraloge (= alle Bins werden beibehalten); 3b) mit SDs, jedoch nicht mit Paralogen und 3c) mit keinen SD-Sequenzen. Diese Filterung der Binpaare ergibt, im Vergleich zur ungefilterten Referenz 3a, eine Reduktion der Binpaare (bei gleichzeitiger Anwendung der Filterkategorien > 15 IAs/Binpaar und > 25 Mb Interaktionsdistanz) um 0,14 % (für Kategorie 3b) bzw. 59,77 % (für Kategorie 3c) (Tab. S1). Die Filterung wurde unter Verwendung von BEDTools „pairtopair“ [326] durchgeführt. Die resultierenden Datensätze der verschiedenen Kombinationen an Schwellenwerten werden im Folgenden genutzt, um die Reproduzierbarkeit des resultierenden Interaktionsmusters des Chromosom 7 zu überprüfen.

Zusammenfassung parallel verlaufender interagierender Binpaare zu Interaktionsbündel

Die nach allen Filterschritten verbleibenden Binpaare (Abb. S2) wurden nun mittels dem Circos Tool „bundlelinks“ [210] zu interagierenden chromosomalen Regionen zusammengefasst (Abb. 2.2 B). Hierbei wurden benachbarte Binpaare zu einem Interaktionsbündel fusioniert, wenn mindestens fünf parallel verlaufende Binpaare in einem 500-kb-Fenster vorliegen (für die Alternativdatensätze der unterschiedlichen Schwellenwerte wurde für das Bündeln der Interaktionsbins die sich ändernde Anzahl von Bins berücksichtigt). Die

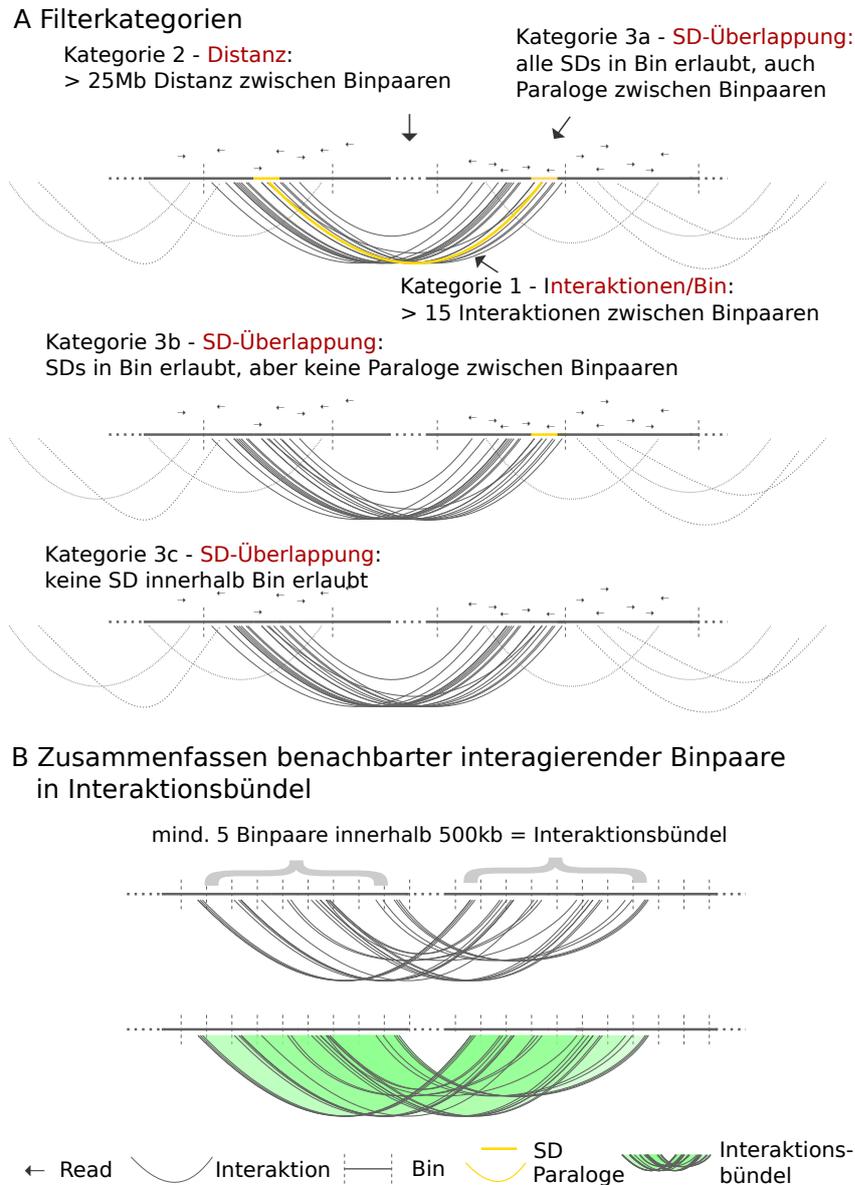


Abbildung 2.2.: Prozessierung der normalisierten IMR90 Hi-C-Daten.

A Schema der Filterkategorien zur Reduzierung unzuverlässiger Binpaare. Readpositionen werden zu Bins zusammengefasst und die Interaktionen zwischen den Binpaaren gezählt. Anhand der folgenden Kategorien wurden die Bins gefiltert: Kategorie 1 - Anzahl der Interaktionen pro Binpaar, Kategorie 2 - Distanz zwischen Binpaaren, Kategorie 3 - Überlappung der Bins mit SD-Sequenzen **B** Benachbarte parallel verlaufende Binpaare werden zu Interaktionsbündeln zusammengefasst. Read - Pfeil, Bin - waagerechte Linie mit Start- und Endposition, SDs - gelbe Linie, Interaktionen zwischen einem Binpaar - Bogenlinie, Interaktionsbündel - grün schattiert

2.2. Methoden

resultierenden Interaktionsbündel wurden über Circos visualisiert.

Berechnung des überlappenden Grundmusters aller gebündelter Filterdatensätze

Zur Untersuchung der Übereinstimmung der von intrachromosomalen Interaktionen betroffenen Regionen aller 12 Filterdatensätze wurden entlang des Chromosom 7 für jedes 20-kb-Bin die Häufigkeit der Überlappung mit intrachromosomal interagierenden Regionen bestimmt (max=12). Zusätzlich wurde untersucht, welche gebündelten intrachromosomalen Interaktionen in allen 12 Datensätzen auftreten. Mittels BEDTools „pairtopair“ unter Verwendung der Option *both* wurden die überlappenden Interaktionsbündel zweier Datensätze ermittelt. Um die resultierenden Interaktionsbündel auf die Regionen einzuschränken, welche an allen Interaktionsbündeln beteiligt sind, wurden die Datensätze über BEDTools „intersect“ sukzessive verglichen und das überlappende Grundmuster bestimmt.

Filterung des IMR90-Hi-C-Datensatzes auf Readebene

Parallel zur Analyse des Einflusses der Sequenzsimilaritäten auf das Interaktionsmuster der Bins (Filterkategorie 3) wurden auf Ebene der *paired-end* Reads die FASTQ-Rohdaten der intrachromosomalen Interaktionen des Chromosom 7 (GSM862724; hg18) [92] auf Überlappungen mit SD-Sequenzen gefiltert. Basierend auf der Methode zur Identifizierung von *singly unique nucleotides* (SUNs) [397] wurden alle Regionen, welche von SD-Sequenzen abgedeckt werden, als Intervalle zusammengefasst, in 30 bp lange Reads unterteilt und mittels des Read Mapping Tools RaserS 3 [433] im Referenzgenom verortet. Dabei wurden nur 30mere berücksichtigt, welche einmalig im Genom und mit einer maximalen Abweichung von 2 bp vorkommen. Die resultierenden 30mere wurden als eindeutig zuordenbar angesehen und als Intervalle zusammengefasst. Alle Reads der ungefilterten Rohdaten, welche Überlappungen mit diesen Intervallen aufwiesen, wurden für die Analyse beibehalten, alle weiteren verworfen. Die verbliebenen Reads wurden in 20-kb-Bins zusammengefasst, die Readanzahl pro 20-kb-Bin berechnet und über die erwarteten Kontaktwahrscheinlichkeiten des Binpaars (entsprechend der ungefilterten Readpaare [92]) gemäß hicpipe [448]) normalisiert. Die resultierenden normalisierten Binpaare und ihre zugeordneten Kontaktwahrscheinlichkeiten wurden auf Binpaare mit mindestens 15 IAs/Bin und einer Interaktionsdistanz von mindestens 25 Mb gefiltert und zu Interaktionsbündeln zusammengefasst (entsprechend der oben behandelten Datensätze).

Resampling der SDs in Interaktionsbündel

Die Anreicherung von SDs innerhalb von Interaktionsbündelregionen (mit einer Interaktionsdistanz von mindestens 25 Mb) wurde für vier verschiedene SD-Kategorien überprüft. Diese berücksichtigen alle intra- und interchromosomalen SDs [15] [19] des humanen Chromosoms 7, mit folgenden Filteroptionen: SDs mit über 98 % Sequenzsimilarität, 98 % bis 99 % Sequenzsimilarität bzw. alle Sequenzsimilaritäten vereint. Hierfür wurde die Summe der überlappenden Basenpaare von Regionen mit SDs und Regionen mit Interaktionsbündel des Chromosoms 7 berechnet. Letzterer Datensatz wurde mit den Filteroptionen von über 15 IAs/Bin und einer Interaktionsdistanz von über 25 Mb erstellt (siehe oben). Folgende Schritte wurden für jede einzelne SD-Kategorie durchgeführt:

1. Überlappende Intervalle innerhalb der einzelnen SD- bzw. Interaktionsbündel-Datensätze wurden fusioniert mithilfe von BEDTools „mergeBed“ [326].
2. Der Basenpaarüberlapp der SD-Datensätze mit den Interaktionsbündelregionen wurde berechnet (= beobachteter Basenpaarüberlapp) mithilfe von BEDTools „coverageBed“.
3. Für den Vergleich des beobachteten Basenpaarüberlapps mit zufällig generierten Werten wurden die vier verschiedenen SD-Kategorien einem Resampling unterzogen. Dabei wurden folgende Konditionen eingehalten: Die Größe und das Chromosom jedes Intervalls wurde auf den Randomdatensatz übertragen, die Intervalle zeigen keine Überlappungen und befinden sich nicht in Annotationslücken des humanen Genoms. Es wurden insgesamt für jeden Datensatz 10.000 zufällig generierte Datensätze unter Verwendung von BEDTools „shuffleBed“ erstellt. Für jedes der 10.000 Datensätze wurde anschließend der Basenpaarüberlapp mit den Interaktionsbündelregionen berechnet (= erwarteter Basenpaarüberlapp).
4. Der *Fold Change* des beobachteten Basenpaarüberlapps wurde als Verhältnis von beobachteten Basenpaarüberlapp zu dem Durchschnitt aus 10.000 erwarteten Werten der Basenpaarüberlappung berechnet.
5. Die Anzahl der erwarteten Basenpaarüberlappungen, die größer oder gleich des beobachteten Basenpaarüberlapps sind, wurde für jede SD-Kategorie ausgezählt und für die Berechnung des p -Wertes verwendet, die für Monte-Carlo-Simulationen von

2.2. Methoden

Smyth et al. [318] empfohlen wird:

$$p_u = P(B \leq b) = \frac{b + 1}{m + 1} \quad (2.1)$$

Die Angleichung der resultierenden p -Werte für die multivariate Berechnung erfolgte nach der Benjamini-Hochberg-Methode. Die Histogramme der Verteilung der erwarteten Basenpaarüberlappungen jeder SD-Kategorie wurden mittels des R-Paketes „ggplot2“ [438] generiert.

Einige zusätzliche Datensätze wurden mithilfe des Resamplings verglichen. So wurden die Intervalle aller SDs mit dem überlappenden Grundmuster der intrachromosomalen Interaktionsmuster des Chromosoms 7 verglichen (siehe oben). Weiterhin wurden Interaktionsbündelregionen aller Chromosomen mit allen SDs bzw. nur intrachromosomalen SDs auf Überlappungen untersucht.

2.2.2. Zellkultur

Die beiden Fibroblasten-Zelllinien IMR90 und IMR91L wurden in Eagle's Minimum Essential Medium (EMEM) mit den Zusätzen 10 % fetales Kälberserum (*fetal calf serum*; FCS), 2 mM UltraGlutamin I, 1 mM Natriumpyruvat und 100 Units/ml Penicillin/Streptomycin bei 37 °C und 5 % CO₂-Atmosphäre kultiviert.

Seax-Zellen wurden in RPMI 1640, supplementiert mit 2 mM Ultraglutamin I, 15 % FCS und 100 U/ml Interleukin-2 (IL-2), und bei 37 °C und 5% CO₂-Atmosphäre kultiviert. Die Kultivierung der HH-, Hut78- und Myla-Zellen erfolgte in RPMI 1640 supplementiert mit 2 mM UltraGlutamin I und 15 % FCS bei 37 °C und 5 % CO₂-Atmosphäre.

Die weiteren T-ALL-Zelllinien Molt-14, Jurkat, KE-37 und H9 wurden entsprechend [260] kultiviert. Die Zelllinien Mewo, SH-SY5Y, Hey und SK-BR-3 wurden entsprechend der Vorgaben des Herstellers (ATCC, Teddington, UK) kultiviert.

2.2.3. Chromatinimmunpräzipitation

Die Chromatinimmunpräzipitation (ChIP) wurde mit dem Transcription Factor ChIP Kit und nach den Empfehlungen des Herstellers (Diagenode, Liège, BEL) durchgeführt. Die Zelllinien IMR90 bzw. IMR91L wurden in einer 150 cm² Zellkulturflasche mit jeweils 2×10^6 Zellen ausgesät und 48 Stunden bei 37 °C und 5 % CO₂ inkubiert. Bereits vor der Zellernte wurden die Zellen über Formaldehyd fixiert. Dazu wurde der Zellkultur Formaldehyd (Endkonzentration: 1 %) zugegeben und anschließend für weitere 10 Minuten bei Raumtemperatur inkubiert. Zum Abstoppen der Reaktion wurde 125 mM Glycin zugegeben und die adhärenen Zellen zweimal mit eiskaltem DPBS gewaschen. Die Ernte der Zellen erfolgte mit einem Zellschaber unter Zugabe von eiskaltem DPBS. Durch das Zentrifugieren bei $1350 \times g$ für 10 Minuten bei 4 °C wurden die Zellen in einem Röhrchen gesammelt. Nach Resuspension in eiskaltem DPBS wurden 3×10^6 der geernteten Zellen abgenommen, über einen Zentrifugationsschritt pelletiert, in flüssigem Stickstoff schockgefroren und bei -80 °C aufbewahrt. Das Pellet wurde für das Experiment entsprechend dem Transcription Factor ChIP Kit resuspendiert und die Zellen für 60 Zyklen mit jeweils 30 Sekunden Puls und 30 Sekunden Pause mithilfe des Bioruptors bei 4 °C sonikiert. Anschließend wurden die gescherten Proben bei $14.000 \times g$ (13.000 rpm) und 4 °C abzentrifugiert. Für die Immunpräzipitation wurden 30 µl des ChIP-Ansatzes ($\cong 1 \times 10^6$ Zellen), 5 µg des Antikörpers gegen acetyliertes Histon Lysin 8 (pAb-103-050) und 30 µl aufbereitete A/G-beschichtete Beads zugegeben und über Nacht und bei 4 °C unter Rotation inkubiert. 10 µl des gescherten Chromatins wurden für die Kontrolle aufbewahrt. Anschließend wurden die Proben für zwei Minuten bei $500 \times g$ (3.000 rpm) und 4 °C pelletiert und entsprechend den Anleitungen des Herstellers mit den Transcription Factor ChIP Kit -Waschpuffern bei 4 °C gewaschen. Für die Eluierung der DNA von den Beads wurde das IPure Kit (Diagenode, Liège, BEL) unter Verwendung des dazugehörigen Protokolls genutzt. Hierzu wurden die gewaschenen Beads in 100 µl Elutionspuffer und 10 µl der Kontroll-DNA in 90 µl Elutionspuffer aufgenommen und für vier Stunden bei 65 °C unter schütteln im Thermomixer inkubiert. Die Beads wurden für zwei Minuten bei $500 \times g$ und Raumtemperatur pelletiert und in 100 µl 100%igen Isopropanol und 2 µl Carrier Buffer aufgenommen bzw. die Input-DNA damit versetzt. Anschließend wurden 15 µl magnetische Beads des Kits dazugegeben und für eine Stunde bei Raumtemperatur rotiert. Die Waschstritte der Beads wurden entsprechend des IPure Kit Protokolls mit jeweils 100 µl Waschpuffer und einer fünfminütigen Inkubationszeit auf dem Rotator durchgeführt. Die sukzessive Eluierung der DNA erfolgte in Buffer C bei einer Inkubationszeit von zweimal

2.2. Methoden

15 Minuten auf dem Rotator.

Die folgenden ChIP- und methylDNA-Experimente wurden bereits während meiner Diplomarbeit und in Zusammenarbeit mit Ines Müller und Vivien Boldt angefertigt. Für die ChIP-Experimente der Zelllinie Seax bzw. HH, Hut78, Myla, SkBr3, SH-SY5Y, Mewo und Hey wurden pro Ansatz 5×10^7 bis 1×10^8 Zellen geerntet und entsprechend des Mammalian ChIP-on-chip-Protokoll von Agilent quervernetzt, an magnetische Beads gekoppelt, über Nacht mit dem entsprechenden Antikörper inkubiert, lysiert und aufgereinigt. Für die Immunpräzipitation der methylierten DNA (*Methylated DNA Immunoprecipitation*; **MeDIP**) wurden 5 µg genomische DNA für 4 Sekunden sonikiert. Anschließend wurden 2 µg des Ansatzes abgenommen, bei 95 °C für 10 Minuten denaturiert, auf 4 °C abgekühlt und mit kaltem IP-Puffer (10 mM Natriumphosphat, 140 mM NaCl, 0,05 % Triton X-100, pH 7.0) auf 245 µl aufgefüllt. Zusammen mit 5 µg Antikörper gegen 5-Methylcytidin (BI-MECY-0100) wurde der Ansatz für drei Stunden bei 4 °C unter Rotation inkubiert. Anschließend wurden 15 µl vorbereitete Dynabeads M-280 Anti-Maus IgG (Invitrogen, Groningen, NL) zugegeben. Die Reaktionsgefäße wurden für drei Minuten in den Dynal MPC-S Magnetständer (Invitrogen, Groningen, NL) platziert und der Überstand abgenommen. Zu den Beads wurden 700 µl IP-Puffer gegeben und diese unter auf- und abpipettieren gewaschen und erneut im Magnetständer platziert. Dieser Waschschrift wurde dreimal wiederholt. Die DNA wurde dann über die Zugabe von 15 µl Ampullenwasser von den Beads eluiert.

Die resultierende DNA der ChIP- und methylDNA-Experimente wurde mit dem GenomePlex Whole Genome Amplification (WGA) Kit amplifiziert, über das QIAquick PCR Purification Kit erneut aufgereinigt und über das BioPrime Total Genomic Labeling System fluoreszenzmarkiert. Dies und die Hybridisierung auf den 36k *whole human genome tiling path* BAC Array (GPL5114) und die anschließende Prozessierung mittels der Analysesoftware GenePix Pro 5.0 erfolgte nach [109].

2.2.4. Analyse der früh-apoptotischen DNA-Degradation

Jeweils 2×10^6 Zellen der Fibroblastenzelllinien IMR90 und IMR91L wurden für 4 Stunden bei 37 °C dem Apoptose-induzierenden Staurosporin (1 µmol/L Staurosporin/ 0,1 % DMSO), einem Proteinkinaseinhibitor, ausgesetzt. Als Kontrollansatz wurden der gleichen Anzahl an Zellen nur 0,1 % DMSO zugesetzt. Von diesem Ansatz wurden 5 bis 10×10^6

Zellen entnommen, mit $1 \times$ Annexin-Bindungspuffer gewaschen und mit Annexin V-APC und 7-Aminoactinomycin D (**7-AAD**) für 15 Minuten angefärbt und das Fortschreiten der Apoptose über Durchflusszytometrie (*fluorescent activated cell sorting*; **FACS**) verfolgt. Die restlichen Zellen wurden mit einem Lysepuffer und $15 \mu\text{g/ml}$ RNase A versetzt und einem RNA-Verdau für eine Stunde bei 37°C unterzogen. Zur Deproteinierung der Zellysate wurden dem Ansatz anschließend 1 M Natriumperchlorat und 1 Volumen Chloroform hinzugegeben und anschließend einer DNA-Präzipitation über Ethanol unterzogen. Das Ausmaß der DNA-Fragmentierung wurde dann mit Genomic DNA Screentape der Agilent 2200 Tap2station analysiert. Zur Trennung von wenig degradierter DNA ($> 48 \text{ kb}$) von degradierter apoptotischer DNA ($\sim 4 \text{ kb}$) wurde der Ansatz mit einem 1%igen *low melt* Agarosegel über Elektrophorese aufgetrennt, die entsprechenden Banden ausgeschnitten und mit β -Agarase I entsprechend der Herstellerangaben verdaut.

2.2.5. Mikroarrayhybridisierung der H4K8ac-ChIP- und DNA-Degradationsexperimente

Zur Amplifizierung der isolierten DNA des ChIP-Experimentes bzw. des DNA-Degradationsexperimentes wurde das GenomePlex Whole Genome Amplification Kit genutzt. Die amplifizierten Ansätze der hochmolekularen DNA ($> 48 \text{ kb}$) und degradierter apoptotischer DNA ($\sim 4 \text{ kb}$) bzw. ChIP DNA und Input DNA wurden auf einem 400k *whole genome* Oligonukleotidarray (GPL9777, hg19) und einen für die Region chr7:69936560-70795513 spezifischen Oligonukleotidarray (GPL17964, hg19) bei 65°C und 40 Stunden kohybridisiert. Die Hybridisierung und das Scanning erfolgte entsprechend den Anweisungen des Herstellers (siehe auch Abschnitt 2.2.11). Die erste Auswertung inklusive der Bildanalyse, Normalisierung und Annotation erfolgte mit Feature Extraction 10.5.1. unter Anwendung der voreingestellten Optionen. Diese Daten sind unter der Identifikationsnummer GSE41356 in der Datenbank Gene Expression Omnibus (**GEO**) zu finden. Die Visualisierung der resultierenden Daten erfolgte mit GenomeCAT und des Human Epigenome Browsers.

2.2.6. RNA-Expressionsprofiling

Die RNA der Zelllinie IMR91L wurde mittels TRIzol extrahiert. Als Qualitätskontrolle wurden anschließend zu 10 µg der aufgereinigten RNA der ERCC Spike-In Control Mix zugegeben. Die Entfernung der rRNA wurde mittels des RiboMinus Kits realisiert. Die RNA wurde dann entsprechend der Protokolle des SOLiD 5500xl Genetic Analyzers über das SOLiD Total RNA-Seq Kit für die Sequenzierung aufgearbeitet (für Geräte und zusätzliche Materialien siehe Protokoll des Herstellers bzw. [99]). Die Sequenzierung erfolgte durch den SOLiD 5500xl Genetic Analyzer unter Verwendung des SOLiD 6-Lane Flowchip. Die resultierenden Readpaare entsprechen einer Länge von 50 bp (*forward*) bzw. 35 bp (*reverse*). Die Daten der RNA-Seq der IMR91L-Zelllinie sind unter der Identifizierungsnummer SRS366467 in der Datenbank Sequence Read Archive (SRA) veröffentlicht.

Für die weitere Auswertung wurden Reads, welche innerhalb eines kodierenden Exons (RefSeq) verortet wurden und dem kodierenden Strang entsprechen den kodierenden RNAs zugeordnet, während alle anderen Reads zu den nichtkodierenden gruppiert worden sind. Die beiden Datensätze wurden anschließend für die Visualisierung über Circos in 100-kb-Bins unterteilt. Die Anzahl der Reads pro Bin wurde dann im Circos-Plot aufgetragen.

2.2.7. Feinkartierung evolutionärer Bruchpunkte und *in silico* Revertierung der Inversionen

Die Alignments wurden über Ensembl *release* 67 unter Verwendung von Perl API heruntergeladen: *Gorilla gorilla gorilla* (gorGor3.1; Mai 2011), *Pongo abelii* (WUGSC2.0.2/-ponAbe2; Juli 2007) und *Homo sapiens* (GRCh37/hg19; Februar 2009). Die proximalen und distalen Bruchpunkte beider Inversionen wurden über das Plotten des Orang Utan-Genoms gegen das humane Genom identifiziert. Die Visualisierung erfolgte als Dot Plot. Inversionen werden hier über invertierte Homologien des Alignments identifiziert und sind als absteigende Linien repräsentiert, wobei die Bruchpunkte des Rearrangements als Unterbrechungen der einzelnen Segmente erscheinen. Anhand der identifizierten Bruchpunkte für die parazentrische Inversion (chr7:76646908 und chr7:102118853) und der perizentrischen Inversion (chr7:6875820 und chr7:80857936) wurde die genomische Position der entsprechenden Segmente revertiert und damit auch die der dort lokalisierten intrachro-

mososomalen Interaktionen und SDs. Die resultierende Genomorganisation wurde mittels Circos visualisiert.

2.2.8. Analysen zur Syntenie des humanen Chromosom 7 und zur Anreicherung von SDs, *Alu*-Repeats und G4-Motiven

Die syntenischen Regionen (entsprechend Ensembl *release* 67) zwischen dem humanen Chromosom 7 und dem Weißbüschelaffen (*Callithrix jacchus*) wurden mittels LiftOver als hg19-Koordinaten konvertiert. Das humane Chromosom 7 wurde in 200-kb-Bins segmentiert ($n=795$), von denen 125 Sequenzhomologien zu Chromosom 2 des Weißbüschelaffen aufweisen. Der UCSC Genome Browser-Ansicht des Chromosom 7 wurden die syntenischen Segmente und die Verteilung von SDs, *Alu*-Repeats und G4-Motiven gegenübergestellt. Letztere sind dabei als Anzahl des Sequenzelementes pro Bin aufgetragen. Zur statistischen Analyse der Anreicherung von Sequenzelementen innerhalb der syntenischen Regionen wurde der hypergeometrische Test nach [100] angewandt. Dazu wurden die Intervalle der 200-kb-Bins auf neue zufällige Positionen verteilt, um den Einfluss benachbarter Bins mit gleichen Werten (aufgrund der natürlichen genomischen Reihenfolge) zu minimieren. Alle Bins wurden dann in absteigender Reihenfolge entsprechend ihrer assoziierten Werte (Anzahl des Sequenzelements pro Bin) sortiert. Die Anreicherung von Bins, welche zu *C. jaccus* Chromosom 2-syntenisch sind, innerhalb der Bindaten mit der höchsten Anzahl der Sequenzelemente wird über den hypergeometrischen Wert berechnet und durch den minimalen hypergeometrischen Wert (mHG) und dem assoziierten p -Wert repräsentiert.

2.2.9. Weitere Analysen im Zusammenhang der IMR90-Chromatinorganisation des Chromosom 7 bzw. der WBS-kritischen Region

Der Locus der WBS-kritischen Region und die umliegenden Regionen wurden im Zusammenhang mit eigenen und publizierten Daten der Chromatinorganisation im Human Epigenome Browser gegenübergestellt (siehe Tabelle Öffentliche Daten und Datenbanken unter Abschnitt 2.1). Die Daten des apoptotischen DNA-Degradationsmusters und

2.2. Methoden

der H4K8-Acetylierung der Zelllinie IMR90 bzw. IMR91L wurden in 100-kb-Bins zusammengefasst und mittels Spearman's Rangkorrelationskoeffizienten unter Verwendung der *cor.test*-Funktion des R-Paketes „stats“ verglichen. Die 100-kb-Bins wurden weiterhin einer zweidimensionalen Kerndichteschätzung mittels des „MASS“-Paketes in R unterzogen (unter Ausschluss von Bins mit fehlenden Werten) und anschließend mithilfe der *filled.contour*-Funktion des R-Paketes „graphics“ visualisiert.

Die Verteilung der Intensitäten der H4K8ac-Modifikationen und früh-apoptotischen DNA-Degradation (als 20-kb-Bin-Datensätze) wurde auch im Zusammenhang mit LAD-Grenzen untersucht. Die entsprechenden Intensitäten der Regionen gesamtgenomischer LAD-Grenzen wurden (500 kb stromabwärts und stromaufwärts der LAD-Grenze) in einer Heatmap wiedergegeben. Dabei wurden nur LADs berücksichtigt, die einen Abstand von mindestens 500 kb zu benachbarten LAD-Grenzen und selbst eine Mindestgröße von 500 kb aufweisen. Zur Visualisierung wurden die IMR90-Datensätze entsprechend ihrer Intensitäten der LAD-Grenzregionen geclustert (mit dem R-Paket „cluster“) und mit „gplots“ visualisiert.

Zur Bestimmung der Gendichte und Introngröße der Gene des 7q11-Segmentes, der randomisierten 4,8-Mb-Segmente bzw. der 7q11-Segment-flankierenden Regionen von 4,8 Mb stromaufwärts und stromabwärts wurden der Datensatz *canonical genes* und die dazugehörigen Introns (in hg19) über den UCSC Table Browser bezogen. Die Anzahl der Gene pro Segment und die durchschnittliche Intronlänge wurden über BEDTools „intersectBed“ berechnet. Die Gendichte wurde dann als Anzahl der Gene pro Mb wiedergegeben. Für den Vergleich mit den umliegenden Bereichen die einer dem 7q11-Segment entsprechenden Genanzahl aufweisen wurden 115 Gene (entspricht der Genanzahl des 7q11-Segmentes) der flankierenden Regionen des 7q11-Segmentes auf ihre durchschnittliche Intronlänge untersucht. Der statistische Nachweis erfolgte über den zweiseitigen exakten Test nach Fisher. Der Gendichte und Intronlänge des 7q11-Segmentes wurden der durchschnittliche Wert aus 100.000 randomisierten 4,8-Mb-Intervallen des gesamten Chromosom 7 gegenübergestellt (unter Verwendung von BEDTools „shuffle“ und unter Beschränkung der Intervallzuteilung auf das Chromosom 7). Der statistische Nachweis erfolgte über die Bestimmung des *p*-Wertes für Monte-Carlo-Simulationen nach Empfehlung von Smyth et al. [318] (siehe Abschnitt 2.2.1).

Zur Berechnung der durchschnittlichen Interaktionsdistanzen des Chromosom 7 wurden alle normalisierten intrachromosomalen 20-kb-Bins mit mindestens 1 IA/Bin in sechs Distanzklassen eingeteilt: 1) < 500 kb, 2) 500 kb bis weniger als 1 Mb, 3) 1 Mb bis weniger als 5 Mb, 4) 5 Mb bis weniger als 10 Mb, 5) 10 Mb bis weniger als 25 Mb und 6) ≥ 25 Mb.

Für jedes Bin wurde die Anzahl der Interaktionen pro Distanzklasse summiert. Der relative Anteil jeder Kategorie im Bezug zur Gesamtanzahl der Interaktionen pro Bin wurde über die Division der Interaktionsanzahl der jeweiligen Distanzkategorie pro Bin durch die Gesamtinteraktionsanzahl jedes Bins bestimmt. Für den Vergleich mit den weiteren Daten der Chromatinorganisation in Abb. 3.6 wurden die Bins mittels dem LiftOver Tool in hg19 konvertiert.

Für den Vergleich der humanen und murinen TADs am Locus der WBS-kritischen Region bzw. dessen homologe Regionen der Maus wurden die murinen TAD-Intervalle (mm9) mittels LiftOver in hg19 konvertiert. Der visuelle Vergleich der Regionen und die orthologen Gene wurde im Humane Epigenome Browser durchgeführt. Die Positionen orthologer Gene, die an den TAD-Grenzen der Maus lokalisiert sind wurden über den Multi-Genome Synteny Viewer (mGSV) in der entsprechenden humanen Region aufgetragen.

2.2.10. Vorbereitung der Sézary-Tumorzellen der Patienten

Mononukleäre Zellen des peripheren Blutes (*peripheral blood mononuclear cell*; **PBMC**) wurden dem Blut direkt oder aus Leukapherese-Proben entnommen und über die Ficoll-Dichtegradientenzentrifugation isoliert. Zur Anreicherung von Tumorzellen wurden die Zellproben mit Fluorochrom-konjugiertem V β -Antikörper 30 Minuten inkubiert und anschließend über magnetische anti-Fluorochrom Beads isoliert. Einige Tumorproben wurden auch über die anti-CD4 Beads selektioniert (angegeben unter Tab. Patientendaten, siehe S. 33). Über die FACS-Analyse wurde anschließend die Anreicherung der Tumorzellen unter Verwendung von Antikörpern gegen CD3, CD4, CD8 und V β kontrolliert. Die den Patienten entnommenen PBMC-Proben wurden einem TCR-Gen-Klonalitäts-Assay (siehe [389]) unterzogen. Der Nachweis einer großen Zellpopulation mit klonalem Rearrangement weist hierbei auf das Vorliegen von malignen Zellen hin. Nach der DNA-Isolierung mittels Proteinase K-Verdau und anschließender Verwendung des *QIAamp DNA purification Kits* wurde die genomische DNA über das GenomePlex Whole Genome Amplification (WGA) Kit amplifiziert.

2.2. Methoden

2.2.11. Array CGH der Sézary-Proben

Die genomische DNA der Sézary-Patienten und Referenz-DNA wurde nach dem *Array-based CGH for Genomic DNA Analysis - Enzymatic Labeling v6.2*-Protokoll für die Hybridisierung vorbereitet. Für die Fragmentierung wurden 0,5-1 µg DNA (2×Array) mit Ampullenwasser auf ein Volumen von 20,2 µl gebracht und mit 5,8 µl Restriktionsmix versetzt:

2× Mikroarray	
2,0 µl	H ₂ Oamp
2,6 µl	10x Buffer C
0,2 µl	BSA
0,5 µl	Alu I
0,5 µl	Rsa I
<hr/>	
5,8 µl	per Reaktion

Nach dem Mischen und kurzem Zentrifugieren der Proben wurden diese für zwei Stunden bei 37 °C und anschließend für 20 Minuten bei 65 °C inkubiert. DNA die über das GenomePlex Kit amplifiziert worden ist wurde direkt ohne vorherige Fragmentierung eingesetzt mit 0,5 bis 1 µg DNA (2×Array) bzw. 0,2 bis 0,5 µg (8×Array).

Alle weiteren Schritte wurden mit dem Agilent DNA Enzymatic Labeling Enzyme Module auf Eis, unter Ausschluss von Licht und unter Verwendung eines Luftreinigers durchgeführt. Zur Vorbereitung des Labelings der DNA wurden 5 µl Random Primer zu der fragmentierten DNA gegeben, für drei Minuten bei 95 °C denaturiert und für fünf Minuten bei 4 °C abgekühlt. Anschließend wurden die Fluoreszenzfarbstoffe zugegeben, die Ansätze gemischt und abzentrifugiert.

2× Mikroarray		8× Mikroarray
10,0 µl	5x Reaktionspuffer	5,0 µl
5,0 µl	10x dNTPs	2,5 µl
3,0 µl	Cyanine 3-UTP bzw. Cyanine 5-UTP	1,5 µl
1,0 µl	Exo (-) Klenow Fragment	0,5 µl
<hr/>		<hr/>
21,0 µl	per Reaktion	11,5 µl

Die Ansätze wurden dann für zwei Stunden bei 37 °C und anschließend für 10 Minuten bei 65 °C inkubiert und bei 4 °C abgekühlt. Die Aufreinigung der Ansätze erfolgte mit

Amicon-Filtersäulen. Dazu wurden die Proben mit 430 μl 1 \times TE Puffer gemischt und auf die Filtersäulen gegeben. Die beladenen Säulen wurden dann für 10 Minuten bei 14.000 \times g und Raumtemperatur zentrifugiert. Der Durchfluss wurde verworfen und der Säule erneut 480 μl zugegeben und für 10 Minuten bei 14.000 \times g und Raumtemperatur zentrifugiert. Die Säule wurde dann invertiert und die aufgereinigte gelabelte DNA für eine Minute bei 1.000 \times g und Raumtemperatur eluiert. Das Volumen der einzelnen Proben wurde mit 1 \times TE Puffer auf 41 μl (2 \times Mikroarray) bzw. 9,5 μl (8 \times Mikroarray) gebracht und gegebenenfalls mit der Vakuumzentrifuge eingengt. Die Konzentration der einzelnen Proben wurden vermessen. Bei Konzentrationsschwankungen zwischen Patienten- und Referenz-DNA, wurden diese angeglichen. Patienten- und Referenz-DNA wurden vermischt und mit dem Hybridisierungsmastermix (Oligo aCGH/ChIP-on-chip Hybridization Kit) versetzt:

2 \times Mikroarray		8 \times Mikroarray
25,0 μl	humane Cot DNA	2,0 μl
26,0 μl	10x aCGH Blocking Agent	4,5 μl
130,0 μl	2x HI-RPM Hybridization Buffer	22,5 μl
<hr/>		<hr/>
181,0 μl	per Reaktion	29 μl

Zu den Proben wurden jeweils 181 μl (2 \times Mikroarray) bzw. 29 μl (8 \times Mikroarray) des Hybridisierungsmastermixes zugegeben. Die Ansätze wurden dann gemischt, abzentrifugiert, für drei Minuten bei 95 $^{\circ}\text{C}$ denaturiert und für 30 Minuten bei 30 $^{\circ}\text{C}$ inkubiert. Auf die in Hybridisierungskammern platzierten Gasket Slides wurden anschließend 245 μl (2 \times Mikroarray) bzw. 40 μl (8 \times Mikroarray) aufgetragen und verschlossen. Die Arrays wurden im vorgeheizten Hybridisierungsöfen für 40 h (2 \times Mikroarray) bzw. 24 h (8 \times Mikroarray) bei 65 $^{\circ}\text{C}$ inkubiert. Nach dieser Zeit wurden die Gasket Slides in Agilent Oligo aCGH/Chip-on-ChIP Wash Buffer I entfernt und die Arrays in demselben Puffer für fünf Minuten auf einem Magnetprüher inkubiert. Anschließend wurden diese in vorgewärmten Agilent Oligo aCGH/Chip-on-ChIP Wash Buffer II überführt und auf einem Magnetprüher mit Heizelement bei 37 $^{\circ}\text{C}$ für eine Minute inkubiert. Die Slides wurden dann langsam aus dem Puffer entfernt und unter Ausschluss von Licht kurz getrocknet. Für das Scannen der Slides wurden diese in Slidehalter gespannt und unter den folgenden Bedingungen eingescannt (DNA Microarray Scanner G2565BA):

Scan resolution:	2 μm
Scan region:	61 \times 21,6 mm
Green PMT:	100 %
Red PMT:	100 %

2.2. Methoden

Das durch den Scanner generierte TIFF-Bild wurde mit der Agilent Feature Extraction Software unter Verwendung des vorgegebenen CGH_105_Dec08 Protokolls und des jeweiligen Mikroarray-Grids prozessiert. Für die weitere Analyse wurden die .txt-Files der Feature Extraction in die GenomeCAT Software geladen und die \log_2 -Verhältnisse der prozessierten Signale subtrahiert: $\log_2(\text{"gProcessedSignal"}) - \log_2(\text{"rProcessedSignal"})$. Die DNA-Kopienzahländerungen wurden mittels *circular binary segmentation (CBS)* [302] bei einem Schwellenwert von $\pm 0,2$ definiert (Deletionen: $\log_2 < -0,2$; Duplikationen: $\log_2 > 0,2$). Visualisiert wurden die arrayCGH-Daten mithilfe von GenomeCAT (Tebel et al., Manuskript eingereicht) und Circos [210]. Die Array-Daten der 20 S ezary-Patienten sind in der GEO-Datenbank unter der Nummer GSE21730 zu finden.

Plattformzuordnung: Zuordnung der verwendeten Arrayplattformen zu den Tumorzellproben der 20 S ezary-Patienten und der Zelllinie Seax.

In blau markierte Experimente wurden f ur die Abb. 3.8 genutzt. BAC: 36k genomweiter *tiling path* BAC-Array. 400k: 400k genomweiter Oligonukleotidarray. 60k: spezifischer 60k Oligonukleotidarray der 19p-Region (Tabelle aus [389]).

Patient Nr.	Arrayplattform
1	BAC 400k -
2	BAC - 60k
3	BAC 400k -
4	BAC - 60k
5	BAC 400k -
6	BAC 400k -
7	BAC - 60k
8	BAC 400k -
9	BAC - -
10	BAC - -
11	BAC 400k -
12	BAC 400k -
13	BAC 400k -
14	BAC - -
15	- 400k -
16	BAC 400k -
17	BAC 400k -
18	BAC 400k -
19	BAC 400k -
20	BAC 400k -
Seax	BAC - -

Entsprechend der Tabelle Plattformzuordnung wurden zur Detektion von DNA-Kopien-

zahlveränderungen der Patientenproben und der Zelllinie Seax verschiedene Mikroarray-Plattformen eingesetzt. Die erste Analyse wurde während meiner Diplomarbeit mittels der BAC-Arrays und entsprechend der Angaben unter [109] durchgeführt unter Verwendung des BioPrime Total Genomic Labeling System, der 36k *whole human genome tiling path* BAC Arrays (GPL5114) und der Analysesoftware GenePix Pro 5.0. Die Identifizierung der DNA-Kopienzahlveränderungen erfolgte nach Anwendung der Subgrid LOWESS (Locally Weighted Scatter Plot Smooth)-Normalisierung und nach Segmentierung der Daten über den CBS-Algorithmus [302] (Deletionen: $\log_2 < -0,2$, Duplikationen: $\log_2 > 0,2$). Die genaue Bruchpunktbestimmung erfolgte dann während der Doktorarbeit mit 400k *whole genome* Oligonukleotidarrays (GPL9777) und einem regionenspezifischen Oligonukleotidarray für 19p (GPL10304). Dieser umfasst 52.828 gleichmäßig verteilte Oligonukleotide der Region chr19: 1–28.500.000 (NCBI36; HG18).

2.2.12. FISH

Die FISH-Analyse der Patientenproben und der Seax-Zellen wurde von Prof. Dr. E. Schröck (TU Dresden) durchgeführt. Die angereicherten Zellen wurden hierfür mit einer fluoreszenzmarkierten DNA-Sonde gegen das *E2A*-Gen (BAC-Klon: RP11-690N6) bzw. der Kontrollsonde gegen 19qter (D19S989; Kreatech) inkubiert. Hybridisierung, Detektion, Zwei-Farben-FISH-Bildaufnahmen und Bildanalyse durchgeführt nach [359].

2.2.13. Immunhistochemie

Die Immunhistochemie wurde in der Arbeitsgruppe von Dr. S. Mathas (MDC Berlin) angefertigt. Dazu wurden 4 μm Schnitte von formalinfixierten und paraffineinegebetteten Hautbiopsien entsprechend nach [260] analysiert. Dabei wurden monoklonale Antikörper gegen E12/E47 (Klon G98-271; BD Biosciences, San Jose, USA), CDK6 (Klon DCS-83; Progen Biotechnik, Heidelberg, D), myc (Klon Y69; Epitomics), and CD3 (Klon LN10; Novocastra Laboratories) verwendet.

2.2.14. Transfektion, RNA-Isolierung und *real-time* PCR der Seax-Zellen

Transfektion, RNA-Isolierung und *real-time* PCR der Seax-Zellen wurden in der Arbeitsgruppe von Dr. Stephan Mathas (MDC) durchgeführt. Hierzu wurden die kultivierten Zellen mit 30 µg myc-getaggten E47, *E47-forced dimer expression construct* [370], [230] bzw. Kontrollplasmid pcDNA3.1 (Invitrogen, Carlsbad, USA) transfiziert. Gleichzeitig wurden 10 µg pEGFP-N3 (Takara Bio Inc., Kusatsu, JPN) zur Transfektionskontrolle eingebracht. Die Selektion von GFP-exprimierenden Zellen erfolgte 48 bis 72 Stunden nach der Transfektion durch FACS-Zellsortierung. Die RNA-Extraktion der resultierenden Zellen wurde mithilfe des RNeasy Mini Kits durchgeführt. Die Ansätze der *Knock-downs* wurden als biologische Duplikate angefertigt. Für weitere Spezifikationen der Transfektion bzw. *real-time* PCR, inklusive der verwendeten Primer, siehe [389].

2.2.15. EMSA

Electrophoretic mobility shift assay (EMSA) wurde in der Arbeitsgruppe von Dr. S. Mathas (MDC) entsprechend nach [261] durchgeführt. Zur Detektion wurden folgende Oligonukleotide genutzt: E2A (µE5) *sense* 5'-AGCTCCAGAA-CACCTGCAGCAG-3' und E2A (µE5) *antisense*, 5'-AGCTCTGCTGCA-GGTGTTCTGG-3'; Sp1 *sense*, 5'-AGCTATTCGATCGGGGCGGGGCG-AGC-3' and Sp1 *antisense*, 5'-AGCTGCTCGCCCCGCCCCGATC-GAAT-3'. Nach dem Annealing wurden die Oligonukleotide mittels des Klenow-Fragmentes unter Verwendung von [α -³²P]dCTP radioaktiv markiert. Die Position der Protein-DNA-Komplexe wurden anschließend über Autoradiographie bestimmt. Für die Supershift-Analyse wurden der monoklonale Maus-Antikörper gegen E12/E47 (Klon G98-271; BD Biosciences, San Jose, USA) verwendet, während für die Immunblot-Analysen folgende Antikörper genutzt wurden: monoklonaler Maus-Antikörper gegen E12/E47 (Klon G98-271; BD Biosciences, San Jose, USA), monoklonaler Maus-Antikörper gegen E47 (Klon G127-32; BD Biosciences, San Jose, USA), polyklonaler Ziegen-Antikörper gegen PARP-1 (sc-1561; Santa Cruz Biotechnology, Dallas, USA) und monoklonaler Maus-Antikörper gegen β -Aktin (A5316; Sigma-Aldrich, St. Louis, USA). Die Filter wurden mit Meerrettichperoxidase-konjugierten sekundären Antikörpern inkubiert und die Banden schließlich über das Amersham ECL Western Blotting Detection System (GE Healthcare, Little Chalfont, UK) vi-

sualisiert.

2.2.16. Proliferationsanalysen

Proliferationsanalysen wurden von Dr. M. Möbs (Charité Berlin) durchgeführt. Dabei wurde die Proliferationsrate der Zellen über die Einbaurate von [³H]-Thymidin während der DNA-Synthese unter Verwendung von Standardprotokollen ermittelt. Weiterhin wurde für eine Zellzyklusanalyse die Einbaurate von Bromdesoxyuridin (BrdU) detektiert unter gleichzeitiger Anfärbung der Zellen mit 7-AAD zur Bestimmung des DNA-Gehaltes der Zellen. Die Zellzyklusposition der Zellen wurde mit dem APC BrdU Flow Kit (BD Biosciences, San Jose, USA) identifiziert. Dabei wurde 48 h nach der Transfektion den Zellen BrdU zugegeben und 30 Minuten inkubiert. Anschließend wurden BrdU-Einbaurate und 7-AAD-Färbung mittels Durchflusszytometrie der EGFP⁺-Zellen analysiert.

2.2.17. Genexpressionsarray der Zelllinie Seax

Die aus den *E2A-knock-down* und zugehörigen Kontroll-Experimenten isolierte RNA wurde nach dem One-Color Microarray-Based Gene Expression Analysis Protocol (Quick Amp Labeling) Version 5.7 (G4140-90040v5.7; Agilent, Santa Clara, USA) und unter dem Einsatz des Gene Expression Hybridization Kits weiter prozessiert. 500 ng RNA wurden zu 1,2 µl T7 Promoter Primer (Low RNA Input Linear Amplification Kit) und 5 µl 1:5000 in einer Verdünnungsreihe verdünnten Agilent One-Color Spike-Mix gegeben. Mit nuklease-freiem Wasser wurde der Ansatz auf 11,5 µl aufgefüllt und anschließend für 10 Minuten bei 65 °C denaturiert. Nach einem 5-minütigen Abkühlungsschritt bei 4 °C wurde 8,5 µl des folgenden cDNA-Mastermixes zugegeben:

4,0 µl	vorgewärmter 5x First Strand Buffer
2,0 µl	0,1 M DTT
1,0 µl	10 mM dNTP-Mix
0,5 µl	RNaseOut
<hr/>	
8,5 µl	per Reaktion

Der Ansatz wurde für zwei Stunden bei 40 °C und anschließend für 15 Minuten und bei 65 °C inkubiert. Zum Abkühlen wurden die Proben für 5 Minuten auf 4 °C gehalten. Die

2.2. Methoden

nachfolgenden Schritte wurden unter Lichtabschluß und mit laufenden Luftreinigungsmaschinen durchgeführt. Den Ansätzen wurden 60 µl Transcription Master Mix zugegeben:

15,3 µl	nukleasefreies H ₂ O
20,0 µl	4x Transcription Buffer
6,0 µl	0,1 M DDT
8,0 µl	NTP mix
6,4 µl	vorgewärmtes 50 % PEG
0,5 µl	RNaseOUT
0,6 µl	Inorganic Pyrophosphatase
0,8 µl	T7 RNA Polymerase
2,4 µl	Cyanine 3-CTP
60,0 µl	per Reaktion

Die Ansätze wurden für zwei Stunden bei 40 °C gelabelt. Zu der cRNA wurden 20 µl nukleasefreies Wasser, 350 µl Buffer RLT und 250 µl Ethanol zugegeben. Dieser Ansatz wurde über Säulen des RNeasy Mini Kit aufgereinigt. Die beladenen Säulen wurden für 30 Sekunden bei 13.000 rpm zentrifugiert und zweimal mit 500 µl RPE Buffer gewaschen, wobei beim letzten Zentrifugationsschritt für eine Minute zentrifugiert worden ist. Zur Eluierung der cRNA wurde 30 µl RNase-freies Wasser auf die Säulen gegeben und nach 60 Sekunden für 30 Sekunden bei 4 °C und 13.000 rpm zentrifugiert. Die Konzentration und die Labelingrate der RNA wurde mittels des NanoDrop bestimmt. Die Ansätze wurden mit folgendem Fragmentierungsmix versetzt:

1,65 µg	cRNA
11,0 µl	10x Blocking Agent
auffüllen auf 52,8 µl	nukleasefreies Wasser
2,2 µl	25x Fragmentation Buffer
55,0 µl	Volumen des Reaktionsansatzes

Anschließend wurden die Ansätze für 30 Minuten bei 60 °C inkubiert, mit 55 µl GEx Hybridization Buffer HI-RPM versetzt und je 100 µl auf ein Gasket Slide geladen. Der 4x44K Agilent Whole Human Genome Microarray (GPL6480) wurde auf die so vorbereiteten Proben platziert und mit der Hybridisierungskammer im vorgewärmten Hybridierungs-ofen für 17 Stunden bei 10 rpm und 65 °C inkubiert. Anschließend wurde das Gasket Slide unter Gene Expression Wash Buffer I entfernt und der Mikroarray in eine Küvette mit Wash Buffer I für eine Minute unter Rühren inkubiert. Anschließend wurden die Slides in

vorgewärmten Wash Buffer II überführt und erneut für eine Minute unter Rühren inkubiert. Die Slides wurden langsam aus dem Puffer gehoben, kurz an der Luft getrocknet und mit dem Agilent DNA Microarray Scanner G2565BA und folgenden Einstellungen gescannt:

Scan region:	61 x 21.6 mm
Scan resolution:	5 μm
Extended Dynamic range:	selected
TIFF:	16 bit
Dye channel:	green
Green PMT XDR Hi:	100 %
Green PMT XDR Lo:	10 %

Das durch den Scanner generierte TIFF-Bild wurde mit der Agilent Feature Extraction v10.5.1.1 Software unter Verwendung des vorgegebenen GE1_105_Dec08 Protokolls und des jeweiligen Mikroarray-Grids prozessiert. Für die Datenanalyse wurden die Feature Extraction-Ausgabertextdateien in das Statistikprogramm R v2.9.1 [328] bzw. RStudio 0.95.265 [329] hochgeladen. Mittels des R-Paketes „Limma“ [431] wurden die Genlisten der durch *E2A*-Wiedereinbringung verändert exprimierten Gene in der Sézary-Zelllinie Seax erstellt (unter Verwendung der Optionen: backgroundcorrection=half, log₂-Transformation, normalize between Arrays: quantile). Die *in silico* Analyse GO-Annotationen der durch *E2A*-deregulierten Gene erfolgte weiterhin mit den web-basierten Tools bzw. Datenbanken TRANSFAC, PANTHER, STRING, KEGG und EnrichR.

2.2.18. Hi-C der Zelllinie Seax

Das Hi-C-Experiment dient der Analyse der Chromosomenkonformation über die Identifizierung der Kontakthäufigkeiten zwischen inter- und intrachromosomalen Regionen [229]. Über diese lassen sich auch chromosomale Rearrangements identifizieren. Die Etablierung des Hi-C-Protokolls in der Arbeitsgruppe von PD Dr. Reinhard Ullmann (MPIMG) und die Laborarbeiten zur Hi-C-Analyse der Seax-Zelllinie wurden von G. Ebert durchgeführt (Protokoll entspricht ihrer Vorlage).

Es wurden 2×10^7 bis $2,5 \times 10^7$ Zellen der Zelllinie Seax über die Zugabe von 16%igem Formaldehyd und anschließender Inkubation für 10 Minuten und Raumtemperatur einer

2.2. Methoden

DNA-Protein-Quervernetzung unterzogen. Das Abstoppen der Reaktion erfolgte mittels der Zugabe von 125 mM Glycin und einer Inkubationszeit von fünf Minuten. Die Zellen wurden für 15 Minuten unter Schwenken auf Eis gehalten und anschließend mittels Zentrifugation pelletiert. Nach einem Waschschrift mit kaltem DPBS wurde das Zellpellet schockgefroren und bei -80 °C gelagert.

Bei Durchführung des Hi-C-Experimentes wurde das Zellpellet unter Zugabe von 800 µl kaltem Hi-C-Lysepuffer (FA ChIP Buffer; Abcam, Cambridge, UK) resuspendiert und für 15 Minuten auf Eis inkubiert. Anschließend wurden die Zellen in einem vorgekühlten Dounce-Homogenisator über 15 Stöße aufgeschlossen, unterbrochen durch jeweils einminütiger Ruhephasen. Die lysierten Zellen wurden bei $2.500 \times g$ pelletiert und zweimal mit 500 µl $1 \times$ NEBuffer 2 gewaschen. Für den Restriktionsverdau wurde der Ansatz in 370 µl $1 \times$ NEBuffer 2 resuspendiert und auf sieben 50 µl Aliquots verteilt:

- 1) 3C-Kontrolle,
- 2-5) Hi-C-Bibliothek,
- 6) Kontrolle des *HindIII*-Verdau,
- 7) Kontrolle des unverdauten Ansatzes.

Zu diesen Ansätzen wurde anschließend 312 µl NEBuffer 2 gegeben. Danach erfolgte eine Inkubation der Proben für 10 Minuten und bei 65 °C mit 38 µl 1 % SDS. Anschließend wurden die Ansätze auf 4 °C abgekühlt und durch die Zugabe von 44 µl 10%iger Triton-100-Lösung das SDS inaktiviert. Der *HindIII*-Verdau der Ansätze 1 bis 6 erfolgte über Nacht rotierend unter Einsatz von 400 U *HindIII*.

Eine besondere Modifikation des Hi-C-Experimentes gegenüber anderen 3C-basierten Experimenten ist die Biotinmarkierung der quervernetzten DNA-Fragmente. Dazu wurden den vier Ansätzen, die für die Generierung der Hi-C-Bibliothek vorgesehen sind (Ansätze 2 bis 5) biotinylierte Nukleotide zugegeben, die die 5'-Überhänge zu *blunt ends* auffüllen (Biotin-14-dCTP). Der Reaktionsansatz wurde für 45 Minuten bei 37 °C inkubiert und enthält hierbei: 1,5 µl 10 mM dATP, 1,5 µl 10 mM dGTP, 1,5 µl 10 mM dTTP, 37,5 µl 0,4 mM Biotin-14-dCTP und 10 µl 5 U/µl Klenow-Fragment. Durch die anschließende Zugabe von 86 µl 10 % SDS und einer Inkubationzeit von 10 Minuten bei 65 °C wurden bei allen sieben Ansätzen die Enzyme inaktiviert.

Die nachfolgende *blunt end*-Ligation der biotinmarkierten Fragmentenden erfolgte mit den Hi-C-Bibliotheks-Ansätzen (Ansätze 2 bis 5) und der 3C-Kontrolle (Ansatz 1) und wurde unter stark verdünnten Bedingungen durchgeführt um eine Ligation der direkt querver-

netzten Fragmente gegenüber weiteren nicht-quervernetzten Fragmenten zu begünstigen. Je 7,61 ml Ligationsmix wurde zu den entsprechenden Ansätzen hinzugegeben und Ansatz 1 mit 1,5 µl bzw. Ansätze 2 bis 5 mit 7,5 µl 2000 U/ml T4-DNA Ligase für 4 h bei 16 °C inkubiert.

Die anschließende Aufhebung der Quervernetzung erfolgte mittels der Zugabe von 25 µl Proteinase K (20 mg/ml) zu den Ansätzen 1 bis 5 bzw. 5 µl Proteinase K (20 mg/ml) zu den Ansätzen 6 bzw. 7 und einer Inkubation über Nacht bei 65 °C. Für Ligationskontrollansätze wurden 480 µl des ligierten Ansatzes 1 und je 120 µl der ligierten Ansätze 2 bis 5 entnommen und mit 25 µl Proteinase K (20 mg/ml) für zwei Stunden und 65 °C einem Protein-Verdau unterzogen.

Die Kontroll-Proben des unverdauten, des *Hind*III-verdauten und der Ligationsansätze wurden währenddessen einer DNA-Aufreinigung unterzogen. Nach Abkühlen der Proben auf 37 °C erfolgte ein RNA-Verdau durch die Zugabe von 10 µl RNase A (10 mg/ml) und einer Inkubation bei 37 °C für 30 Minuten. Zur Entfernung von Proteinresten wurden die Ansätze mit einem Volumen Phenol:Chloroform:Isoamylalkohol (25:24:1) versetzt und damit Phase-Lock-Gel-Tubes beladen. Die Phasenauftrennung erfolgte dann über die Zentrifugation für fünf Minuten bei 14.000 rpm, wobei anschließend die wässrige Phase entnommen und davon 15 µl auf ein 0,8%iges Agarosegel aufgetragen wurde. Die Kontrollen der Ansätze wurden dann elektrophoretisch bei 100 V und für 40 Minuten aufgetrennt.

Die ligierten Ansätze wurden anschließend ebenfalls einer DNA-Aufreinigung unterzogen. Nach Abkühlen der Proben auf 37 °C erfolgte ein RNA-Verdau durch die Zugabe von 50 µl RNase A (10 mg/ml) und einer Inkubation bei 37 °C für 45 Minuten. Zur Entfernung von Proteinresten wurden die Ansätze nach der Überführung in 50 ml Reaktionsgefäße mit einem Volumen Phenol:Chloroform:Isoamylalkohol (25:24:1) versetzt, mehrmals invertiert und bei 3.500 rpm und 21 °C für 10 Minuten abzentrifugiert. Die wässrige Phase wurde abgenommen und mit nukleasefreiem Wasser auf 12 ml aufgefüllt. Dieser Ansatz wurde mit 1/10 Volumen 3 M Natriumacetat pH 5,2 und 2,5-fachen Volumen eiskaltem Ethanol (absolut) versetzt und bei -80 °C für eine Stunde gefällt. Anschließend erfolgte eine Zentrifugation für 30 Minuten bei 4 °C. Die resultierenden Pellets wurden in je 450 µl TE Puffer pH 8.0 resuspendiert und in 1,5 ml Reaktionsgefäße überführt und einer erneuten Extraktion unterzogen. Die DNA-Pellets wurden mit 70 % Ethanol gewaschen, luftgetrocknet und in 60 µl Ampullenwasser resuspendiert unter Inkubation bei 37 °C für 10 Minuten. Nachfolgend wurden die Ansätze 2 bis 5 vereint (= Hi-C-Bibliothek).

2.2. Methoden

Zur Kontrolle wurden je 6 µl einer 1:10 Verdünnung des Ansatzes 1 bzw. der vereinigten Ansätze 2 bis 5 entnommen, elektrophoretisch aufgetrennt und kontrolliert. Die DNA-Konzentration der Ansätze wurde unter Verwendung des Qubit dsDNA BR Assays und des Qubit Fluorometers bestimmt.

Zur Entfernung des Biotins an den Enden der nichtligierten Fragmente wurden sechs 100 µl Reaktionsansätze mit je 5 µg der Hi-C-Bibliothek erstellt: mit 6,5 µl BSA (10 mg/ml), 6,5 µl 10 mM dGTP, 6,5 µl 10 mM dATP, 65 µl 10× NEBuffer 2 und 5 U T4 DNA Polymerase. Diese wurden für zwei Stunden bei 12 °C inkubiert und anschließend durch 10 mM EDTA pH 8.0 gestoppt. Die Ansätze wurden vereinigt und 10 µl einer 1:10 Verdünnung über eine Gelelektrophorese kontrolliert bzw. über das Qubit Fluorometer vermessen.

Die Vorbereitung der Hi-C-Bibliothek zur *paired-end* Sequenzierung erfolgte nach dem Protokoll *Fragment Library Preparation: 5500 Series SOLiD™ Systems User Guide*, wobei der Schritt der Anreicherung der biotinmarkierten Fragmente des Hi-C-Experimentes hinzugefügt wurde. Hierzu wurden 5 µg der Hi-C-Fragmentbibliothek mit Covaris S2 fragmentiert, die Fragmentenden zu *blunt ends* konvertiert, DNA-Fragmente mit einer Länge von 100 bis 300 bp selektioniert und die DNA-Konzentration über den Agilent 2100 Bioanalyser bestimmt. Zur Anreicherung der biotinmarkierten Ligationsfragmente wurden 50 µl MyOne Streptavidin T1 Dynabeads (10 mg/ml) dreimal mit 1× Bindungs- und Waschlösung gewaschen und in 100 µl 2× Bindungs- und Waschlösung gelöst. Der Hi-C-Ansatz wurde auf 100 µl aufgefüllt, zu den Streptavidin-Beads gegeben und für 15 Minuten bei Raumtemperatur und Rotation inkubiert. Es folgten drei Waschschrte mit dem 1× Bindungs- und Waschlösung und die Resuspension in 34 µl Low TE Buffer. Anschließend werden den 3'-Enden dATPs angefügt, Adaptor ligiert und 5 µl der aufgereinigten DNA für 10 Zyklen amplifiziert. Die amplifizierte DNA wurde von den Beads getrennt, die Ansätze vereinigt und mit Agencourt AMPure XP Reagent Beads aufgereinigt und entsprechend der Fragmentgröße selektioniert. Die Fragmentgrößen wurden mittels des Agilent 2100 Bioanalyzers kontrolliert.

Nachfolgend wurde die DNA an P1 Beads gebunden, über Emulsion-PCR klonal amplifiziert und die Beads angereichert. Die Sequenzierung der so prozessierten DNA-Fragmente erfolgte in zwei Lanes eines FlowChips des SOLiD 5500×1 Genetic Analyzers unter Sequenzierung von 75 Nukleotiden (*forward*) bzw. 35 Nukleotiden (*reverse*). Mithilfe der LifeScope Genomic Analysis Software 2.5.1 wurden die Sequenzen im humanen Referenzgenom (hg19) verortet.

2.2.19. Prozessierung und Kontrolle des Seax-Hi-C-Datensatzes

Hi-C ist eine NGS-basierte genomweite Variante der 3C-Methodik um intra- und interchromosomale Interaktionen zu identifizieren. Ligationsprodukte interagierender chromosomaler Abschnitte werden durch den Sequenzierer von beiden Fragmentenden her (für 75bp bzw. 35bp) sequenziert und als Readpaare (*paired-end reads*) ausgegeben. Die Häufigkeit von chromosomalen Interaktionen zwischen zwei Regionen wird anschließend über die Anzahl der vorliegenden Readpaare bestimmt. Die Readpaaranzahl kann jedoch Schwankungen unterlegen sein, die von anderen Faktoren als der Interaktionshäufigkeit bestimmt werden. Faktoren wie ein unterschiedlicher GC-Gehalt oder Dichte an *HindIII*-Schnittstellen in genomischen Regionen, Einfluss der Interaktionsdistanz zwischen Readpaaren und die Zuordenbarkeit der Reads an das Referenzgenom können systematische Fehler in die Hi-C-Daten einbringen und müssen berücksichtigt werden [448]. Die Read-Daten werden daher einer Reihe von Prozessierungs- und Normalisierungsschritten unterzogen. Im ersten Prozessierungsschritt wird die chromosomale Position der Reads dem Referenzgenom zugeordnet (*read mapping* durch LifeScope Genomic Analysis Software 2.5.1, Abb. 2.3). Die daraus resultierenden Qualitätswerte wie der MapQ-Wert (*mapping quality score*) und Distanzinformationen der Readpaare werden in die nachfolgenden Filterschritte einbezogen. Im nächsten Prozessierungsschritt werden die Daten in eine plattformunabhängige Form umgewandelt (Tag-Verzeichnis, HOMER: Hypergeometric Optimization of Motif EnRichment; v4.7 [156]). Gefiltert wurden die bereits anhand des Referenzgenoms positionierten Reads mithilfe der Funktion „analyzeHiC“ des Softwaretools HOMER. Dabei wurden Readpaare entfernt, welche experimentelle Artefakte repräsentieren. So wurden Reads verworfen, die nicht eindeutig einer Position im Referenzgenom zugeordnet werden konnten, von potentiellen PCR-Artefakten oder mit sich selbst ligierten *HindIII*-Fragmenten stammen, einen MapQ-Wert von unter 10 aufweisen oder die nicht in der Nähe einer *HindIII*-Schnittstelle lokalisiert sind (Abb. 2.3).

Das Softwaretool HOMER bietet die Möglichkeit über Kontrolldatensätze die Readverteilung in Abhängigkeit von verschiedenen Parametern zu überprüfen, wobei Abweichungen von erwarteten Verteilungen auf systematische Fehler des Experimentes hinweisen können. Für Readpaar-Artefakte ist u. a. die Ligation von willkürlichen Fragmenten verantwortlich, welche nicht über die *HindIII*-Restriktion generiert wurden. Differenzieren lassen sich diese von zuverlässigen Reads durch den Abstand zu den Restriktionsschnittstellen.

2.2. Methoden

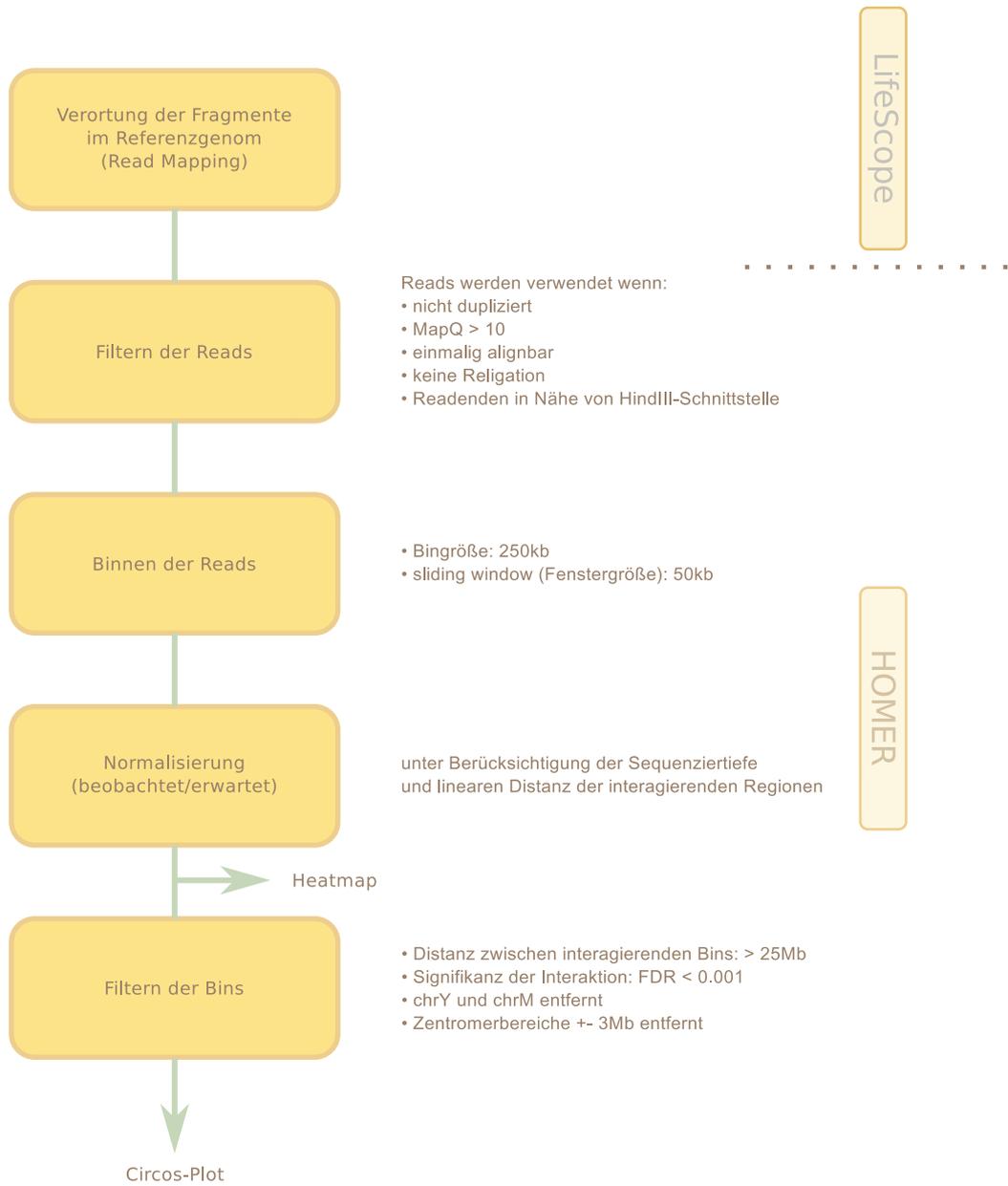


Abbildung 2.3.: Überblick der Prozessierungs- und Normalisierungsschritte des Seax-Hi-C-Experimentes mit HOMER

In Abbildung 2.4 A ist die Anzahl der Reads pro bp ins Verhältnis zu der Distanz zu *Hind*III-Schnittstellen gesetzt. Dabei zeigt sich, dass die gefilterten Reads des Seax-HiC-Experiments vorwiegend innerhalb eines 200 bp Fensters um die Restriktionsschnittstellen lokalisiert sind, und nicht, wie unzuverlässige Reads, davon unabhängig sind. Eine weitere Fehlerquelle bilden GC-reiche Fragmente, welche z. B. im Rahmen der Amplifikation unverhältnismäßig angereichert werden können. Die Gesamtheit der Ligationsfragmente des Seax-HiC-Experimentes (Abb. 2.4 B) weist jedoch, im Vergleich zu dem Gesamtgenom, keinen Anstieg im GC-Gehalt auf. Des Weiteren zeigt die Übersicht über die Verteilung der interagierenden gefilterten Readpaare gemäß ihrer linearen genomischen Distanz (Abb.2.4 C), dass Readpaare mit einer geringeren Interaktionsdistanz gegenüber Readpaaren mit größeren Distanzen untereinander dominieren [229]. Über die Kontrollen konnten somit mögliche Fehlerquellen, wie willkürliche Fragmente oder die vermehrte Anreicherung von GC-reichen Fragmenten, nicht festgestellt werden. Auch die Anteile der Reads in Abhängigkeit zu den Interaktionsdistanzen entspricht dem erwarteten Muster. Im Anschluss

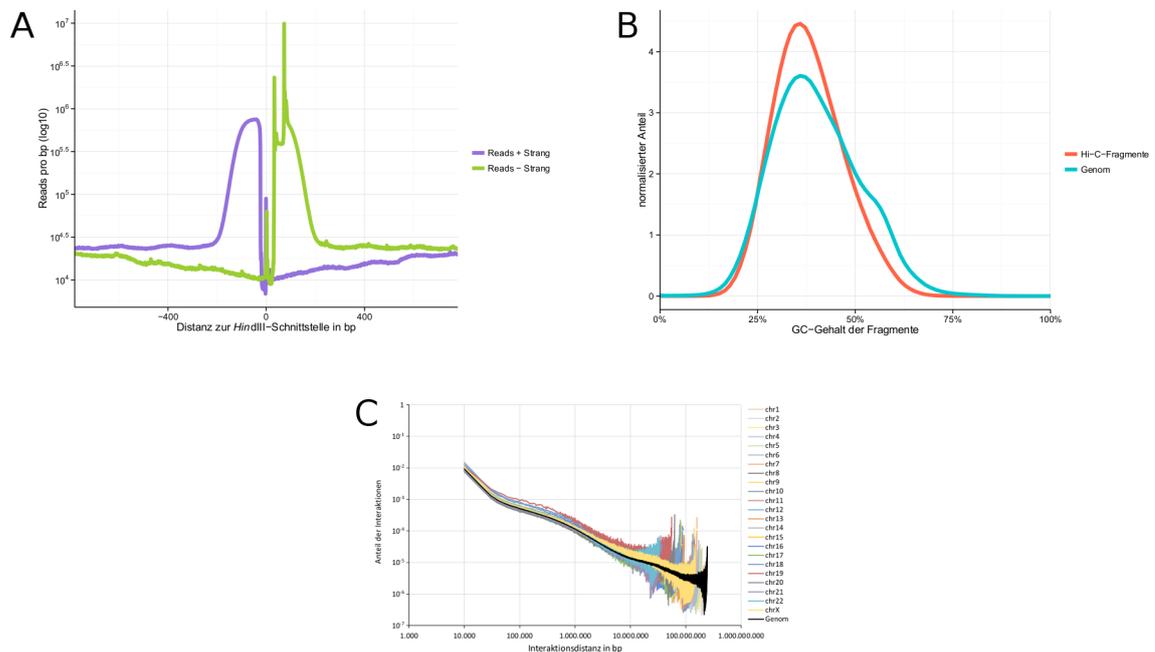


Abbildung 2.4.: Kontrollen des Seax-Hi-C-Experimentes entsprechend der Software HOMER.

A Distanz der Reads zu *Hind*III-Schnittstellen **B** GC-Gehalt der Hi-C-Fragmente im Vergleich zum gesamtgenomischen Anteil **C** Log-Log-Plot der normalisierten relativen Häufigkeiten der Readpaardistanzen gesamtgenomisch und pro Chromosom aufgetragen. Basierend auf Kontrolldatensätze von HOMER.

2.2. Methoden

an die Filterschritte wurden die ca. 91,9 Millionen Readpaare für die Darstellung im Circos-Plot [210] bzw. Heatmaps in 250 kb große Intervalle (Bins) zusammengefasst, die versetzt in 50 kb Abständen positioniert wurden (Binning der Daten nach dem *sliding window-Binning*-Ansatz; für weitere Analysen wurden auch veränderte Bingrößen eingesetzt). Um reale Interaktionen von irrtümlichen Interaktionen zu unterscheiden wurden die Bins anhand eines Referenzmodells hinsichtlich ihrer linearen genomischen Distanz und ihrer Sequenziertiefe normalisiert. Dabei wird davon ausgegangen, dass jede genomische Region von einer ähnlichen Anzahl von Reads abgedeckt ist. Die Sequenziertiefe einer Region ist jedoch abhängig von Faktoren wie z. B. GC-Gehalt oder die Anzahl der dort lokalisierten *HindIII*-Schnittstellen. Abweichungen in der Readzahl können daher entweder auf systematische Fehler oder auf eine tatsächliche Veränderung der Interaktionshäufigkeit hindeuten. Aufgrund dessen wurden die beobachteten Kontakthäufigkeiten (pro Bin) bezüglich der Sequenziertiefe und linearer genomischer Distanz anhand von erwarteten Werten normalisiert. Die erwarteten Werte werden hierbei von einem Referenzmodell bezogen. Dieses Referenzmodell wird dabei für jedes Hi-C-Experiment neu bestimmt.

Anschließend an die Normalisierungsschritte wurden alle Binpaare mit einem höheren Anteil an gemessenen als erwarteten Interaktionen mithilfe der Homer-Funktion „analyzeHiC“ als signifikante Interaktionen eingestuft. Der Schwellenwert für die Bestimmung von signifikanten Bin-zu-Bin-Interaktionen wurde bei einer FDR von 0,001 festgelegt. Da auch nach dieser Filterung sehr viele interagierende Bins vorliegen, wurden zur besseren Übersicht im Circos-Plot Bins entfernt, welche eine lineare Distanz von weniger als 25 Mb aufweisen und die innerhalb der Zentromerregionen voliegen (± 3 Mb). Somit wurden nur über größere Distanzen wirkende Interaktionen berücksichtigt.

Für die Bestimmung von chromosomalen Rearrangements über erhöhte Interaktionshäufigkeiten wurden 100-kb-Bins generiert und diese über Java Treeview [350] visualisiert. Rearrangements zeigen ein spezielles Interaktionsmuster an den Fusionsregionen, bestehend aus einem fokalen Punkt hoher Interaktionsintensität am Bruchpunkt mit einseitig graduell abfallenden Intensitäten der benachbarten Bins (Abb. 3.17 und Abb. S15 A und C; siehe auch [108]). Durch diesen Verlauf der Intensitäten kann die Orientierung des Rearrangements innerhalb des beteiligten Segmentes bestimmt werden. Dabei gibt es mehrere Kombinationen an Regionen, die stromabwärts (tail) oder stromaufwärts (head) des Bruchpunktes mit dem Partnerchromosom fusioniert sind (Abb. 3.17).

Zur Bestimmung der Fusionsregionen auf *HindIII*-Fragmentebene wurden für die über die

100-kb-Heatmaps identifizierten Bruchpunkte ein Intervall von ± 1 Mb um den Bruchpunkt herum definiert. Die Readinformationen der Tag-Verzeichnisse wurden nun nach Reads selektioniert, die innerhalb der 2-Mb-Fenster der interagierenden Regionen positioniert sind. Die resultierende Readverteilung wurde im UCSC Genome Browser visualisiert. Das *HindIII*-Fragment [342], dass den Übergang von readarmen zu readreichen Regionen beinhaltet wurde als Bruchpunktregion definiert (Abb. 3.18).

Die identifizierten Bruchpunktregionen wurden mit Genloci verglichen und auf potentielle Fusionsgene untersucht. Weiterhin wurden die Fusionsregionen mit dem von [178] ermittelten derivaten Chromosom der(3)t(3;6;17;9) verglichen und auf Übereinstimmungen geprüft. Zur Untersuchung der Bruchpunktregionen auf charakteristische Genom- und Chromatineigenschaften wurden die 2-Mb-Intervalle um die Bruchpunkte über BedTools „intersectBed“ der verwendeten Datensätze isoliert, in 100-kb-Intervalle unterteilt und unter Verwendung von R bzw. der R-Pakete „Reshape2“ und „ggplot2“ als Heatmap visualisiert [438] [437] [328]. Die Datensätze setzten sich aus folgenden öffentlichen Daten zusammen: Segmentale Duplikationen [15] [19], *Alu*-Repeats [377], evolutionäre Bruchpunkte [225], *Fragile sites* [131], G4-Motive [54], *Genomic Disorders* (DECIPHER) [119], Subkompartimente (GM12878) [333], Sub-TADs (GM12878) [333], Lamin-assoziierte Domänen (LADs) [143] und non-B-DNA (*Mirror Repeats*, zDNA, G4-Motive) [54]. Weiterhin wurde die Anzahl der unter Abschnitt 2.2.11 20 untersuchten Sézary-Patienten ermittelt, die mit den 100-kb-Bins überlappende Deletionen bzw. Duplikationen aufweisen. Zusätzlich wurden aCGH und ChIP-chip-Daten der Zelllinie Seax mit denen der Zelllinien HH, Hut78, Myla, Hey, Mewo, SH-SY5Y und SK-BR-3 verglichen. Dabei wurden aus diesen Datensätzen die entsprechenden Seax 2-Mb-Bruchpunktregionen selektiert und in 100-kb-Bins unterteilt. Für jede Zelllinie wurde dann die durchschnittliche Intensität der Bruchpunkte pro Bin berechnet (unter Ausschluss der Bins mit fehlenden Daten), die resultierende Linie geglättet (*smoothed*) und so das allgemeine Modifikations- bzw. Aberrationsmuster um den Rearrangementbruchpunkt unter Verwendung von R und der R-Pakete „Reshape2“ und „ggplot2“ bestimmt [438] [437] [328]. Die einzelnen Rearrangementbruchpunktregionen wurden ebenfalls für die Seax-Zelllinie als Heatmap aufgetragen. Des Weiteren wurden die 2-Mb-Rearrangementbruchpunktregionen mit weiteren Daten verglichen. So wurden Hi-C-Daten der Zelllinie GM12878 [363] entsprechend der Seax-Daten mit HOMER prozessiert (Interaktionsdistanzen von > 25 Mb, Binggröße: 250 kb, *sliding window*: 50 kb, FDR: $< 0,001$) und die signifikanten Interaktionen mittels des Circos-Plots mit den Seax-Rearrangementbruchpunktregionen verglichen. Weiterhin wurden *self chain*-Paare [73, 194, 362], Segmentalen Duplikationen [15] [19] und Daten

2.2. Methoden

zu Chromatininteraktionen verschiedenster Experimente der 4DGenome-Datenbank [406] zwischen den Bruchpunktregionen aufgetragen. Der DNA-Kopienzahlstatus der Zelllinie Seax wurde ebenfalls in diese Darstellung integriert und berücksichtigt alle chromosomalen Regionen mit den assoziierten CBS-Werten $< -0,2$ für Deletionen bzw. $> 0,2$ für Duplikationen. Gene in den Bruchpunktregionen wurden mit Genlisten zur GO-Annotation „Lymphozyt“ (CL:0000542, 163 Gene nach der Ausgabe der Genloci durch den UCSC Table Browser) bzw. Gene der Transkriptionsfaktoren der T-Zell-Entwicklung (*ID3*, *GFI1*, *RORC*, *MYCN*, *BCL11A*, *MEIS1*, *GATA2*, *HES1*, *LEF1*, *KIT*, *MEF2C*, *TCF7*, *MYB*, *HOXA9*, *IKZF1*, *KCNH2*, *AHR*, *GFI1B*, *NOTCH1*, *GATA3*, *HHEX*, *LMO2*, *CD44*, *SPI1*, *CD3E*, *ETS1*, *CD4*, *FLT3*, *BCL11B*, *TCF12*, *E2A*, *LYL1*, *NOTCH3*, *SPIB*, *RUNX1*, *ERG*, *ETS2*, *C21orf33*) verglichen. Zum Abschluss wurde die 19p13.3-Region der Seax-Hi-C-Daten mittels visPIG [354] visualisiert. In dieser Region wurden Seax-TADs aufgetragen, deren Ermittlung durch die intrachromosomalen Interaktionen der Zelllinie Seax unter Verwendung der Software TADBit [365] mit einer Auflösung von 50 kb erfolgte. Die resultierenden TAD-Intervalle wurden mit den Sub-TAD-Intervallen der Zelllinien GM12878, HeLa, HMEC, HUVEC, IMR90, K562, KBM7 und NHEK aus [333] und den 19p13.3-Deletions-Intervallen der Sézary-Patienten (aus Abschnitt 2.2.11) verglichen. Zur Untersuchung ob E2A-deregulierte Gene aus Abschnitt 2.2.17 innerhalb der Regionen von signifikanten intra- und interchromosomalen Interaktionen der Zelllinie Seax angereichert sind wurde eine Randomisierung von 222 Genen (=Anzahl der durch E2A deregulierten Gene nach *E2A*-Rekonstitution in Seax) durchgeführt. Dazu wurde diese Anzahl der Gene aus dem UCSC Table Browser *canonical genes*-Datensatz $10.000\times$ zufällig ausgewählt (unter Ausschluss des chrY) und die Anzahl der Gene bestimmt, welche in Regionen inter- und intrachromosomaler Interaktionen lokalisiert sind. Der *p*-Wert wurde nach der Empfehlung für Monte-Carlo-Simulationen von Smyth et al. [318] berechnet (siehe Abschnitt 2.2.1).

2.2.20. Aufschlüsselung der Zusammenarbeit

Die Bearbeitung des Abschnitts 3.1 erfolgte insbesondere in Kooperation mit Grit Ebert (MPIMG) und zusätzlich den Ko-Autoren von [99]. Eine genaue Trennung zwischen meinem Anteil und dem Anteil von Grit Ebert ist nicht immer möglich, da die Analysen gemeinsam bearbeitet wurden. Der Abschnitt 3.2 wurde in Zusammenarbeit mit Dr. Markus Möbs (Charité Berlin), PD Dr. Chalid Assaf (Charité Berlin), Dr. Stephan Mathas

(MDC Berlin) und weiteren Ko-Autoren aus [389] erstellt.

Experiment/ Analyse /Abbildung	Hauptbearbeiter
ChIP der IMR90-Zellen	GE; AS
Früh-apoptotische DNA-Degradation	GE
Mikroarray-Hybridisierung von der IMR90-ChIP und DNA-Degradation	GE; AS
Filterung der IMR90-Hi-C-Interaktionen auf Binebene	AS; ^a
Filterung der IMR90-Hi-C-Interaktionen auf Readebene	GE; ^a
Bündelung der gefilterten Interaktionen	AS
SD-Verteilung und Resampling	AS
Überlappendes Grundmuster	AS
Bestimmung Bruchpunkte der evolutionären Inversionen	^a
Reversion der evolutionären Inversionen	^a
Syntenie zwischen humanen chr7 und <i>C. jacchus</i> chr2	GE
Anreicherung von SDs, G4-Motiven, <i>Alus</i> in Cj syntenischen Regionen	^a
Korrelationen apoptotisches DNA-Degradations- und H4K8ac-Muster	AS
DNA-Degradations- und H4K8ac-Muster an LAD-Grenzen	AS
Gendichte, Intronlänge, GC-Gehalt und G4-Motive des 7q11-Segmentes	GE, ^a
Profil der Interaktionsdistanzen am 7q11-Segment	^a
Vergleich der humanen und murinen topologischen Domänen	GE
RNA-Expressionprofiling	^a
Abb. 3.1	GE; AS
Abb. 3.2	^a ; GE
Abb. 3.3	GE; AS
Abb. 3.4	^a ; GE
Abb. 3.5	AS
Abb. 3.6	^a ; GE
Abb. 3.7	GE
Abb. S1	^a
Abb. S2	^a ; AS
Abb. S3	AS
Abb. S4	^a
Abb. S5	GE
Abb. S6	AS
Anreicherung der Sézary-Zellen und DNA-Isolierung	^b
DNA-Amplifizierung der Sézary-DNA	AS
Hybridisierung und Auswertung der oligo aCGH	AS
Hybridisierung und Auswertung der BAC aCGH	AS*
FISH-Analysen der Sézary-Zellen	^b
Immunhistochemie der Sézary-Zellen	^b
Transfektion der <i>E2A</i> -Konstrukte	^b
Genexpressionsanalyse nach <i>E2A</i> -Rekonstitution	AS

Fortsetzung auf der nächsten Seite

2.2. Methoden

Experiment/ Analyse /Abbildung	Hauptbearbeiter
Auswertung der Genexpressionsanalyse	AS**
EMSA nach <i>E2A</i> -Rekonstitution	<i>b</i>
Proliferationsanalysen nach <i>E2A</i> -Rekonstitution	<i>b</i>
RT-PCR nach <i>E2A</i> -Rekonstitution	<i>b</i>
Abb. 3.8	AS
Abb. 3.9	AS
Abb. 3.10	AS
Abb. 3.11	AS; <i>b</i>
Abb. 3.12	AS
Abb. 3.13	AS
Abb. 3.14	AS
Abb. S7	AS
Abb. S8	<i>b</i>
Abb. S9	<i>b</i>
Abb. S10	<i>b</i>
Abb. S11	<i>b</i>
Abb. S12	<i>b</i>
Hi-C der Zelllinie Seax	GE
Auswertung der Hi-C der Zelllinie Seax	AS
BAC-Arrays der aCGH-, ChIP- und MeDIP-Experimente	AS*; ^c
Abb. 3.15 bis 3.30	AS
Abb. S13 bis S16	AS

AS, Anne Steininger

GE, Grit Ebert

^a, Ko-Autoren aus [99]

^b, Ko-Autoren aus [389]

^c, in Zusammenarbeit mit Vivien Boldt und Ines Müller

^{*}, Durchführung bereits während der Diplomarbeit

^{**}, eigene Auswertung; in [389] ausgeführt von Dr. K. Köchert (MDC Berlin)

3. Ergebnisse

Strukturelle Chromosomenveränderungen sind nicht zufällig im Genom verteilt, sondern weisen eine Abhängigkeit ihrer Fixierung von regionalen Unterschieden in Selektionsdruck und Sequenzzusammensetzung auf [223] [384] [378] [372] [177] [353] [349] [249]. Neue, auf 3C basierende Techniken, erlauben anhand der Detektion chromosomaler Interaktionen die Untersuchung des Einflusses eines weiteren Faktors: der 3D-Chromosomenkonformation. Obwohl sich in der Literatur bereits Hinweise auf die gegenseitige Abhängigkeit von dreidimensionaler Chromatinorganisation, Genregulation und Genomintegrität finden, wurde dieser Aspekt bislang nur wenig berücksichtigt [455] [129]. Am Beispiel verschiedener struktureller Chromosomenveränderungen, wie evolutionärer Rearrangements des Chromosom 7, der mit WBS assoziierten Deletionsregion und der Aberrationsmuster der Tumorentität des Sézary Syndroms, werden chromosomale Rearrangements im Kontext der 3D-Organisation untersucht.

Die Ergebnisse des Abschnitts 3.1 sind dabei in Zusammenarbeit mit Grit Ebert ermittelt und unter [99] publiziert worden. Abschnitt 3.2 hingegen ist Teil der Publikation [389], welche in Zusammenarbeit mit Dr. Markus Möbs entstand (siehe Material und Methoden).

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

Zur Analyse der 3D-Konformation des Chromosom 7 wurden aus publizierten Hi-C-Daten der Zelllinie IMR90 [92] intrachromosomale Interaktionsmuster des Chromosoms berechnet. Die räumliche Nähe zwischen zwei chromosomalen Regionen wird dabei über die

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

Anzahl der Interaktionen (repräsentiert durch *paired-end* Reads) innerhalb eines Binpaars (= IAs/Bin) definierter Größe von 20 kb ermittelt. Um Bins mit unzuverlässigen Werten aus der Analyse auszuschließen, wurden die Bins nach folgenden Kriterien und den entsprechenden Schwellenwerten gefiltert (siehe Material und Methoden, Abb. 2.2 A):

- 1) Interaktionsanzahl der Binpaare, um Bins mit einer höheren Interaktionsanzahl zu bewahren (> 10 , > 15 bzw. 10-50 IAs/Binpaar)
- 2) Interaktionsdistanz der Binpaare, um über große Distanzen interagierende Binpaare zu bewahren (> 10 bzw. > 25 Mb) und
- 3) Überlappung der Binpaare mit SD-Sequenzen, um einen Einfluss von Alignierschwierigkeiten aufgrund der hohen Sequenzsimilaritäten der SDs auf das Interaktionsmuster auszuschließen (Kategorie 3a - SD-Überlappung erlaubt, Kategorie 3b - SD-Überlappung erlaubt, aber keine paraloge SDs bzw. Kategorie 3c - keine SDs erlaubt).

Die gefilterten Bins wurden anschließend zu Interaktionsbündeln zusammengefasst und die resultierende Anordnung der Bündel im Chromosom, das Interaktionsmuster, mittels der Circos Software visualisiert (Abb. 3.1). Welchen Einfluss die Anwendung unterschiedlicher Schwellenwerte auf die Reproduzierbarkeit des Interaktionsmusters hat, zeigt ein Vergleich der resultierenden Interaktionsbündel. Hierbei wurden die Schwellenwerte jeder Filterkategorie der Binpaarselektion variiert und die daraus folgende Veränderung des Interaktionsmusters anhand von Circos-Plots gegenübergestellt (Abb. 3.1, sechs der 12 generierten Interaktionsbündel datensätze sind unter A-F visualisiert; Tab. S1 und Tab. S2). Dabei zeigt sich, dass, unabhängig von den gewählten Schwellenwerten, ein allgemein übereinstimmendes Dreiecksmuster der über lange Distanzen interagierenden Regionen des Chromosom 7 auftritt. Diese Übereinstimmung lässt sich über eine Intersektion aller generierten Muster verdeutlichen und resultiert in ein überlappendes Grundmuster aller untersuchten Datensätze (Abb. 3.1 H). Auch zeigt die Analyse des Datensatzes nach Entfernung von Bins mit SD-Überlappungen dieses trianguläre Muster der Interaktionen. Ein ähnliches Grundmuster wird erkennbar, wenn ein zusätzlicher Datensatz zur Generierung der Interaktionsbündel verwendet wird, bei dem die SD-spezifische Filterung nicht auf Ebene genomischer Bins erfolgte, sondern vor Beginn der Filterschritte alle nicht eindeutig im Referenzgenom verortbaren Reads entfernt wurden (Abb. 3.1 J). Damit ist das Auftreten dieses Interaktionsmusters aufgrund von Alignierschwierigkeiten oder der Schwellenwertdefinition weniger wahrscheinlich.

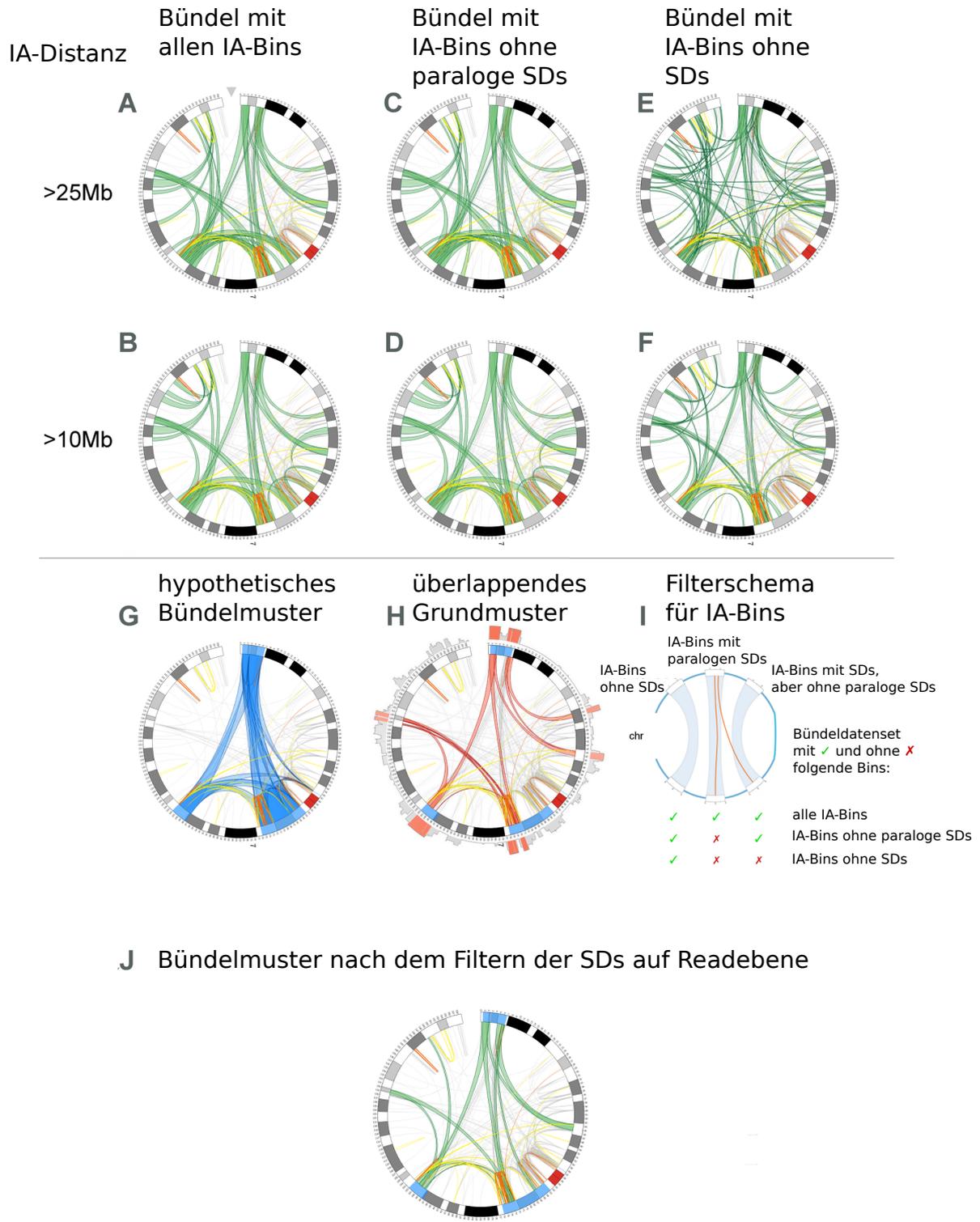
Für die weiteren Analysen wurde als Beispiel das resultierende Interaktionsmuster mit den Schwellenwerten der Filterkategorien von „ > 15 IAs/Binpaar“, „ > 25 Mb Interaktionsdistanz“, „Bins mit und ohne SD-Überlappung (3a)“ und dem Bündelkriterium „mindestens

fünf parallel verlaufende Binpaare in einem 500-kb-Fenster“ verwendet. Bei diesem Datensatz reduziert sich die Binanzahl im Vergleich zu dem ungefilterten Bindatensatz auf nur noch 0,11 % (Tab. S1). Die gefilterten Bins resultieren in 33 Interaktionsbündel, die einzeln mehrere Mb des Chromosom 7 umspannen können und zusammen das Chromosom zu 23,4 % (37,2 Mb) abdecken.

Abbildung 3.1. : (Nächste Seite) Reproduzierbarkeit des triangulären Interaktionsmusters des Chromosom 7 bei der Anwendung unterschiedlicher Schwellenwerte zur Generierung der intrachromosomalen Interaktionsbündel.

Für die Generierung des Interaktionsmusters wurden die Bindaten der intrachromosomalen Interaktionen des Chromosom 7 [92] anhand drei verschiedener Kategorien gefiltert: 1) Anzahl an Interaktionen (IAs) pro Bin, 2) Interaktionsdistanz und 3) SD-Überlappung der interagierenden Binpaare. Durch die Variation der Schwellenwerte wurden insgesamt 12 verschiedene Bin-Datensätze generiert und auf ihre Interaktionsmuster (repräsentiert durch grüne Bänderung) untersucht. Sechs der 12 resultierenden Interaktionsmuster sind unter **A-F** dargestellt, wobei nur diejenigen mit dem Schwellenwert „> 15 IAs/Bin“ der Filterkategorie 1, in Kombination mit den verschiedenen Schwellenwerten der beiden anderen Filteroptionen, SD-Überlappung und Interaktionsdistanz, als Circos-Plot visualisiert sind. Die Reihe des Plots gibt jeweils die Interaktionsdistanz (erste Reihe: > 25 Mb, zweite Reihe: > 10 Mb), die Spalte den Überlapp mit SDs an. **A-B** Bins mit und ohne SDs (Kategorie 3a) **C-D** Entfernung von Bins mit parallel verlaufenden paralogen SDs (Kategorie 3b) **E-F** Entfernung von allen Bins mit SDs (Kategorie 3c); siehe auch Filterschema für SD-überlappende Bins unter (I) oder Abb. 2.2. Darstellung der SDs: < 98 % (grau), 98-99 % (gelb) und > 99 % Sequenzsimilarität (orange). Der Circos-Plot mit den Filteroptionen zur Generierung der Interaktionsbündel der Abb. 3.2 und Abb. S2 ist mit einem grauen Dreieck markiert (A). **G** Hypothetisches Interaktionsmuster, das aus der angenommenen Verbindung der zu *Callithrix* Chromosom 2-homologen Regionen (blau innerhalb des Idiogramms) resultiert (siehe Abschnitt 3.1.2). **H** Überlappendes Grundmuster der 12 verschiedenen Bin-Datensätze. Das äußere Histogramm repräsentiert dabei die Häufigkeit der Integration eines 20-kb-Bins in die 12 Interaktionsmuster. Die rote Schattierung des Histogramms markiert in allen Datensätzen vorkommende interagierende Regionen. Entsprechend veranschaulichen die roten Bänder Intervalle, welche in den Interaktionsbündeln aller 12 Datensätzen vorkommen und zusätzlich die gleichen interagierenden Partnerregionen verbinden. **I** Filterschema für die Entfernung SD-überlappende Bins. **J** Interaktionsmuster nach Entfernung der nicht eindeutig im Referenzgenom verortbaren SD-Sequenzen auf Read-Ebene. Nach dem Verwerfen von Reads, welche eine Überlappung mit SD-Sequenzen aufweisen und nicht eindeutig zuordenbar waren, wurden die oben genannten Filterkriterien angewendet (mit den Schwellenwerten für A). Abb. aus [99].

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7



3.1.1. Anreicherung von SDs in intrachromosomal interagierenden Regionen des Chromosom 7

Die durch das Interaktionsbündel abgedeckten Regionen des Chromosom 7 wurden auf eine Anreicherung von genomischen Eigenschaften und funktionellen Gengruppen untersucht. Ein Vergleich der über eine Distanz von mindestens 25 Mb intrachromosomal interagierenden Regionen mit publizierten Genomeigenschaften und eigenen Acetylierungsdaten des Histons H4K8 bzw. Genexpressionsdaten zeigen, dass diese chromosomale Regionen mit H4K8-Hyperacetylierung, früh-apoptotischer DNA-Degradation und hoher transkriptioneller Aktivität umspannen (Abb. 3.2 und Abb. S1; Abb. S2 Interaktionsmuster mit interagierenden Binpaaren ohne Bündelung). Dies weist, in Übereinstimmung mit der Literatur [229], auf euchromatische Eigenschaften und eine offene, transkriptionell aktive Konformation der interagierenden Regionen hin. Die Grenzen zu eher heterochromatischen Bereichen stimmen dabei mit LAD-Grenzen überein (Abb. S3).

Eine Analyse der in den Interaktionsregionen vorliegenden Gene mit Gene Ontology (GO)-Annotationen ergab jedoch keine Anreicherung gemeinsamer biologischer Funktionen oder Signalwege (analysiert mit PANTHER [274], Daten nicht dargestellt).

Entsprechend der bevorzugten Insertion von SDs in euchromatischen Regionen findet sich eine Anhäufung von intrachromosomalen SD-Sequenzen innerhalb der Regionen der Interaktionsbündel (Abb. 3.2). Ob es sich dabei um eine signifikante Anreicherung von SDs innerhalb der Bündelregionen handelt, wurde mithilfe von Monte-Carlo-Simulationen untersucht. Hierbei wurden mit SD-Sequenzen überlappende Regionen des Chromosom 7 über das gesamte Chromosom 7 $10.000\times$ zufällig verteilt, d. h. mit einer neuen genomischen Lokalisation assoziiert (Randomisierung). Der Anteil dieser zufällig verteilten SD-Regionen an den Interaktionsbündelregionen (in bp) wurde mit dem der tatsächlich überlappenden SDs verglichen (Abb. 3.3 und Tab. 3.1). SDs können entsprechend ihrer Position und Sequenzsimilarität unterschiedlichen Kategorien zugeordnet werden [16]. Die SD-Kategorien beziehen sich hierbei 1) auf welchem Chromosom die SDs lokalisiert sind, 2) ob intra- und/oder interchromosomal interagierende SDs berücksichtigt wurden und 3) welche Sequenzsimilarität die paralogen SDs untereinander aufweisen. Somit wurde ein Einfluss der SD-Kategorien auf die SD-Verteilung berücksichtigt und verschiedene SD-Datensätze als Grundlage der Randomisierung eingesetzt (Tab. 3.1). Die Ergebnisse zeigen, dass alle SD-Datensätze, gegenüber den jeweiligen zufällig verteilten Datensätzen, eine mehr als zweifache Anreicherung von SDs in den intrachromosomalen Interaktions-

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

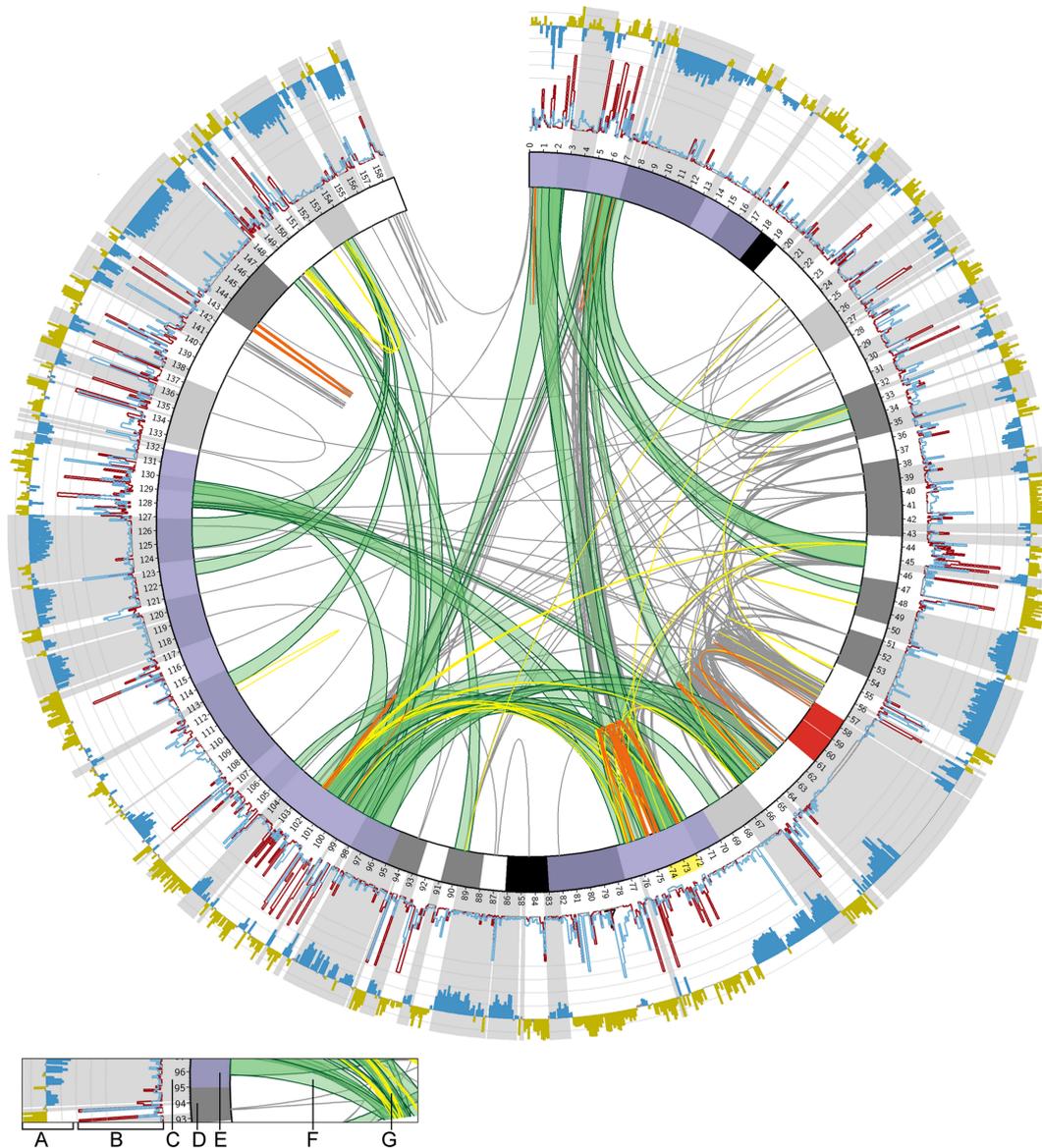


Abbildung 3.2.: Das intrachromosomale Interaktionsmuster des humanen Chromosom 7 überlappt mit Regionen offenen, transkriptionell aktiven Chromatins und SD-Clustern.

Die Regionen intrachromosomaler Interaktionen verbinden Domänen mit hoher transkriptioneller Aktivität und SD-Cluster (Genomische Daten entsprechen, soweit nicht anders vermerkt, der Zelllinie IMR90). **A** Acetylierungsprofil für das Histon H4K8; dunkelgelb: Hyperacetylierung, blau: Hypoacetylierung **B** RNA-Seq Readanzahl pro 100-kb-Bin für kodierende (in rot) und nichtkodierende RNA (in blau) der Zelllinie IMR91L **C** LAD-Grenzen als grau unterlegte Intervalle (Tig3-Zellen; Lungenfibroblasten) **D** Idiogramm des Chromosom 7 mit in gelb markierter Williams-Beuren-Deletionsregion (72-74 Mb; hg18) **E** blau-transparente Schattierung markiert die von der Inversion betroffenen Segmente aus Abb. 3.4 **F** Interaktionsbündel als grün-transparente Bänder und **G** segmentale Duplikationen mit < 98 % (grau), 98-99 % (gelb) und > 99 % Sequenzsimilarität (orange). Abb. aus [99].

bündelregionen aufweisen.

Daher kann von einer signifikanten Anreicherung von SD-Sequenzen innerhalb der intrachromosomal interagierenden Regionen des Chromosom 7 ausgegangen werden.

Tabelle 3.1.: Übersicht der gefilterten SD-Datensätze für die Randomisierungen der SD-Positionen nach Monte-Carlo-Resampling.

Sequenzsim. - Sequenzsimilarität. Tab. aus [99].

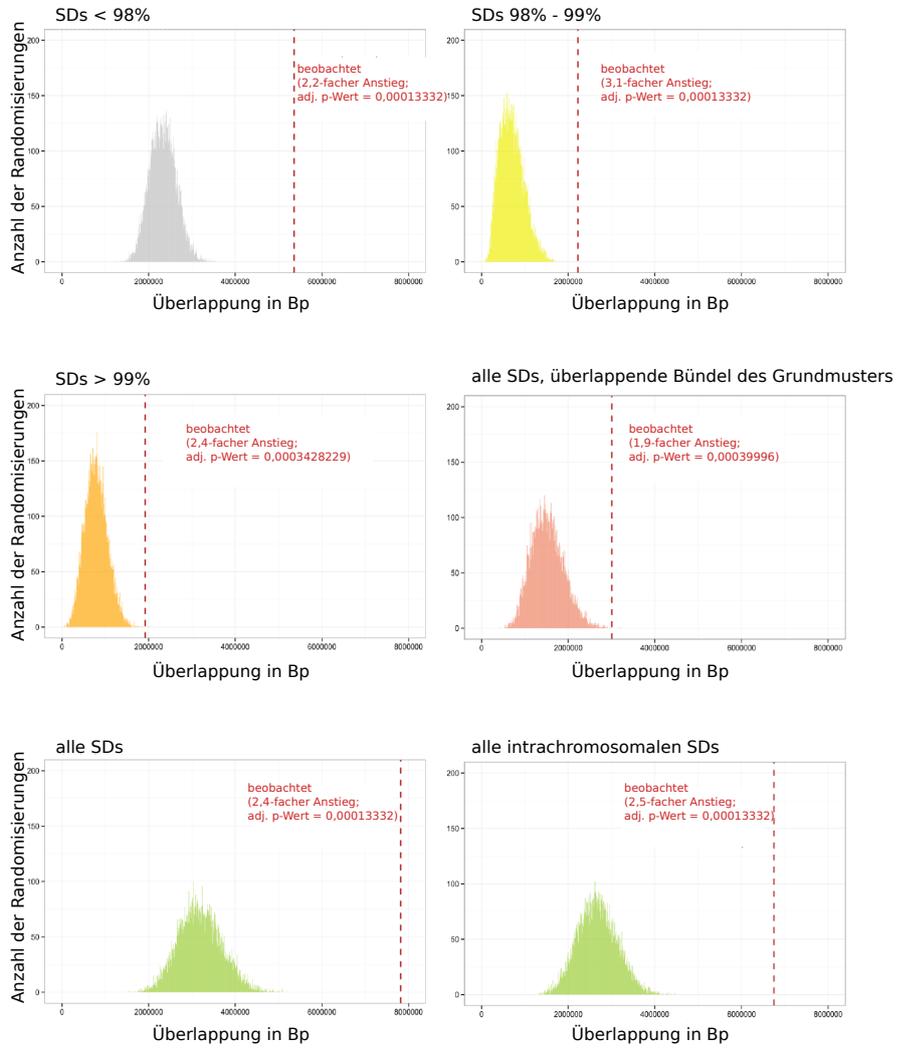
chr	SDs intra/inter	Sequenzsim.	Vergleich mit Interaktionsbündel	adj. <i>p</i> -Wert
chr7	intra + inter	< 98 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332
chr7	intra + inter	98-99 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332
chr7	intra+ inter	> 99 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,0003428229
chr7	inter + inter	> 90 %	Grundmuster aus Abb. 3.1 H	0,00039996
chr7	intra + inter	> 90 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332
chr7	intra	> 90 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332
alle chr	intra + inter	> 90 %	<15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332
alle chr	intra	> 90 %	< 15 IAs/Bin + > 25 Mb IA-Distanz	0,00013332

Abbildung 3.3. : (Nächste Seite) Anreicherung von SD-Sequenzen in Regionen der intrachromosomalen Interaktionsbündel des Chromosom 7.

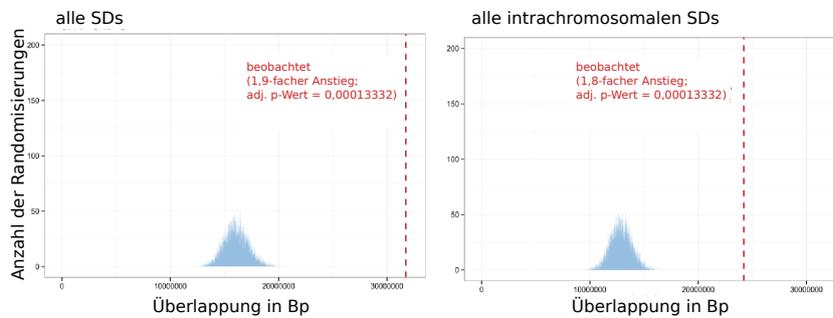
Monte-Carlo-Resampling der SD-Positionen und Bestimmung des Basenpaaranteils der SDs an den Regionen der Interaktionsbündel. Der tatsächlich beobachtete Anteil der SD-Sequenzen an Interaktionsbündelregionen des Chromosom 7 wurde mit Werten aus 10.000 Randomisierungen der SD-Positionen verglichen. Im Histogramm aufgetragen sind dabei die Anzahl der randomisierten Datensätze und deren entsprechende Basenpaarüberlappungen mit den Interaktionsbündeln. Diese Verteilung des Basenpaarüberlapps wurde mit den beobachteten Werten verglichen. Dabei wurden die verwendeten SD-Datensätze in drei Kategorien eingeteilt: alle SDs mit intrachromosomalen und genomweiten Paralogen (alle SDs), alle SDs mit ausschließlich intrachromosomalen Paralogen (alle intrachromosomalen SDs) und die drei Kategorien der Sequenzsimilaritäten einzeln (SDs < 98 %, SDs 98-99 %, SDs > 99 %). Zusätzlich wurden randomisierte SD-Positionen im Vergleich zu dem in Abb. 3.1 H dargestellten Grundmuster von Chromosom 7 und der intrachromosomalen Interaktionsregionen aller Autosomen betrachtet. Der beobachtete Wert des Anteils der SD-Sequenzen an den Interaktionsbündeln wird durch die vertikale gestrichelte Linie markiert. Die Anreicherung der SDs im Datensatz im Vergleich zu den zufälligen Positionen und der zugehörige adjustierte *p*-Wert sind in rot angegeben. Abb. aus [99].

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

chr7



alle chr



3.1.2. Evolutionäre Rearrangements und intrachromosomale Interaktionen des Chromosom 7

Unter Abschnitt 3.1.1 konnte gezeigt werden, dass SDs des Chromosom 7 in Regionen angereichert sind, welche über große Distanzen hinweg interagieren. SDs akkumulieren aber auch u. a. an Synteniebruchpunkten [9] [193] [49]. Während die illegitime homologe Rekombination von SDs in vielen *Genomic Disorders* auch als Ursache für das Auftreten der Rearrangements identifiziert werden konnte [16] [9], ist bei evolutionären chromosomalen Veränderungen der Zusammenhang mit den SD-Sequenzen nicht eindeutig geklärt. Untersuchungen der Einflüsse der 3D-Organisation auf die Verteilung der strukturellen Chromosomenveränderungen können hierbei neue Einblicke in die molekularen Mechanismen chromosomaler Rearrangements geben. Anhand der bioinformatischen Analyse publizierter Hi-C-Daten und eigenen Daten zur Chromatinorganisation der humanen Fibroblastenzelllinie IMR90 werden intrachromosomale Interaktionen des Chromosom 7 auf SD-Verteilung und positionelle Änderung während der Primatenevolution untersucht.

Durch evolutionäre Rearrangements separierte Segmente bleiben über intrachromosomale Interaktionen in der Kernorganisation in räumlicher Nähe zueinander

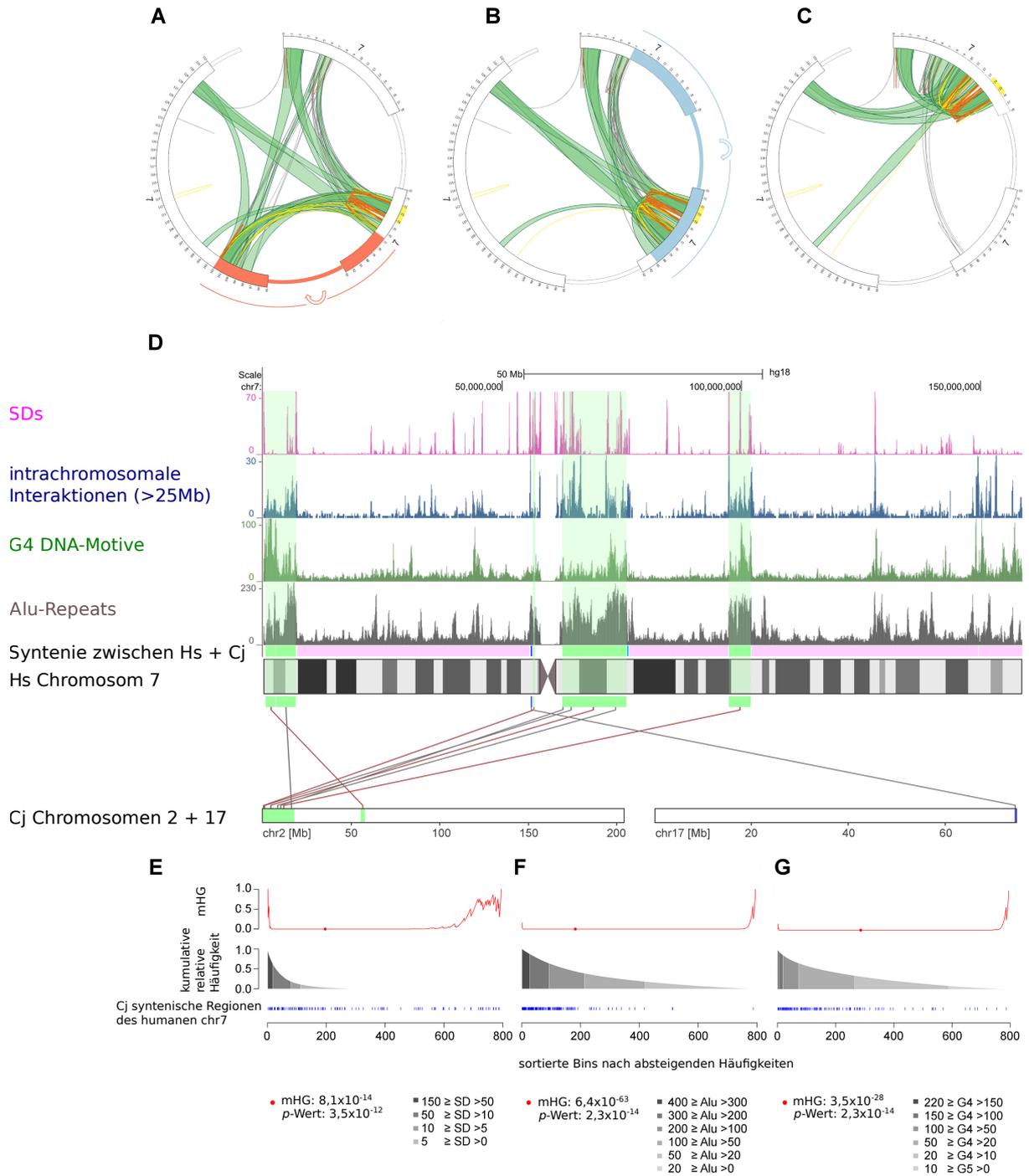
Das Chromosom 7 unterlag während der Evolution einer Reihe von strukturellen Chromosomenveränderungen. Unter anderem kam es während der Primatenevolution zu zwei sukzessiven Inversionen (Abb. 3.4 und Abb. 1.5). Bevor sich die Taxa des Menschen, Schimpansen und Gorillas aufspalteten, erfolgte eine perizentrische Inversion (Abb. 1.5), danach trat eine parazentrische Inversion im letzten gemeinsamen Vorfahren des Menschen und Schimpansen ein. Beide Inversionen sind im Genom des Orang-Utan nicht präsent [287]. Daher ist es möglich die Bruchpunktregionen der Inversionen über ein Alignment des humanen Chromosom 7 mit dem Homolog des Gorilla-Genoms zu bestimmen (Abb. S4, Feinkartierung der Bruchpunkte durchgeführt von Dr. P. Arndt, MPIMG; parazentrische Inversion: 76646908-102118853, perizentrische Inversion: 6875820-80857936; hg18). Eine Rückrechnung dieser Bruchpunkte rekonstruiert die ursprüngliche lineare Anordnung der chromosomalen Segmente vor den Inversionen (Abb. 3.4 A-C). Mittels der Rückrechnung der Segmentanordnung lässt sich auch das hypothetische Interaktionsmuster des Chromosom 7 im letzten gemeinsamen Vorfahren von Mensch und Gorilla *in silico* simulieren

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

(Abb. 3.4 C). Dies wurde berechnet unter der Voraussetzung, dass diese Interaktionsregionen und Interaktionspartner mit denen des humanen Chromosom 7 übereinstimmen und nur ihre Position (durch die Inversionen) verändert haben. Dabei wird deutlich, dass durch die Inversionen die Interaktionen, die über Distanzen von über 25 Mb wirken, in einem vormalig zusammenhängenden telomernahen Segment akkumulieren (Abb. 3.4 C). Das lässt darauf schließen, dass das vormalig linear zusammenhängende chromosomale Segment durch die Inversionen getrennt wurde, jedoch die räumliche Nähe in der 3D-Organisation der Nukleus beibehält.

Die so zusammengeführten Regionen sind syntenisch zu Abschnitten des Chromosom 2 des Weißbüschelaffen *Callithrix jacchus*, einem Neuweltaffen (Cj) (Abb. 3.4 D). Das Schema unter Abb. 3.5 verdeutlicht die veränderte Verteilung dieser zu Cj syntenischen Regionen durch die evolutionären Rearrangements. Fast alle zu Cj Chromosom 2 homologen Segmente des humanen Chromosom 7 weisen eine Übereinstimmung mit intrachromosomal interagierenden Regionen auf, wobei der größte Teil der Interaktionsregionen syntenisch zu einem 17,9 Mb großen Segment des p-Arms des Cj Chromosom 2 ist (Abb. 3.4 D). Ein

Abbildung 3.4. : (Gegenüberliegende Seite) Zu Cj Chromosom 2 syntenische Regionen des humanen Chromosom 7 zeigen eine räumliche Nähe in der Kernorganisation und eine Anreicherung an SD-, *Alu*- und G4-Motiv-Sequenzen. A-C Anordnung des intrachromosomalen Interaktionsmusters und der SDs **A** des humanen Chromosom 7 und nach *in silico* Reversion **B** der parazentrischen und **C** der perizentrischen Inversion. Interaktionsmuster - grüne Bänder, SDs -Farbkodierung gemäß Abb. 3.2. Die an den Inversionen beteiligten Segmente für die para- bzw. perizentrische Inversion sind rot bzw. blau markiert (Abb. erstellt von PD Dr. Reinhard Ullmann). **D** Verteilung der SDs, intrachromosomale Interaktionen über eine Distanz von 25 Mb, G4-Motive und *Alu*-Repeats pro 100-kb-Bin im humanen (*Homo sapiens*; **Hs**) Chromosom 7 und deren Anreicherung in zu Cj syntenischen Regionen (grüne Blöcke). Ein kurzes Segment zeigt ebenfalls eine Syntenie zu Cj Chromosom 17 (blauer Block), während das restliche Chromosom syntenisch zu Cj Chromosom 8 ist (violetter Block)(Abb. von Grit Ebert). **E-F** Anreicherung der SDs, *Alu*-Repeats und G4-Motive in zu Cj Chromosom 2 syntenischen Segmenten des humanen Chromosom 7. Die Anreicherung wurde anhand eines hypergeometrischen Tests nachgewiesen. Alle 200-kb-Bins des Chromosom 7 wurden entsprechend der Anzahl der untersuchten Sequenzen sortiert und der Anteil der Cj syntenischen Bins an den hoch-gerankten Bins untersucht. Die sortierten Bins, die in zu Cj Chromosom 2 syntenischen Segmenten lokalisiert sind, wurden in blau hervorgehoben. Die zu den Bins korrespondierenden Werte des hypergeometrischen Tests entsprechen der roten Linie, wobei das Minimum (**mHG**) als roter Punkt markiert ist (Abb. erstellt von Katrin Tebel). Abb. aus [99].



3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

hypothetisches Interaktionsmuster der zu Cj Chromosom 2 syntenischen Regionen des humanen Chromosom 7 ergibt ein trianguläres Muster, ähnlich den über Hi-C ermittelten intrachromosomalen Interaktionsbündeln (Abb. 3.1 G). Neben der Akkumulation von intrachromosomalen Interaktionen, und damit der räumlichen Nähe zueinander, weisen die zu Cj Chromosom 2 syntenischen Regionen als genomische Eigenschaften eine Anreicherung an SDs (p -Wert = $3,5 \times 10^{-12}$, Abb. 3.4 E), *Alu*- (p -Wert = $2,3 \times 10^{-14}$, Abb. 3.4 F) und G4-Motiv-Sequenzen auf (p -Wert = $2,3 \times 10^{-14}$, Abb. 3.4 G).

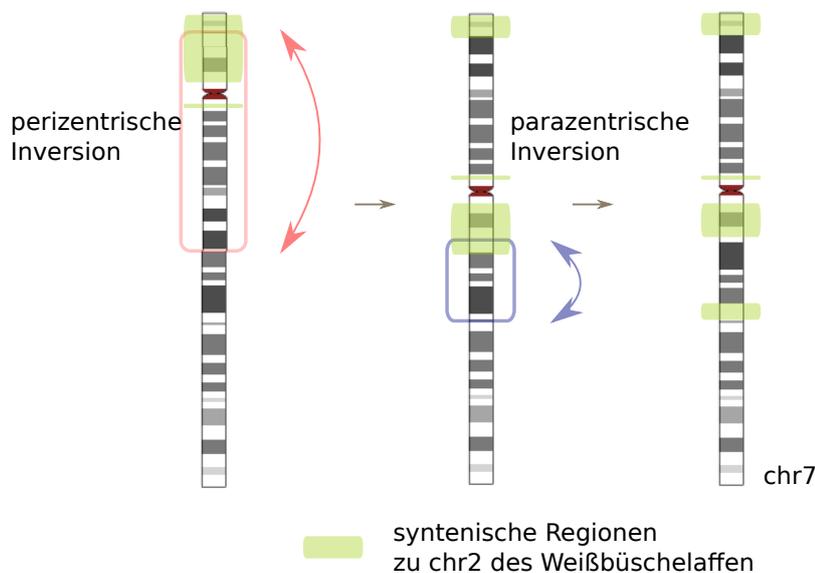


Abbildung 3.5.: Schema der Separation der zu Cj Chromosom 2 syntenischen Region durch die peri- und die nachfolgende parazentrische Inversion des Chromosom 7 während der Primatenevolution.

Eine perizentrische Inversion, gefolgt von einer parazentrischen Inversion während der Primatenevolution führen zu einer Aufspaltung des proximalen mit Cj Chromosom 2 syntenischen Segmentes auf drei verschiedene Regionen des humanen Chromosom 7. Die schematische Reversion der Idiogramme des Chromosom 7 berücksichtigt nur die peri- und parazentrischen Inversionen.

3.1.3. Pathogene Chromosomenveränderungen und intrachromosomale Interaktionen des Chromosom 7

Chromosomenveränderungen der Keimbahnzellen sind zumeist phänotypisch neutral, einige haben aber auch phänotypische Konsequenzen [241]. So sind z. B. pathogene CNVs der 7q11.23-Region an der Ausbildung von *Genomic Disorders* beteiligt, die schwerwiegende Beeinträchtigungen des Patienten zur Folge haben: das Williams-Beuren-Syndrom bzw.

das distale 7q11.23-Deletionssyndroms (siehe Abschnitt 1.2.4). Den *Genomic Disorders* zugrunde liegen instabile Regionen des Genoms, die häufig durch SDs verursacht werden [384]. Da auch die Chromosomenarchitektur Einfluss auf die Genomstabilität ausübt [278], soll hier eine Untersuchung der an WBS-beteiligten SDs in Zusammenhang mit intrachromosomalen Interaktionen erfolgen. Zur Analyse der instabilen 7q11.23-Region, in von den 7q11.23-CNVs nicht betroffenen Zellen der Fibroblastenzelllinie IMR90, wurden die unter Abschnitt 3.1 analysierten publizierten Hi-C-Daten der intrachromosomalen Interaktionen des Chromosom 7 mit verschiedenen Eigenschaften der Chromatinorganisation verglichen.

Die Williams-Beuren-Syndrom-kritische Region verfügt über eine distinkte Chromatinkonformation mit den Eigenschaften einer topologischen Domäne

Das humane Chromosom 7 ist auch Lokalisationsort einer instabilen genomischen Region, die anfällig für pathogene CNVs ist und zu der Ausbildung des Williams-Beuren-Syndroms bzw. des distalen 7q11.23-Deletionssyndroms führen kann. Flankiert werden diese mit den *Genomic Disorders* assoziierten Regionen von drei SD-Clustern, deren nicht-allelische Rekombination Ursache der CNVs ist [271] (Übersicht in Abb. 3.6). Die von den SDs flankierte gesamte Region von ca. 4,8 Mb, im Folgenden als 7q11-Segment benannt (siehe Abb. 3.6), und die angrenzenden Sequenzen werden von den intrachromosomalen Interaktionsbündeln des Chromosom 7 umspannt (Abb. 3.2, die WBS-kritische Region ist am Idiogramm gelb markiert). Damit ist das 7q11-Segment Teil der über eine Distanz von über 25 Mb intrachromosomal interagierenden Regionen.

Eine nähere Betrachtung der Organisation des 7q11-Segmentes mittels eigener und publizierter Daten zeigt chromosomale Chromatineigenschaften, die sich von den angrenzenden Regionen unterscheiden. So kolokalisiert das proximale SD-Cluster mit Änderungen im Chromatinstatus, die auf einen Übergang von Eigenschaften des heterochromatischen, inaktiven Chromatins zu Eigenschaften des euchromatischen, aktiven Chromatins hinweisen. Dies zeigt sich z. B. in einem Übergang von H4K8-hypoacetylierten zu H4K8-hyperacetylierten Regionen am proximalen SD-Cluster, das direkt an die WBS-kritische Region angrenzt (Abb. 3.6 und Abb. S5). Zusätzlich zeigt sich, dass die Regionen stromabwärts des proximalen SD-Clusters ein geringeres Signal für Lamin B1 und damit eine verminderte Wahrscheinlichkeit zur Assoziation mit der Kernmembran [143] und eine früh-apoptotische DNA-Degradation als Chromatineigenschaften aufweisen (Abb. 3.6). Dabei konnte, gemäß der Literatur, eine genomweite signifikante Korrelation

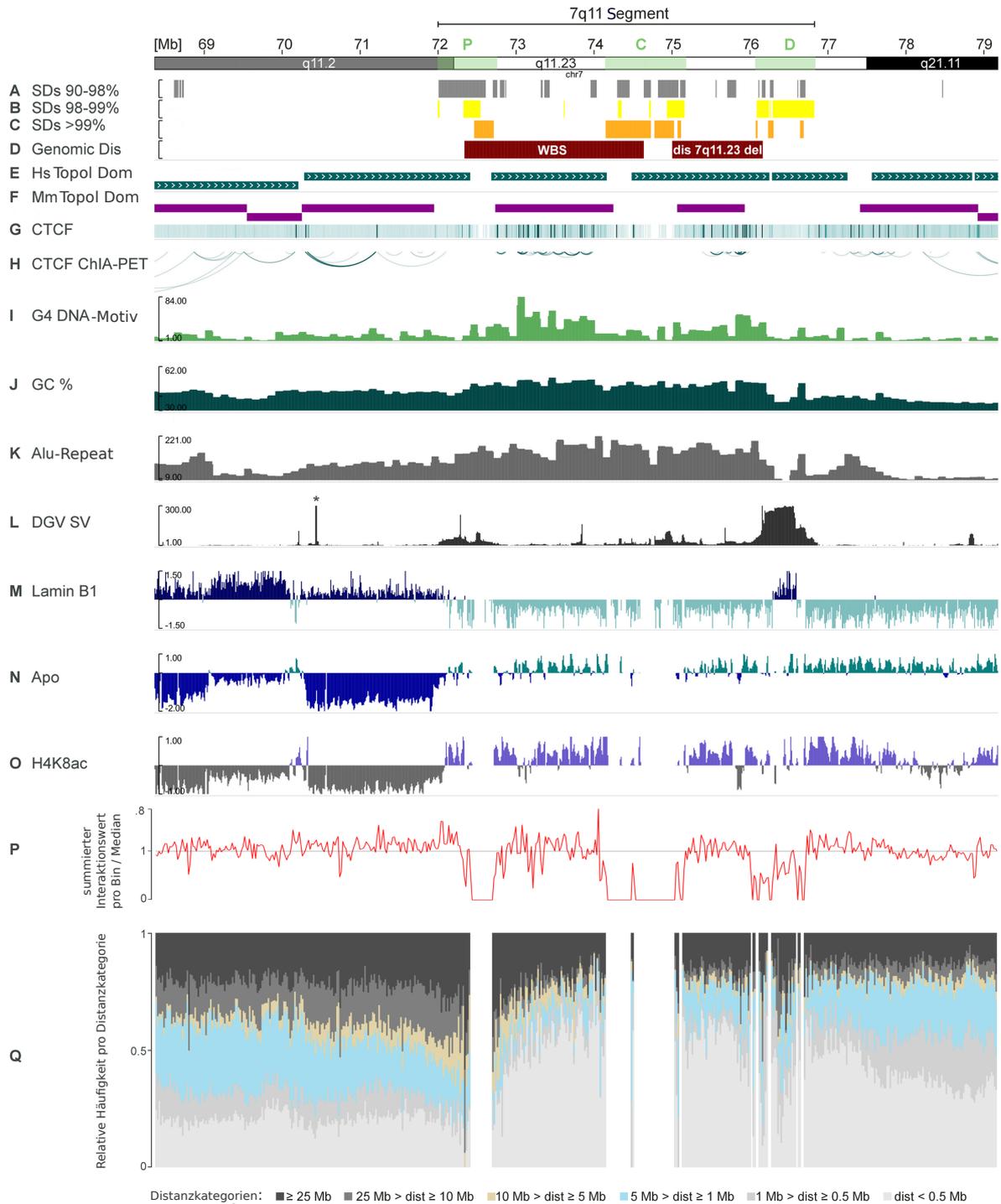
3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

der apoptotischen DNA-Degradation mit folgenden Chromatineigenschaften demonstriert werden: Verteilung der Replikationszeiten ($\rho = 0,65$, p -Wert $< 2,2 \times 10^{-16}$, Abb. S5), H4K8-Acetylierungsmuster (für IMR90-Zellen: $\rho = 0,86$, p -Wert $< 2,2 \times 10^{-16}$, Abb. S6) und des Lamin-B1-Signals der Assoziation chromosomaler Bereiche an die Kernmembran ($\rho = -0,62$, p -Wert $< 2,2 \times 10^{-16}$, siehe Übereinstimmung von LAD-Grenzen mit einer Änderung im H4K8ac- und DNA-Degradationsprofil unter Abb. S3).

Als Sequenzeigenschaften wurden zusätzlich die Gendichte und Introngröße des 7q11-Segmentes untersucht, da die Gendichte, laut Literatur, mit der Chromatinorganisation assoziiert ist [213]. Dabei zeigt sich, dass, im Vergleich zum restlichen Chromosom 7, die Gendichte innerhalb des 7q11-Segmentes signifikant erhöht (23,86 Gene im Vergleich zu durchschnittlich 9,38 Gene pro Mb; geschätzter p -Wert $< 0,0441$ bei 100.000 Randomisierungen im gesamten Chromosom 7) und die Introngröße signifikant verringert ist (3760 bp im Vergleich zu 9827 bp; geschätzter p -Wert $< 0,0453$ bei 100.000 Randomisie-

Abbildung 3.6. : (Gegenüberliegende Seite) Änderung der Chromatinorganisation an der Grenze der WBS-kritischen Region.

Alle Angaben sind in hg19 angegeben und stammen aus der Zelllinie IMR90 (soweit nicht anders beschrieben). **A-C** SD-Positionen: Die entsprechenden Intervalle der innerhalb des 7q11-Segmentes vorliegenden proximalen, zentralen und distalen SD-Cluster (P, Z, D) sind im Idiogramm grün unterlegt; SDs mit Sequenzsimilaritäten von < 98 % (grau), $98-99$ % (gelb) und > 99 % (orange) **D** Häufig deletierte Intervalle des WBS bzw. des distalen 7q11.23-Deletionssyndroms **E** Topologische Domänen nach [92] **F** Topologische Domänen in homologen Segmenten der Maus [92] nach Konvertierung der Daten in hg19 (Topologischen Domänen der Maus, die homolog zu Regionen des humanen distalen 7q11.23-Deletionssyndroms sind, werden durch einen Bruch der Syntenie innerhalb des Intervalls nicht vollständig repräsentiert, siehe Abb. 3.7) **G-H** CTCF-Bindestellen nach ChIA-PET als Heatmap und Bogenansicht (Zelllinie MCF7) **I** Anzahl der G4-Motive pro 100-kb-Bin **J** Durchschnittlicher GC-Gehalt pro 100-kb-Bin **K** Anzahl der *Alu*-Repeats pro 100-kb-Bin **L** Anzahl der Strukturvarianten (nach DGV [170]) pro 100-kb-Bin * Das Maximum bei 1080 CNVs ist nicht dargestellt **M** \log_2 -Intensitäten des LaminB1 DamID-Profiles pro 100-kb-Bin (Zelllinie Tig3, nach [143]) **N** \log_2 -Intensitäten des DNA-Degradationsprofils pro 20-kb-Bin mit früh-apoptotischen (türkis) und nicht-degradierten Regionen (blau) **O** \log_2 -Intensitäten des H4K8-Acetylierungsprofils in 20-kb-Bins, blau: Hyperacetylierung, grau: Hypoacetylierung **P** Summe aller intrachromosomalen Interaktionen pro Bin, normalisiert durch den Median an Interaktionen für das gesamte Chromosom 7 **Q** Häufigkeit (in %) der Interaktionsdistanzen: $< 0,5$ Mb (hellgrau), $0,5-1$ Mb (grau), $1-5$ Mb (hellblau), $5-10$ Mb (hellbraun), $10-25$ Mb (dunkelgrau), ≥ 25 Mb (schwarz). Lücken des Plots entstehen durch Alignmentprobleme der Regionen mit SDs hoher Sequenzsimilaritäten (Abb. erstellt von Grit Ebert und Katrin Tebel und entnommen aus [99]).



3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

rungen im gesamten Chromosom 7). Einen noch deutlicheren Unterschied gibt es, wenn der Vergleich nicht in Bezug auf das gesamte Chromosom 7 erfolgt, sondern auf die direkt angrenzenden Regionen des 7q11-Segmentes beschränkt wird (4,8 Mb stromaufwärts bzw. stromabwärts des 7q11-Segmentes). Hierbei erhöht sich der Unterschied in der Gendichte (stromaufwärts 1,45 Gene pro Mb; p -Wert = $5,829 \times 10^{-14}$; stromabwärts 5,19 Gene pro Mb; p -Wert = $4,661 \times 10^{-7}$, zweiseitiger exakter Test nach Fisher) und Introngröße im Vergleich zu dem 7q11-Segment (stromaufwärts 13772 bp bzw. stromabwärts 9420 bp bei gleicher Genanzahl entsprechend des 7q11-Segmentes, p -Wert $< 2,2 \times 10^{-16}$; zweiseitiger exakter Test nach Fisher). Eine weitere Auffälligkeit in den Sequenzeigenschaften des 7q11-Segmentes ist, dass sich die Grenzen des Segmentes auch in der Verteilung des GC-Gehaltes, der G4-Motive und *Alu*-Repeats widerspiegelt (Abb. 3.6). Besonders stromabwärts der distalen Grenze des 7q11-Segmentes zeigt sich, im Vergleich zu der Region innerhalb des Segmentes, eine Verminderung des GC-Gehaltes, des Anteils an *Alu*-Repeats und der G4-Motive.

Des Weiteren wurde untersucht, ob sich die 3D-Chromatinkonformation des 7q11-Segmentes von den angrenzenden Regionen unterscheidet. Die Analyse der intrachromosomalen Interaktionen des Chromosom 7 ergab dabei eine Reduktion der Interaktionsdistanzen innerhalb der Grenzen des 7q11-Segmentes (Abb. 3.6 Q). Für die Berechnung wurden die den Bins zugrundeliegenden Readpaare entsprechend ihren Interaktionsdistanzen in sechs Distanzkategorien eingeordnet: $< 0,5$ Mb, 0,5-1 Mb, 1-5 Mb, 5-10 Mb, 10-25 Mb und > 25 Mb. Der relative Anteil der einzelnen Distanzkategorien an der Gesamtinteraktionsanzahl pro 20-kb-Bin gibt Aufschluss über die hinsichtlich der Interaktionsdistanz bevorzugte Art der Interaktion des 7q11-Segmentes (Abb. 3.6). Dabei reduziert sich die Anzahl der über eine Distanz von 0,5-5 Mb und 10-25 Mb agierenden Interaktionen innerhalb des 7q11-Segmentes, wohingegen Interaktionen mit einer geringeren Distanz zwischen den Bins ($< 0,5$ Mb) zunehmen. Die Gesamtanzahl der Interaktionen pro Bin zeigt hingegen keine Reduktion innerhalb des 7q11-Segmentes (Abb. 3.6 P). Stromabwärts des distalen SD-Clusters sind die Veränderungen in der Chromatinstruktur nicht so prägnant, jedoch nimmt die Anzahl der über längere Distanzen wirkenden Interaktionen wieder zu (besonders bei 0,5-5 Mb). Damit zeigt das 7q11-Segment, welches die Regionen des WBS und des 7q11.23-Deletionsyndroms umspannt, einen höheren Anteil an Interaktionen mit Distanzen von weniger als 0,5 Mb, im Vergleich zu den angrenzenden Regionen, und bestätigt die spezielle Chromatinkonformation dieser Region. Die bevorzugte chromosomale Interaktion über kurze Distanzen innerhalb des 7q11-Segmentes weist darauf hin, dass die Regionen innerhalb des Segmentes bevorzugt innerhalb der eigenen Grenzen interagieren,

ähnlich einer topologischen Domäne. Somit könnte die Deletion dieses Segmentes die Reduktion einer gesamten topologischen Domäne zur Folge haben.

Publizierte Daten zu den Grenzen der topologischen Domänen beruhen auf der Auswertung von Hi-C-Daten. Diese weisen jedoch an den SD-Loci des 7q11-Segmentes, nahe des *FKBP6*- bzw. *WBSCR16*-Gens, Lücken auf (Abb. 3.6 P bzw. Q und Abb. 3.7 oberere Heatmap). Ursache dafür sind Alignmentprobleme bei der Prozessierung der Daten, die durch die hohen Sequenzsimilaritäten der SD-Sequenzen auftreten können. Diese Lücken können dabei zu einer inkorrekten Definition der topologischen Domänen und deren Grenzen führen. Da TAD-Grenzen im Allgemeinen evolutionär konserviert sind [92], bietet der Vergleich zu Hi-C-Daten der homologen Region der Maus eine Möglichkeit zur Betrachtung der intrachromosomalen Interaktionen und der topologischen Domänen ohne das Vorhandensein der SD-Sequenzen. Auch die zu dem 7q11-Segment homologe Region der Maus weist keine großen SD-Cluster auf. Die Hi-C-Daten der Maus bestätigen das Vorliegen von in sich abgegrenzten topologischen Domänen auch ohne das Vorliegen von flankierenden SD-Clustern (Abb. 3.7 untere Heatmap und blaue Balken). Tatsächlich sind die Grenzen der topologischen Domänen der Maus in dieser Region, wie auch in dem homologen humanen Abschnitt, nahe des *Wbscr16*- und *Fkbp6*-Gens (bzw. *WBSCR16* und *FKBP6*) lokalisiert. Dies weist darauf hin, dass die SD-Blöcke im Laufe der Evolution an den zur Maus homologen TAD-Grenzen des 7q11-Segmentes inserierten. Weiterhin kann man davon ausgehen, dass im 7q11-Segment tatsächlich von SDs flankierte topologische Domänen vorliegen, unabhängig von möglichen Alignmentproblemen durch die SD-Blöcke.

3.1. Chromosomenveränderungen und intrachromosomale Interaktionsmuster des humanen Chromosom 7

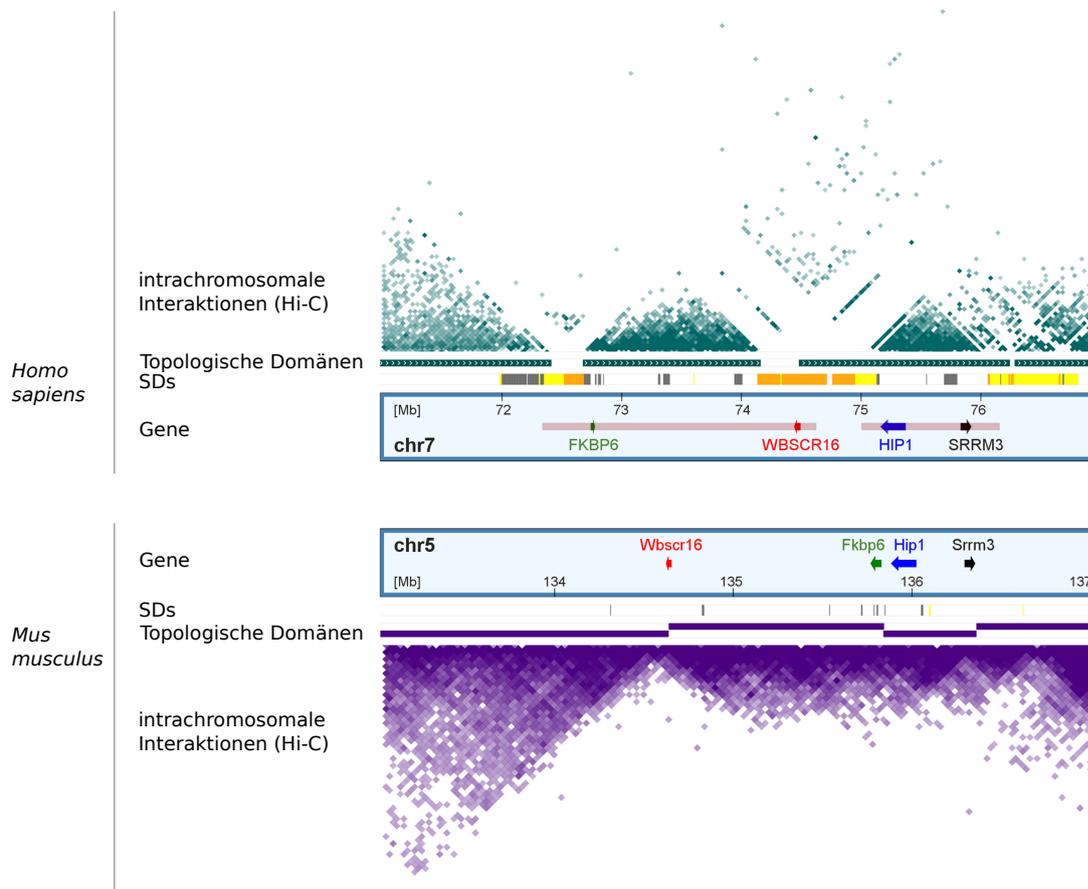


Abbildung 3.7.: SD-Cluster der WBS-kritischen Region inserierten während der Evolution in den homologen Grenzen der topologischen Domänen.

Vergleich der TADs des humanen 7q11-Segmentes mit den homologen Segmenten der Maus auf 5qG2. Die Hi-C-Interaktionen und TADs der Zelllinie IMR90 werden als Heatmap dargestellt bzw. als Balken zur Markierung der Domänengrenzen. Oberer Abschnitt: Intrachromosomale Interaktionen, TAD-Grenzen, SD-Verteilung und Gen-Loci des humanen 7q11-Segmentes. Unterer Abschnitt: Intrachromosomale Interaktionen, TAD-Grenzen, SD-Verteilung und Gen-Loci der zum 7q11-Segment homologen Abschnitte der Maus (5qG2). SD-Cluster (Farbgebung entsprechend Abb. 3.2) kolokalisieren mit Lücken in den humanen Daten der intrachromosomalen Interaktionen. Das entsprechende Intervall des Mausgenoms weist hingegen keine großen SD-Cluster auf. Trotzdem entsprechen die relativen Grenzen der murinen TADs denen des humanen Genoms in dieser Region. Sichtbar wird dies durch den Vergleich der Gen-Loci relativ zu den Domänengrenzen. Die humanen Orthologe der murinen Gene *Wbscr16* (in rot) und *Fkbp6* (in grün) sind ebenfalls an den berechneten Domänengrenzen lokalisiert. Daher ist davon auszugehen, dass sich in dieser Region evolutionär zuerst die Grenzen der TADs etablierten und die SD-Cluster hier inserierten. Die in WBS oder dem 7q11.23-Deletionssyndrom häufig deletierten Intervalle sind innerhalb der humanen Genbox durch hellrote Balken hervorgehoben. Die Region distal des *SRRM3*-Gens inklusive des distalen SD-Clusters ist homolog zu einem anderen murinen Chromosom (Abbildung erstellt von Grit Ebert und aus [99]).

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

Komplexe chromosomale Rearrangements sind ein Kennzeichen vieler Tumorentitäten. Dabei variieren die Aberrationsmuster stark zwischen den verschiedenen beteiligten Zelltypen. In vielen Fällen können diese, besonders bei dem Auftreten rekurrenter chromosomaler Veränderungen, entscheidende prognostische Parameter aufweisen [150] [337]. Die Mechanismen der Musterbildung von chromosomalen Aberrationen ermöglichen daher einen Einblick in die Tumorgenese. Inwieweit die resultierenden Veränderungen rein stochastischer Natur sind oder gerichtet auftreten, ist Gegenstand vieler Untersuchungen. Bisher konnte gezeigt werden, dass hierbei u. a. die Chromatinorganisation, die 3D-Chromosomenstruktur, aber auch selektive Mechanismen der Tumorpathogenese eine entscheidende Rolle spielen [129] [280] [198] [378] [394].

Genome von Tumorzellen mit aggressiven CTCL, wie das Sézary Syndrom, sind ein Beispiel für sehr komplexe Karyotypen mit großen interindividuellen Unterschieden [27] [74]. Zur Untersuchung der Aberrationsmuster auf rekurrente Aberrationen wurden Tumorzellen von 20 Sézary-Patienten auf Kopienzahlveränderungen mittels der array-basierten komparativen genomischen Hybridisierung (aCGH) überprüft und mit der Sézary-Zelllinie Seax verglichen. Anschließend wurden die chromosomalen Rearrangements dieser Zelllinie im Kontext tumorrelevanter Chromosomenveränderungen bzw. inter- und intrachromosomaler Interaktionen betrachtet.

3.2.1. DNA-Kopienzahlveränderungen von Sézary-Tumorzellen

Für die Analyse des Aberrationsmusters wurde das Genom klonaler atypischer T-Zellen von 20 Sézary-Patienten mithilfe der aCGH untersucht. Mit dieser Methode ist es möglich, Kopienzahlveränderungen der Chromosomen, wie Deletionen oder Duplikationen, zu detektieren [379] [320]. Eine Übersicht der chromosomalen Kopienzahlunterschiede aller 20 untersuchten Sézary-Patienten zeigt die Verteilung der betroffenen Segmente und deren Häufigkeiten (Abb. 3.8, unter Verwendung der Arrayplattformen siehe Plattformzuordnung S. 58).

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

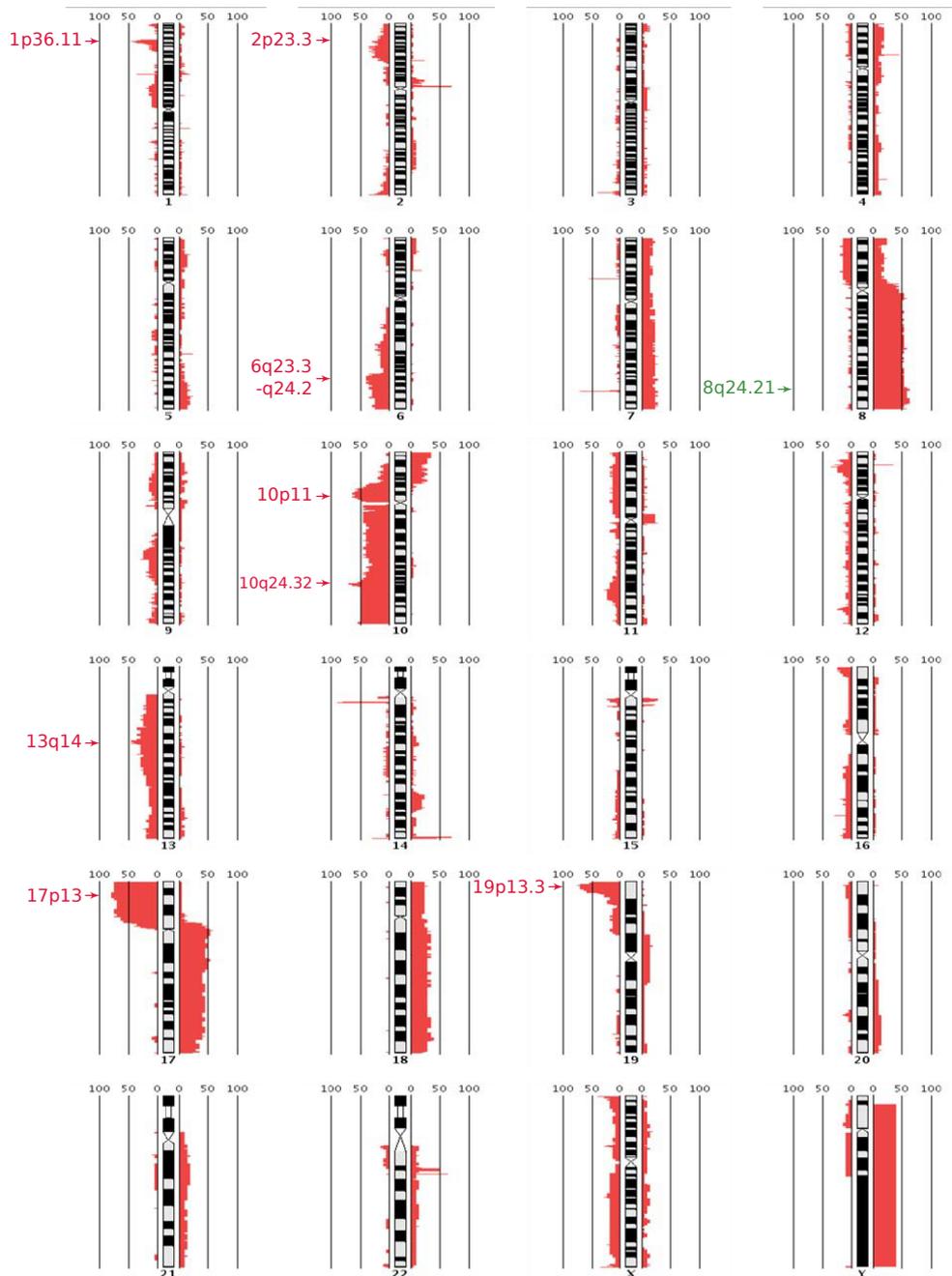


Abbildung 3.8.: Umfangreiche DNA-Kopienzahlveränderungen in Tumorzellproben von 20 Patienten mit Sézary Syndrom.

Die Häufigkeit (in %), mit der Genomabschnitte von Deletionen (links des Idiogrammes) bzw. Duplikationen (rechts des Idiogrammes) betroffen sind, ist entsprechend der chromosomalen Lokalisation aufgetragen. Die Positionen der häufigsten kleinsten gemeinsam veränderten Regionen sind für Deletionen in rot bzw. Duplikationen in grün hervorgehoben. Häufig veränderte Regionen, die mit Telomerbereichen oder Polymorphismen kolokalisieren, z. B. 14q11.2 sowie chrY, sind von der folgenden Analyse ausgeschlossen. Darstellung in hg18 (Abb. aus [389]).

Die Chromosomen der Sézary-Tumorzellen weisen sehr heterogene und komplexe Aberrationsmuster auf. Der Anteil der betroffenen Genomabschnitte an dem Gesamtgenom weicht hierbei mit einer Abdeckung von mindestens 6 % bis über 40 % zwischen den verschiedenen Tumorzellproben stark voneinander ab. Im Durchschnitt sind 17 % des Genoms von DNA-Kopienzahlveränderungen betroffen (Tab. S3). Ob dabei die Segmente eher von Deletionen bzw. von Duplikationen betroffen sind ist individuell verschieden. Im Mittel kann jedoch keine eindeutige Dominanz einer Art der Kopienzahlveränderung an der Abdeckung des Gesamtgenomes aller Tumorzellproben identifiziert werden (Tab. S3). Die größten Segmente chromosomaler Veränderungen finden sich auf den Chromosomen chr8, chr10 und chr17, da diese häufig von Verlusten bzw. Duplikationen ganzer Chromosomenarme betroffen sind. Deletionen gesamter Chromosomenarme sind dabei am häufigsten bei 17p (in 10 von 20 Patienten) zu detektieren, während Duplikationen gesamter Chromosomenarme am häufigsten bei 8q (in 7 von 20 Patienten) zu finden sind. Neben Kopienzahlunterschieden ganzer Chromosomenarme sind auch häufig kleinere Bereiche betroffen. Die chromosomalen Regionen, die die kleinsten gemeinsamen Regionen der Deletionen bzw. Duplikationen aufweisen und die am häufigsten von einer Veränderung der Kopienzahl betroffen sind, finden sich auf den Loci 1p36.11 (in 9 von 20 Patienten), 2p23.3 (in 7 von 20 Patienten), 6q23.3-q24.2 (in 9 von 20 Patienten), 8q24.21 (in 13 von 20 Patienten), 10p11 (in 13 von 20 Patienten), 10q24.32 (in 14 von 20 Patienten), 13q14 (in 8 von 20 Patienten), 17p13 (in 16 von 20 Patienten) und 19p13.3 (in 14 von 20 Patienten) (Abb. 3.9). Die Sézary-Zelllinie Seax wurde ebenfalls auf Kopienzahlveränderungen untersucht und zeigt außer an den Loci 10p11, 10p24.32 und 13q14 ebenfalls eine Überlappung mit den häufigsten kleinsten gemeinsam veränderten Regionen der Patiententumorzellproben (Übersicht des Aberrationsmusters in Seax in Abb. S7).

Eine GO-Analyse mit PANTHER der Gene, die in den häufigsten kleinsten gemeinsam veränderten Regionen und den entsprechenden Grenzregionen (50 kb stromauf- und stromabwärts der Grenzen) lokalisiert sind, ergab keine spezielle Anreicherung von funktionellen Gengruppen bzw. Signalwegen (Ergebnisse nicht dargestellt). Die kleinsten gemeinsam veränderten Regionen wurden ebenfalls auf bekannte, mit Krebs-assoziierte Gene mittels der Datenbank NCG 5.0 Cancer Genes [5] untersucht (Tab. 3.2). Unter diesen Genen befindet sich auch das *E2A*-Gen, welches auf 19p13.3 lokalisiert ist und für einen Transkriptionsfaktor u. a. der Lymphozytenentwicklung kodiert [107]. Die Deletion dieser Region ist die neben der 17p13-Deletion häufigste Veränderung in Tumorzellen der hier untersuchten Sézary-Patienten, wobei die 17p13-Deletion vermehrt durch die Verluste des gesamten 17p gebildet wird und 19p13.3 einer fokalen Aberration entspricht.

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

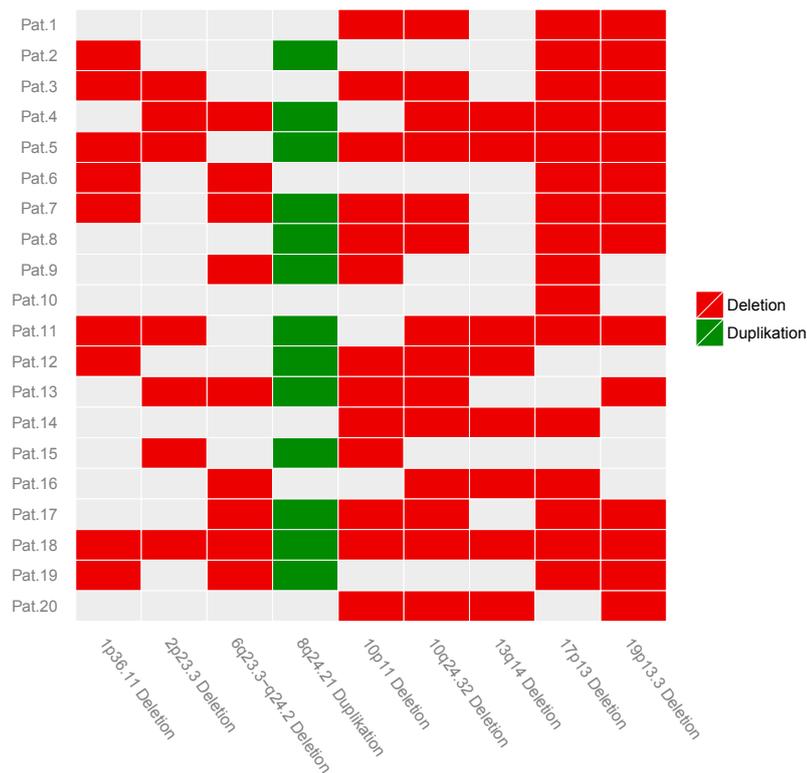


Abbildung 3.9.: Chromosomale Loci der häufigsten kleinsten gemeinsam veränderten Regionen der Deletionen bzw. Duplikationen und deren Zuordnung zu 20 Sézary-Patienten.

Angegebene Positionen enthalten die kleinste gemeinsame Region der chromosomalen Veränderungen. Deletionen sind in rot, Duplikationen in grün markiert.

Die Literaturrecherche ergab, dass der Verlust dieses Gens in Mäusen zur Entwicklung von aggressiven T-Zell-Lymphomen führt und *E2A* eine wichtige Rolle in der murinen Tumorprogression einnimmt [449] [20] [285]. Auch konnte bereits eine Verminderung der E2A-Aktivität auch im Zusammenhang mit humanen Neoplasien nachgewiesen werden [311] [261] [164]. Zu der Zeit der Bearbeitung dieses Projektes konnte in der Literatur jedoch noch kein häufiger Verlust dieses Gens auf genomischer Ebene in Zusammenhang mit humanen T-Zell-lymphoiden Neoplasien belegt werden.

Tabelle 3.2.: Bekannte Krebs-assoziierte Gene innerhalb der kleinsten gemeinsam veränderten Regionen der untersuchten DNA-Kopienzahlveränderungen. Die Gene innerhalb der kleinsten gemeinsam veränderten Regionen wurden auf Übereinstimmungen mit bekannten Krebsgenen der NCG 5.0 Cancer Genes Datenbank [5] untersucht.

Region	bekannte Krebs-assoziierte Gene nach [5]
2p23.3	<i>DNMT3A, NCOA1</i>
6q23.3	<i>TNFAIP3</i>
8q24.21	<i>MYC</i>
10p11	<i>KIF5B</i>
10q24.32	<i>NFKB2, SUFU</i>
17p13	<i>TP53, RABEP1</i>
19p13.3	<i>E2A</i>

3.2.2. Die 19p13.3-Deletion der Sézary-Tumorzellen resultiert in einem Verlust von E2A

Insgesamt zeigen in der aCGH-Analyse 70 % (14 von 20 Patienten) der untersuchten Tumorzellproben eine Deletion auf 19p13.3 mit einer kleinsten gemeinsam deletierten Region, die ca. 1,4 Mb und 54 Gene umspannt (chr19:1368087-2824434, Angabe in hg18; Abb. 3.10 und Tab. 3.3). Um diese Ergebnisse zu verifizieren, wurden zusätzlich FISH-Analysen der 19p13.3-Deletion in Tumorproben durchgeführt. Dabei konnte die heterozygote Deletion des *E2A*-Locus in 8 der 12 untersuchten Sézary-Tumorproben bestätigt werden (Abb. S8 und Tab. S4, durchgeführt von Prof. Dr. E. Schröck). Der Vergleich zwischen beiden Analysemethoden zeigt Unterschiede in vier der Tumorzellproben der Sézary-Patienten, wobei zwei Tumorzellproben eine Deletion des E2A-Locus in der FISH-Analyse zeigen, welche nicht in der aCGH-Analyse zu erkennen ist. Allerdings muss darauf hingewiesen werden, dass die zugrundeliegenden Proben zu unterschiedlichen Terminen entnommen

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

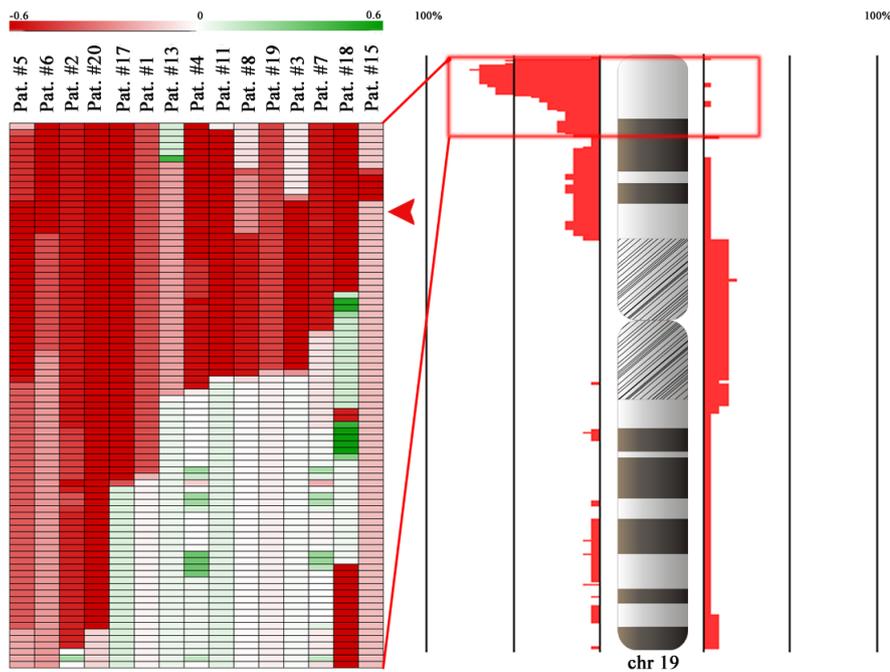


Abbildung 3.10.: Häufigkeit und chromosomale Positionen der 19p13.3-Deletion in Tumorzellen der 20 Sézary-Patienten.

Idiogramm des Chromosom 19 mit Häufigkeit (in %) der Deletionen (links des Idiogrammes) und Duplikationen (rechts des Idiogrammes) der Tumorzellproben von 20 Sézary-Patienten. Die Heatmap links zeigt einen Ausschnitt des chr19 in der Region chr19:200000–8599999 (hg18) mit den individuellen Deletionsintervallen. Die CBS-Werte der aCGH-Analyse werden in 100-kb-Bins repräsentiert und als Intensitäten der Deletionen (rot) und Duplikationen (grün) mit einer maximalen Farbsättigung bei den \log_2 -Verhältnissen 0,6 bzw. -0,6 aufgetragen. Die kleinste gemeinsam deletierte Region der Tumorzellproben umspannt dabei das Gen *E2A*, dessen Genlocus mit einem Pfeil markiert ist. Patient 15 weist eine 19p13.3-Deletion auf, welche jedoch nicht an der *E2A*-Deletion beteiligt ist. Ergebnisse der aCGH-Analyse für Patienten ohne 19p13.3-Deletion sind in der Heatmap nicht dargestellt (Abbildung aus [389]).

Tabelle 3.3.: Chromosomale Bruchpunkte der Deletionen, welche die kleinste gemeinsam deletierte Region auf 19p13.3 umspannen.

Zuordnung zu den untersuchten Sézary-Patienten (Tabelle aus [389]).

Patient Nr.	19p13-Deletion	
	Start (bp)	Stop (bp)
1	19pter	5634354
2	19pter	8313679
3	1368087	4038576
4	19pter	4312210
5	329606	19692685
6	19pter	1881426
7	19pter	3413889
8	902763	4116979
9	–	–
10	–	–
11	265457	4110578
12	–	–
13	794978	4424528
14	–	–
15	–	–
16	–	–
17	19pter	5763726
18	19pter	2824434
19	19pter	4033360
20	19pter	7992388

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

wurden.

Als nächstes wurde die biologische Relevanz des Verlustes des *E2A*-Gens über molekularbiologische Untersuchungen analysiert. Dabei wurden die Auswirkungen der DNA-Kopienzahländerung des *E2A*-Locus auf die Genexpression der beiden *E2A*-Transkriptvarianten *E12* und *E47* ermittelt. Der Vergleich zu CD4⁺-Kontrollzellen zeigt dabei eine Reduzierung von *E12/E47* auf mRNA-Ebene (Abb. S9 A und B, durchgeführt von Dr. S. Mathas und Dr. M. Möbs). Ebenfalls konnte eine fehlende bzw. schwache E2A-Proteinexpression in allen 15 untersuchten Hautbiopsien der Sézary-Patienten nachgewiesen werden (Beispiel unter Abb. S9 C, durchgeführt von Dr. S. Mathas und Dr. M. Möbs).

Aufgrund der Limitierungen von primären Sézary-Tumorzellen in der Zellkultur (z. B. durch die erschwerte Transfektion), wurde für die anschließenden funktionellen Analysen von *E2A* auf eine etablierte Tumorzelllinie zurückgegriffen. Zur Identifizierung eines *E2A*-Verlustes wurden die mRNA-Expression und DNA-Bindeaktivität des Genproduktes in vier T-Zell-Tumorzelllinien (Seax, Myla, Jurkat, Molt-14) untersucht. Ein Vergleich der Zelllinien ergab ausschließlich für die Sézary-Zelllinie Seax eine Reduktion von *E2A*-mRNA und des Proteinlevels (Abb. 3.11 B, durchgeführt von Dr. S. Mathas). Des Weiteren konnte eine verringerte Bindekapazität des E2A-Proteins an das E-Box-DNA-Bindemotiv nachgewiesen werden (Abb. 3.11 C, durchgeführt von Dr. S. Mathas). In Übereinstimmung mit diesen Befunden weist die Seax-Zelllinie in der aCGH- und FISH-Analyse eine heterozygote Deletion der 19p13.3-Region von ca. 3 Mb auf, welche den *E2A*-Locus beinhaltet (Abb. 3.11 A, Tab. Plattformzuordnung unter Abschnitt 2.2.11 und Abb. S8). Aufgrund des heterozygoten *E2A*-Verlustes und der daraus resultierenden Reduktion der E2A-Aktivität eignet sich die Seax-Zelllinie als Modellsystem für eine Analyse des Sézary Syndroms und den Einfluss der 19p13.3-Deletion.

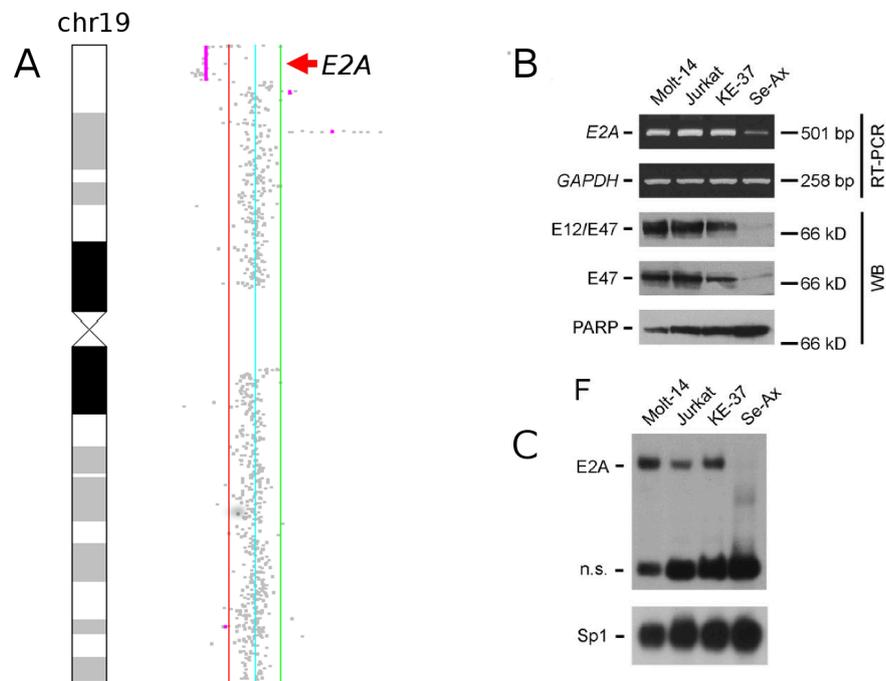


Abbildung 3.11.: *E2A*-Verlust der Sézary-Zelllinie Seax auf DNA-, RNA-, und Proteinebene.

A Die durch aCGH bestätigte Deletion der 19p13.3-Region in Seax-Zellen umspannt auch den *E2A*-Locus (Pfeil). \log_2 -Intensitäten der aCGH-Analyse sind neben dem Idiogramm entsprechend ihrer genomischen Position aufgetragen. Der Schwellenwert von 0,2 für die Bestätigung einer Deletion ist als rote vertikale Linie, der Schwellenwert für Duplikationen als grüne Linie markiert. Der violette Balken entspricht dem durch den CBS-Algorithmus detektierten Aberrationsintervall. **B** Verlust der *E2A*-Expression auf mRNA- und Proteinebene in Seax im Vergleich zu den T-Zell-Tumorzelllinien Molt-14, Jurkat und KE-37. Oberer Abschnitt: RT-qPCR-Analyse der *E2A*-mRNA-Expression und der *GAPDH*-Kontrolle. Unterer Abschnitt: Western-Blot-Analyse der Proteinexpression mit Antikörpern gegen die beiden *E2A*-Transkriptvarianten zusammen (E12/E47), der Transkriptvariante E47 allein bzw. der Kontrolle PARP. **C** EMSA zeigt den Verlust der E-Box-Bindeaktivität des Transkriptionsfaktors *E2A* in Seax verglichen zu Molt-14, Jurkat und KE-37. Als Kontrolle dient der Sp1-Antikörper (n.s.: nichtspezifischer Komplex; Experimente für Abbildung B und C durchgeführt von Dr. S. Mathas und Abb. entnommen aus [389]).

3.2.3. Der Verlust von E2A und die Auswirkungen auf Zell-Proliferation und Genexpression

Um den Effekt einer durch DNA-Verlust reduzierten E2A-Verfügbarkeit auf die Tumorzelle zu analysieren, wurde die Proliferationsgeschwindigkeit der Seax-Zellen nach Rekonstitution des *E2A*-Gens untersucht. Dafür wurden neben einem Kontrollplasmid zwei unterschiedliche *E2A*-Konstrukte unabhängig voneinander in die Seax-Zelllinie eingebracht: das *E47*-Konstrukt, bestehend aus einem mit *Myc*-fusionierten *E47*-Transkript, und das *E47-forced-dimer*-Konstrukt E47-fd, welches zwei über einen flexiblen Linker fusionierte *E47*-Sequenzen vereint [371] [230] (Abb. 3.12, Rekonstitution durchgeführt von Dr. S. Mathas).

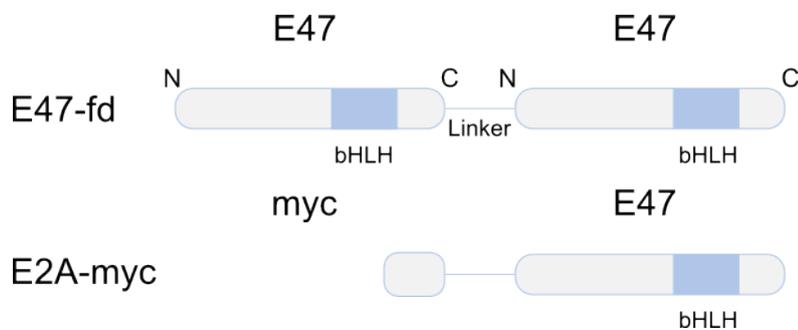


Abbildung 3.12.: Konstrukte zur Rekonstitution der *E2A*-Aktivität in der *E2A*-defizienten Zelllinie Seax.

Schematische Darstellung der verwendeten *E2A*-Konstrukte E47-fd und E2A-myc [371] [230]. E47-fd: Zwei Sequenzen der Transkriptvariante *E47* sind über einen Linker verbunden, was in einem E47-Homodimer resultiert. E2A-myc: Die *E47*-Transkriptvariante ist über einen N-terminalen Linker mit einem *Myc*-Tag verbunden (bHLH: basisches Helix-Loop-Helix-Motiv der DNA-Bindedomäne).

Zur Bestätigung der Rekonstitution der E2A-Aktivität wurde die E2A-Proteinexpression und DNA-Bindeaktivität in den transfizierten Zellen bestimmt (Abb. S10). Dabei zeigen transfizierte Seax-Zellen, gegenüber E2A-positiven Kontrollzelllinien, eine vergleichbare Proteinexpression und E-Box-Bindeaktivität.

Um die Auswirkungen der E2A-Aktivität und dessen biologische Relevanz für die Tumorphagenese zu untersuchen wurde die Änderung in der Zellproliferation 48 Stunden nach Einbringen der Konstrukte analysiert (Abb. S10, Experimente durchgeführt von Dr. S. Mathas). Die mit *E2A* transient transfizierten Seax-Zellen wiesen dabei eine signifikante Reduktion der Zellproliferation im Vergleich zu den Kontrollplasmidexperimenten auf, wobei die Zellzyklusanalyse ein erhöhtes Vorkommen von Zellen der G₀/G₁-Phase

detektierte (Abb. S11, durchgeführt von Dr. M. Möbs).

Des Weiteren wurde der Einfluss der *E2A*-Rekonstitution auf die Genexpression der Seax-Zellen mittels mikroarraybasierter Genexpressionsanalysen untersucht. Die Transfektion der Seax-Zellen mit den *E2A*-Konstrukten E2A-myc und E47-fd (in unabhängigen Duplikaten) führt im Vergleich zu Kontrollexperimenten zu einer allgemein erhöhten Genexpression (Abb. 3.13 A und B). Signifikant deregulierte Gene wurden anhand ihrer assoziierten *p*-Werte (adj. $p < 0,05$) und des \log_2 *Fold Changes* (≥ 1 bzw. ≤ -1 ; entspricht einem *Fold Change* von 2) selektiert. Dabei zeigen sich zwischen den beiden *E2A*-Konstrukten Übereinstimmungen in der signifikanten Deregulation von 222 gemeinsamen Genen, wovon 221 eine Hochregulation aufweisen (Tab. S5 und Abb. 3.13). Zur Kontrolle der Mikroarrayexperimente wurden 7 der deregulierten Gene mittels RT-qPCR auf eine veränderte Genexpression getestet, wobei die Hochregulierung bestätigt wurde (Abb. S12 A, durchgeführt von Dr. S. Mathas). Entsprechend dieser Hochregulation nach Rekonstitution der *E2A*-Aktivität in *E2A*-defizienten Seax-Zellen konnte eine Suppression von vier untersuchten deregulierten Genen in *E2A*-negativen Patiententumorproben, im Vergleich zu CD4-Kontrollen, gezeigt werden (Abb. S12 B und C, durchgeführt von Dr. S. Mathas). Zusätzlich wurden die 222 Gene mittels der TRANSFAC-Option des webbasierten Gen-Annotationstools GATHER [59] auf Transkriptionsfaktorbindestellen untersucht, die 2 kb stromaufwärts bzw. stromabwärts der Transkriptionsstartpunkte (*transcription start site*; TSS) liegen (Tab. 3.4). Dies ergab eine signifikante Anreicherung von Genen, welche in ihren Promotorregionen *E2A*-Bindemotive aufweisen. Anschließend wurde untersucht, ob die signifikant deregulierten Gene gemeinsame Signalwege bzw. molekulare Funktionen einnehmen. Dazu wurde die Genliste mit EnrichR [61] auf ihre GO-Annotationen überprüft (Tab. 3.5 und Tab. 3.6). Dabei zeigt sich u. a. eine Anreicherung von Genen, die mit GO-Annotationen der Zelladhäsions-Regulation und der Ras-Protein-Signaltransduktion assoziiert sind. Anschließend wurden die Proteininteraktionen zwischen den *E2A*-deregulierten Genprodukten mittels der STRING-Datenbank [124] visualisiert (Abb. 3.14). Dabei gibt es zwei Cluster von interagierenden Proteinen (mit mindestens fünf Proteinen), welche mit den gleichen KEGG-Signalwegen assoziiert sind. Das erste Cluster mit LAMB2, VWF, ITGA6, COL5A2 und THBS1 zeigt Übereinstimmungen mit den KEGG-Signalwegen *Extracellular-Matrix-(ECM)-Rezeptor-Interaktion*, *Fokale Adhäsion* und *PI3K-AKT-Signaling*. Das zweite Cluster aus F2RL3, NTSR1, BDKRB1, BDKRB2 und ADRAC zeigt hingegen Übereinstimmungen mit dem KEGG-Signalweg der *neuroaktiven Ligand-Rezeptor-Interaktion*.

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

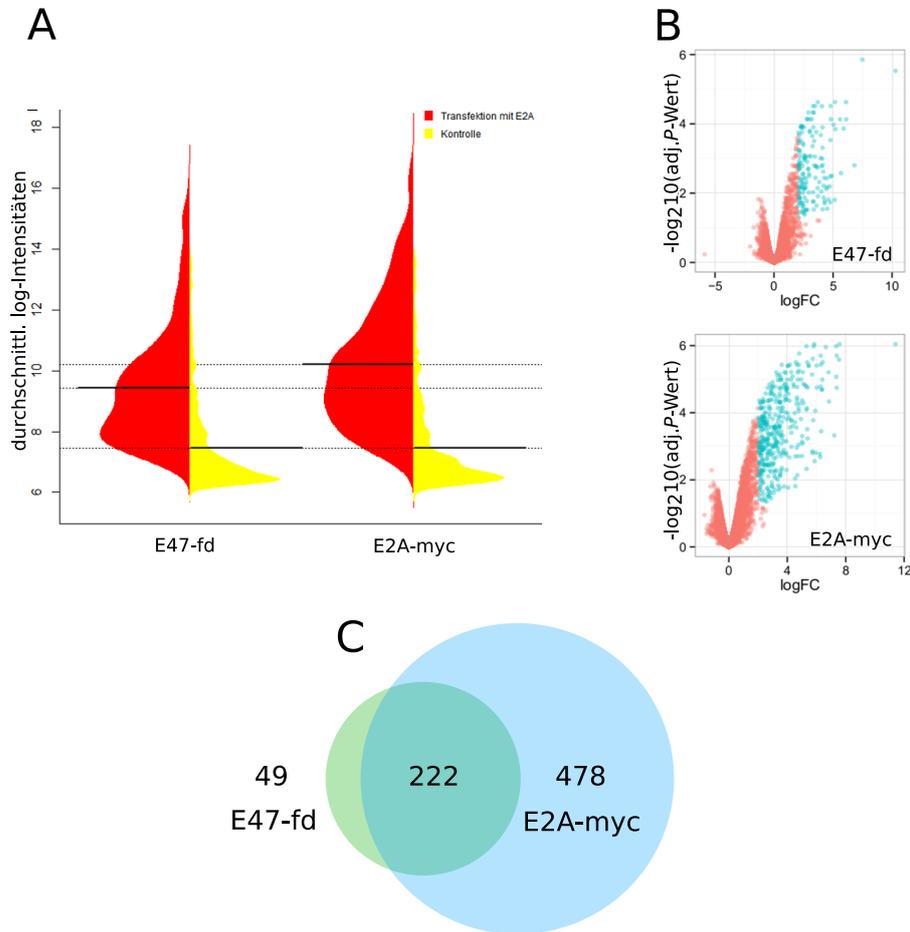


Abbildung 3.13.: Die allgemeine Hochregulation von Genen nach *E2A*-Rekonstitution in Seax-Zellen

A Verteilung der durchschnittlichen Genexpressionsintensitäten nach Rekonstitution von *E2A* durch die Konstrukte *E2A-myc* bzw. *E47-fd* (in rot) im Vergleich zu der Mock-Kontrolle (in gelb). Der Beanplot stellt dabei den Mittelwert (horizontale Linie) und die Dichteverteilung der Intensitäten dar. **B** Volcano Plots der differentiellen Genexpression nach Rekonstitution von *E2A* im Vergleich zur Mock-Kontrolle. Hochregulierte (positiver *Fold Change*) bzw. supprimierte (negativer *Fold Change*) Gene sind entsprechend ihrer Veränderung und der assoziierten Signifikanz aufgetragen. Signifikant deregulierte Gene (adjustierte *p*-Werte < 0,05 und \log_2 *Fold Change* $\log_{FC} > 2$) sind in blau, nicht-deregulierte Gene in rot markiert. **C** Übereinstimmungen der signifikant deregulierten Gene nach *E2A*-Rekonstitution durch die Konstrukte *E2A-myc* bzw. *E47-fd*.

Tabelle 3.4.: Nach *E2A*-Rekonstitution differentiell exprimierte Gene zeigen eine Anreicherung von Promotoren mit *E2A*-Bindemotiven.

Die 222 nach *E2A*-Rekonstitution signifikant differentiell exprimierten Gene wurden mittels der TRANSFAC-Datenbank auf Bindemotive von Transkriptionsfaktoren 2 kb stromabwärts bzw. stromaufwärts der Transkriptionsstartpunkte untersucht. Diese zeigen eine Anreicherung von Promotoren mit *E2A*- (bzw. *E12*/*E47*-) Bindemotiven.

	Transkriptionsfaktor	TRANSFAC ID	Genanzahl	<i>p</i> -Wert	Bayes-Faktor
1	<i>E12</i> (<i>E2A</i>)	V\$E12_Q6	92	<0,001	25
2	<i>E47</i> (<i>E2A</i>)	V\$E47_01	60	<0,001	15
3	GATA3	V\$GATA3_03	54	<0,001	14
4	<i>E2A</i>	V\$E2A_Q2	133	<0,001	13
5	HNF3B	V\$HNF3B_01	74	<0,001	12

Tabelle 3.5.: GO-Analyse der 222 signifikant deregulierten Gene: GO Molekulare Funktion.

Durchgeführt mit EnrichR.

	GO Name	GO ID	Genanzahl	<i>p</i> -Wert	<i>z</i> -Wert	Komb. Wert
1	Integrin-bindend	GO:0005178	8/102	0,00002686	-2,35	10,81
2	Zell-Adhäsionsmolekül-bindend	GO:0050839	9/168	0,0001399	-2,35	8,57
3	Kaliumion-Transmembrantransporteraktivität	GO:0015079	7/138	0,001061	-2,30	5,30
4	Struktureller Bestandteil des Zytoskeletts	GO:0005200	6/100	0,001070	-2,28	5,24
5	Oxidoreduktaseaktivität	GO:0016645	3/28	0,004599	-2,51	4,93
6	Calmodulin-bindend	GO:0005516	7/170	0,003301	-2,32	4,55
7	Wachstumsfaktor-bindend	GO:0019838	6/123	0,002910	-2,31	4,53
8	Serin-Typ-Endopeptidase-Aktivität	GO:0004252	7/167	0,003002	-2,31	4,52
9	Protein-Homodimerisierungsaktivität	GO:0042803	15/653	0,006352	-2,31	4,52
10	G-Protein-gekoppelter Rezeptor-bindend	GO:0001664	8/229	0,004512	-2,30	4,52

Tabelle 3.6.: GO-Analyse der 222 signifikant deregulierten Gene: GO Biologische Funktion.

Durchgeführt mit EnrichR.

	GO Name	GO ID	Genanzahl	<i>p</i> -Wert	<i>z</i> -Wert	Komb. Wert
1	Regulation der Zell-Substrat-Adhäsion	GO:0010810	10/149	0,000004099	-2,37	11,17
2	Regulation der Zelladhäsion	GO:0030155	14/336	0,00001008	-2,46	11,10
3	Regulation der Zell-Matrix-Adhäsion	GO:0001952	7/78	0,00002120	-2,23	9,94
4	Regulation der Ras-Protein-Signaltransduktion	GO:0046578	8/109	0,00002123	-2,19	9,76
5	Stressantwort verursacht durch Salz	GO:0009651	4/25	0,0001822	-2,74	9,55
6	Regulation der neuronalen Differenzierung	GO:0045664	14/405	0,00007300	-2,38	9,17
7	Organisation der extrazellulären Matrix	GO:0030198	13/359	0,00008518	-2,36	9,11
8	Organisation der extrazellulären Struktur	GO:0043062	13/360	0,00008752	-2,36	9,11
9	positive Regulation der Zell-Substrat-Adhäsion	GO:0010811	7/93	0,00006075	-2,31	8,89
10	Regulation des Zellwachstums	GO:0001558	12/322	0,0001242	-2,33	8,40

3.2. Chromosomenveränderungen und chromosomale Interaktionsmuster in den Tumorzellen des Sézary Syndroms

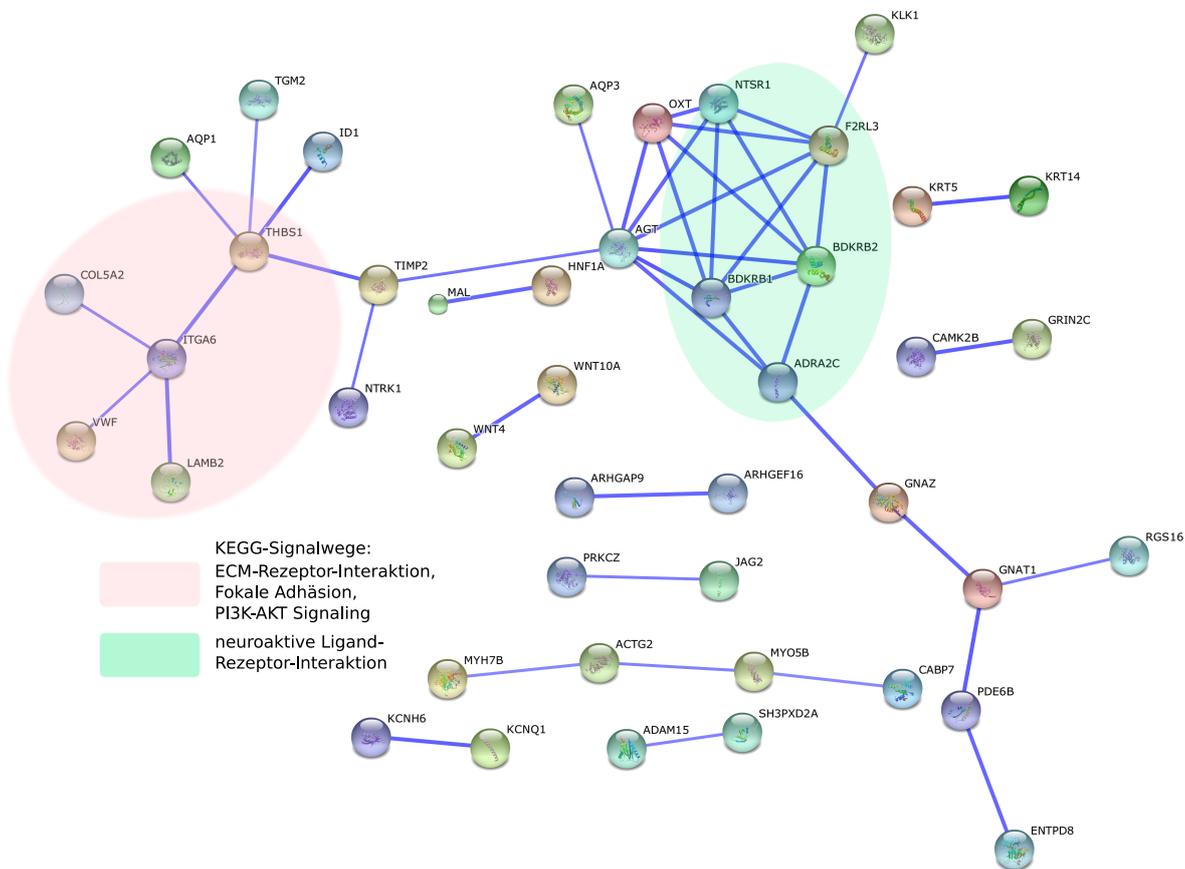


Abbildung 3.14.: STRING-Netzwerk der nach *E2A*-Rekonstitution signifikant deregulierten Genprodukte.

Blaue Linien repräsentieren funktionelle und physische Interaktionen zwischen den Genprodukten. Die Farbintensität der Interaktionslinien entspricht dabei dem Konfidenzwert der Interaktion. Es wurden ausschließlich Interaktionen mit einem Konfidenzwert von $> 0,7$ berücksichtigt (*high confidence*). Durch Interaktionen verbundene Proteine (mit $n \geq 5$), die mit dem gleichen KEGG-Signalweg assoziieren, sind farblich unterlegt. Die Abbildung wurde generiert unter STRING 10.0.

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

Neben Kopienzahlveränderungen liegen auch eine Vielzahl von Translokationen im Genom von Sézary-Zellen vor [74]. Die Detektion von Translokationen ist jedoch durch die Anreicherung von repetitiven Elementen erschwert [312] [1]. Alternativen zu den gebräuchlichen Analysemethoden bieten 3C-basierte Verfahren, welche über die detektierten chromosomalen Interaktionshäufigkeiten Veränderungen in der linearen Anordnung chromosomaler Regionen nachweisen können. Da die Interaktionshäufigkeit von chromosomalen Regionen exponentiell mit der Entfernung abnimmt, kann man daher über erhöhte Interaktionshäufigkeiten zwischen chromosomalen Regionen auch balancierte Rearrangements wie Translokationen oder Inversionen identifizieren [47] [125] [25] [108] [215] [333]. Inwieweit damit komplexe chromosomale Rearrangements analysiert werden können, soll anhand des Beispiels der Sézary-Zelllinie Seax untersucht werden. Dafür wurden die chromosomalen Interaktionen der *E2A*-defizienten Zelllinie Seax mittels Hi-C-Analyse untersucht (durchgeführt von G. Ebert und A. Steininger).

3.3.1. Intra- und interchromosomale Interaktionen der Seax-Zelllinie weisen auf Rearrangementbruchpunkte hin

Interaktionen zwischen chromosomalen Regionen werden durch Binpaarkombinationen repräsentiert (siehe Abschnitt 2.2.1), deren Interaktionshäufigkeiten in zweidimensionalen Matrizen repräsentiert werden. Somit entspricht jede Reihe dem Interaktionsprofil eines Bins mit jedem weiteren Bin desselben Chromosoms (bei Matrizen intrachromosomaler Interaktionen) bzw. weiterer Chromosomen (bei interchromosomalen Interaktionen). Beispiele für intrachromosomale Interaktionsprofile der Chromosomen chr17 bzw. chr6 und deren Normalisierung gibt Abbildung 3.15 wieder. Die Matrizen der nicht normalisierten beobachteten Interaktionshäufigkeiten zeigen hierbei, dass, entsprechend der Literatur, die Interaktionsfrequenz von direkt linear benachbarten Bins erhöht ist (Abb. 3.15). Um Binpaare zu identifizieren, die abweichend von ihrer erwarteten linearen Distanz hohe Interaktionshäufigkeiten aufweisen, werden die beobachteten Interaktionen anhand der zu erwartenden Interaktionswerte korrigiert (siehe Abschnitt 2.2.19).

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

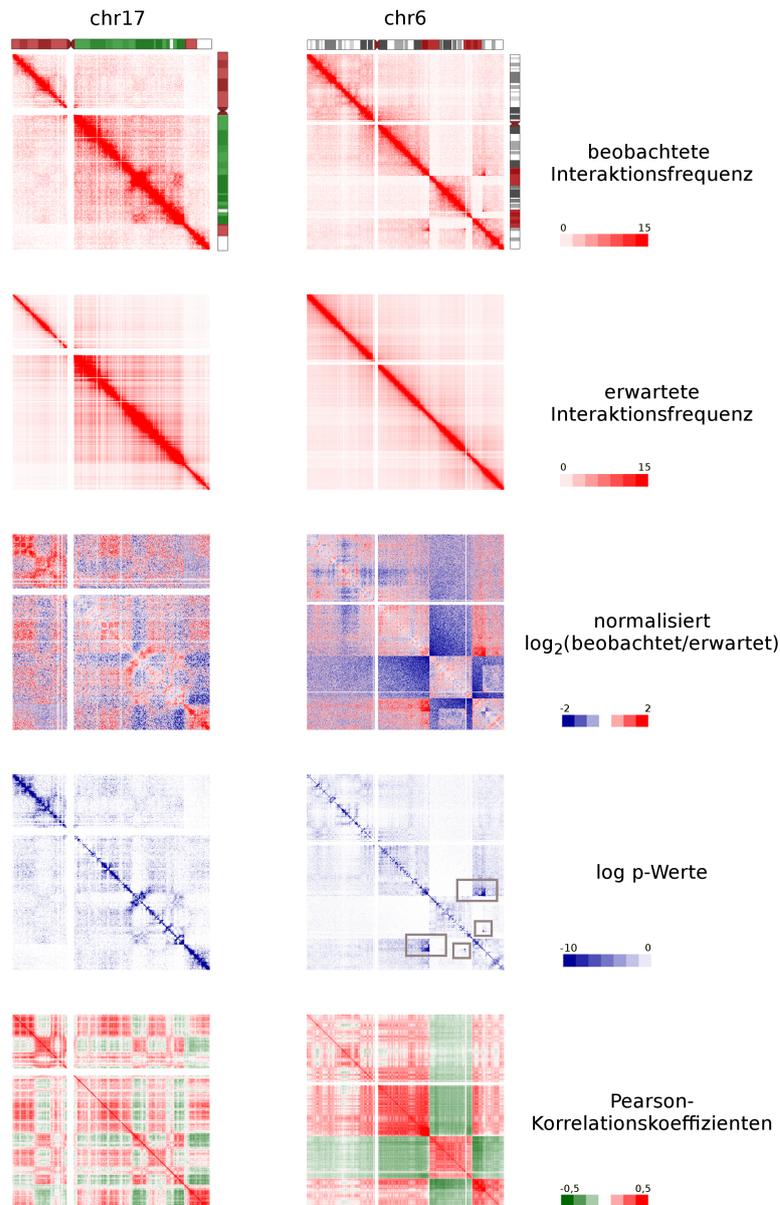


Abbildung 3.15.: Vergleich der Interaktionsmatrizen: Chromosomen 6 und 17. Heatmaps der intrachromosomalen Interaktionsmatrizen der Chromosomen 6 und 17. Reihe 1-2: Heatmaps repräsentieren die beobachteten bzw. erwarteten Werte der Interaktionshäufigkeiten zwischen Bins des gleichen Chromosoms (rot: Intensität entspricht der ansteigenden Interaktionshäufigkeit zwischen Binpaaren). Reihe 3: Heatmaps repräsentieren die normalisierten Werte der Interaktionshäufigkeiten zwischen Bins des gleichen Chromosoms (rot: Binpaar mit häufiger Interaktion; blau: Binpaar mit verminderter Interaktion). Reihe 4: Matrizen der assoziierten $\log p$ -Werte (blau: signifikante Interaktionen; graue Kästen: signifikante Interaktionen zwischen mehrere Mb-entfernten Bins repräsentieren hier intrachromosomale Rearrangements). Reihe 5: Die Pearson-Korrelationskoeffizienten der normalisierten Interaktionshäufigkeiten weisen auf die A- und B-Kompartimentierung der Chromosomen (in nicht intrachromosomal rearrangierten Chromosomen; chr17) bzw. chromosomale Rearrangements (chr6) hin (rot: Binpaare mit korrelierenden Interaktionsprofilen; grün: Binpaare mit nicht-korrelierenden Interaktionsprofilen). Die Binsgröße entspricht 250 kb mit einer Fenstergröße von 50 kb (= Versatz der überlappenden Bins).

Die resultierenden normalisierten und \log_2 -transformierten Werte repräsentieren ein Muster mit geringen Interaktionshäufigkeiten (blau, Abb. 3.15) und erhöhten Interaktionshäufigkeiten (rot, Abb. 3.15) zwischen chromosomalen Regionen.

Chromosomale Rearrangements lassen sich, entsprechend der Literatur, anhand signifikant erhöhter Interaktionshäufigkeiten identifizieren [47] [125] [25] [108] [215] [333]. Das Beispiel der Chromosomen 6 und 17 in Abb. 3.15 stellt einem Chromosom ohne auffällige intrachromosomale Rearrangements (chr17) eines mit intrachromosomalen Rearrangements (chr6) gegenüber. Während man signifikante Interaktionen des Chromosom 17 hauptsächlich bei linear benachbarten Bins nachweisen kann, so finden sich bei Chromosom 6 auch signifikante Interaktionen, welche zwischen mehrere Mb-entfernten Bins interagieren. Diese weisen auf Fusionen innerhalb des Chromosoms hin. Die Rearrangements spiegeln sich auch in der Matrix der Pearson-Korrelationskoeffizienten wider. Mit diesen können Bins, welche ähnliche Interaktionsprofile aufweisen, veranschaulicht werden. Chromosom 17 zeigt hier Interaktionsmuster mit einem der Literatur entsprechenden „Schachbrettmuster“, das die A- und B-Kompartimentierung der Chromosomen repräsentiert (siehe Abschnitt 2.2.1) und bei denen Binpaare der gleichen Kompartimentklasse korrelierende Interaktionsprofile aufweisen (in rot, Abb. 3.15 und Abb. S13; chr17), wohingegen Binpaare unterschiedlicher Kompartimente nicht korrelieren (in grün, Abb. 3.15 und Abb. S13; chr17) [229]. Die intrachromosomalen Rearrangements des Chromosom 6 hingegen zeigen anstatt des „Schachbrettmusters“ der Chromosomenkompartimentierung Brüche im Chromosom, bei denen die resultierenden Segmente sich deutlich voneinander abgrenzen (Abb. 3.15; chr6).

Entsprechend den obigen intrachromosomalen Matrizen werden die Interaktionen zwischen zwei Chromosomen als Interaktionsprofile aller Binpaare der beiden Chromosomen dargestellt. Die Übersicht der Interaktionsmatrizen aller Chromosomen der Zelllinie Seax zeigt auch hier Regionen mit stark erhöhten Interaktionshäufigkeiten, welche Fusionsregionen zwischen verschiedenen Chromosomen repräsentieren. Diese geben Aufschluss über die beteiligten Rearrangementpartnerregionen. Mit Ausnahme der Chromosomen chr14, chr15, chr18, chr20, chr21 und chr22 (Chromosom Y ist nicht in die Analyse integriert), zeigen alle Chromosomen potentielle Fusionsregionen (Abb. 3.16). Die Fusionsregionen der Translokationen zeigen hierbei ein wiederkehrendes Muster aus einem fokalen Punkt hoher Interaktionsintensität am Bruchpunkt mit einseitig graduell abfallenden Intensitäten der benachbarten Bins (Abb. 3.17 und Abb. S15 A und C; siehe auch [108]). Durch diesen Verlauf der Intensitäten kann die Orientierung des Rearrangements innerhalb des

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

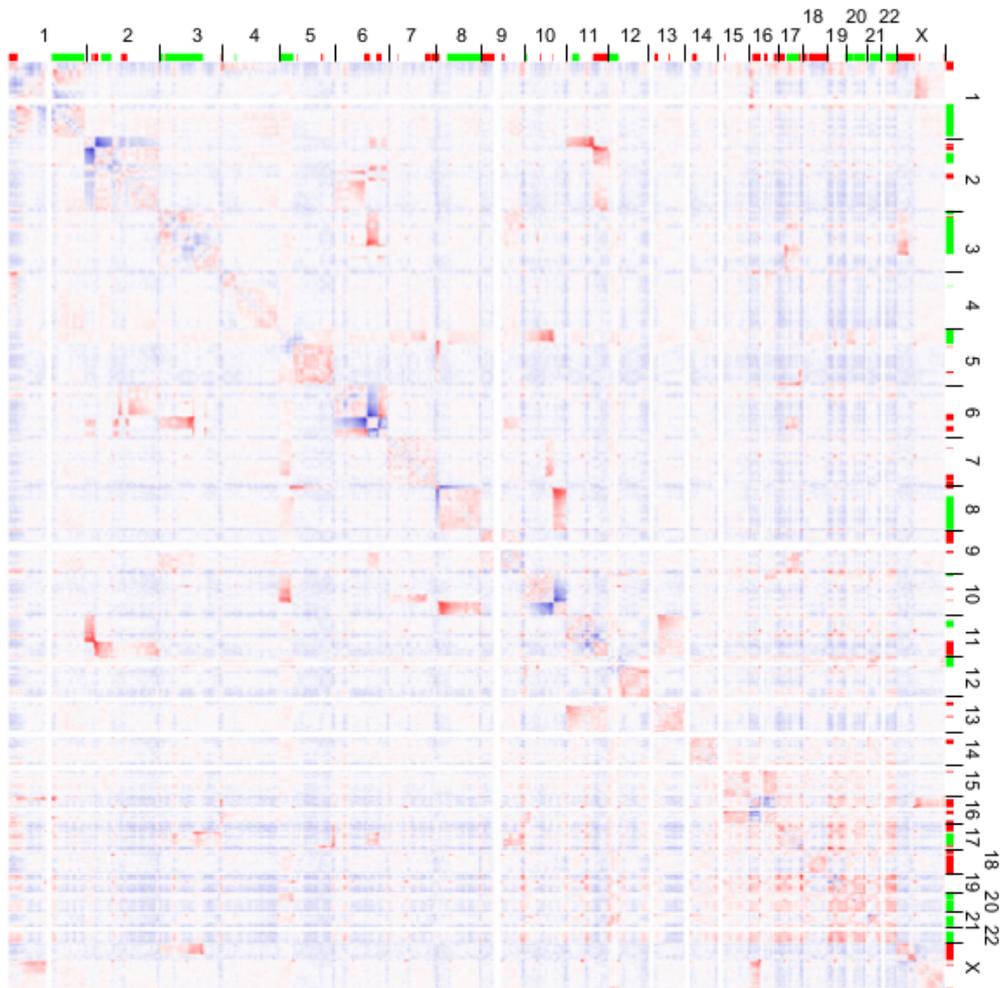


Abbildung 3.16.: Interaktionsmatrix des gesamten Seax-Genoms.

Die normalisierte Interaktionsmatrix der gesamt-genomischen intra- und interchromosomalen Interaktionen weist auf Fusionsregionen chromosomaler Rearrangements hin, dabei sind Binpaare mit hohen Interaktionsfrequenzen in rot und Binpaare mit geringen Interaktionsfrequenzen in blau dargestellt. Die Bingröße entspricht 5 Mb mit einer Fenstergröße von 2,5 Mb. Die mit den Chromosomen assoziierten DNA-Kopienzahlveränderungen sind neben der Matrix in grün (Duplikationen) und rot (Deletionen) entsprechend ihrer chromosomalen Position markiert.

beteiligten Segmentes bestimmt werden (Abb. 3.17). Dies ist jedoch nicht bei kleinen translozierten Segmenten anwendbar, da der Intensitätenverlauf hier nicht eindeutig genug ist, um auf die Richtung der Fusionsregion schließen zu können (z. B. Abb. S15 B: die Fusionsregion zeigt keinen Farbverlauf). Dazu wäre eine Erhöhung der Sequenziertiefe nötig.

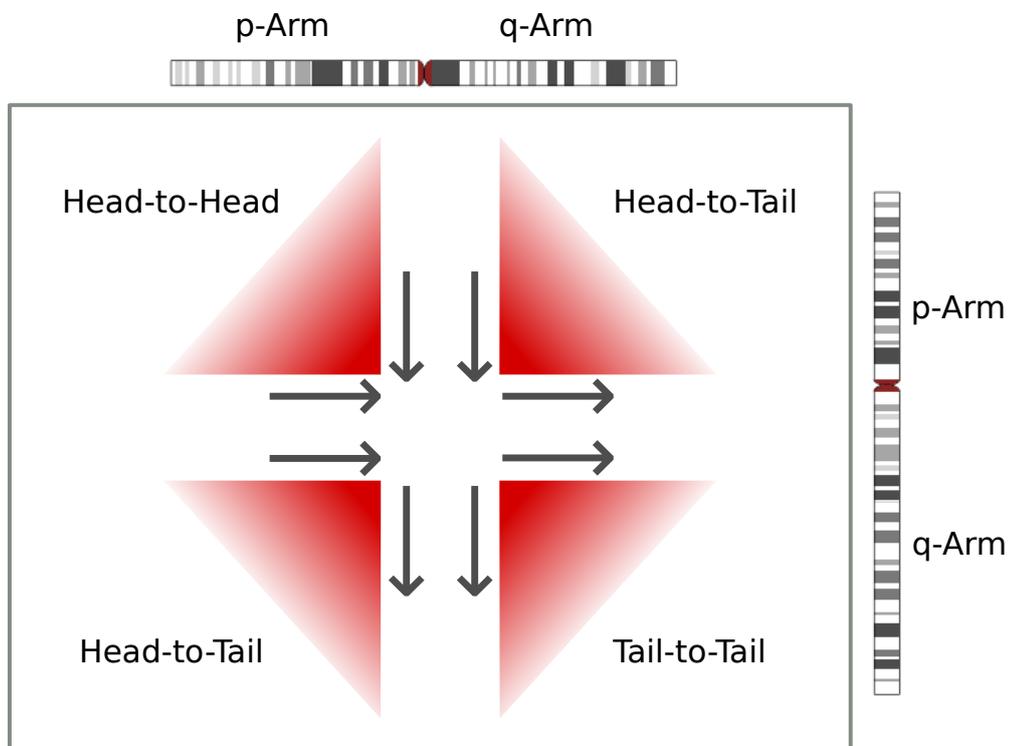


Abbildung 3.17.: Orientierung chromosomaler Fusionsregionen in den Interaktionsmatrizen.

Schema eines Rearrangementbruchpunktes, repräsentiert durch die Interaktionsmatrix und den möglichen Orientierungen der Fusionspartner. Mithilfe der Interaktionsmatrizen lassen sich nicht nur die Positionen der Fusionsregionen, sondern auch deren Orientierung ablesen. Durch die vom Bruchpunkt aus abfallende Intensität der Interaktionen kann bestimmt werden, ob die Region stromabwärts (tail) oder stromaufwärts (head) des Bruchpunktes mit dem Partnerchromosom fusioniert ist. Die Pfeile weisen hier stromabwärts und geben die Orientierung der Segmente innerhalb des Rearrangements an.

3.3.2. Kartierung der Seax-Translokationen durch interchromosomale Interaktionen

Durch die Analyse der Interaktionsmatrizen (Binggröße: 100 kb mit einer Fenstergröße von 10 kb) konnten 22 potentielle Fusionsregionen der Seax-Zelllinie bestimmt und einer Fusionsrichtung zugeordnet werden (Abb. 3.19, Abb. 3.20 und Tab. 3.7). Jedoch lassen sich die Bruchpunkte mithilfe der Interaktionsmatrizen bei der vorliegenden Sequenziertiefe nur in einem 100-kb-Bin eindeutig detektieren. Daher wurde zur genauen Bruchpunktbestimmung der Rearrangements, anstatt der Bins, die Readinformationen in den selektierten Fusionsregionen genutzt (Abb. 3.18). Dazu wurden alle Readpaare selektiert, welche in einem ± 1 -Mb-Fenster um den potentiellen Bruchpunkt vorliegen und gleichzeitig mit der entsprechenden Partnerregion interagieren. Diese wurden anschließend mit dem UCSC Genome Browser visualisiert (Abb. 3.18).

Die Schnittstellen der *HindIII*-Fragmente, die den Übergang zwischen gehäuften und sporadischem Auftreten von Readpaaren markieren, wurden als potentielle Bruchpunktregionen in Tabelle 3.7 aufgeführt. Gene, deren Positionen sich mit dem *HindIII*-Fragment überschneiden, können somit direkt von der Translokation betroffen sein und Fusionsgene repräsentieren. Die in Tab. 3.7 hervorgehobenen Gene wurden dabei mit höherer Verlässlichkeit gewertet, da diese Gene wie im Beispiel unter Abb. 3.18 zusätzliche *HindIII*-Fragmente umspannen. Eine fehlerhafte Assoziation angrenzender *HindIII*-Fragmente mit den Bruchpunkten ist damit bei der Identifizierung der Bruchpunktgene eher zu vermeiden. Einige der Bruchpunktregionen konnten jedoch nicht aufgrund der Readverteilung am Bruchpunkt bestimmt werden, da die 100-kb-Bin-Interaktionsmatrix keine eindeutigen Bruchpunkte ergab, die zur Read-Selektion vonnöten sind.

Die von den Bruchpunkten betroffenen Gene wurden mit PANTHER untersucht, was jedoch keine Anreicherungen von GO-Annotationen ergab (Daten nicht dargestellt). Des Weiteren wurden die Gene der Bruchpunkte auf potentielle Fusionsgene untersucht. Sind an beiden Translokationsbruchpunkten Gene mit gleicher Leserichtung (und zusätzlich Promotor- bzw. Terminatorsequenzen) positioniert, kann es zu einer Ausbildung von Genfusionen mit chimären Genen kommen. So sind die möglichen Fusionsgene mit gleicher Orientierung: *CIITA-CTSK* (t2), *NAA50-GOSR1* (t8), *SLC39A11-KCTD16* (t13), *AIG1-GOSR1* bzw. *GOSR1-AIG1* (t15), *GRM8-HK1* (t16) und *HMGB1-BRSK2* (t19). Ebenfalls fusionieren die Gene des Beispiels in Abb. 3.21 (und Abb. S14): *NIPBL* (chr5) zwischen Intron 38 und Intron 40 bzw. *XKR6* (chr8) am Intron 1. Allerdings fusionieren diese

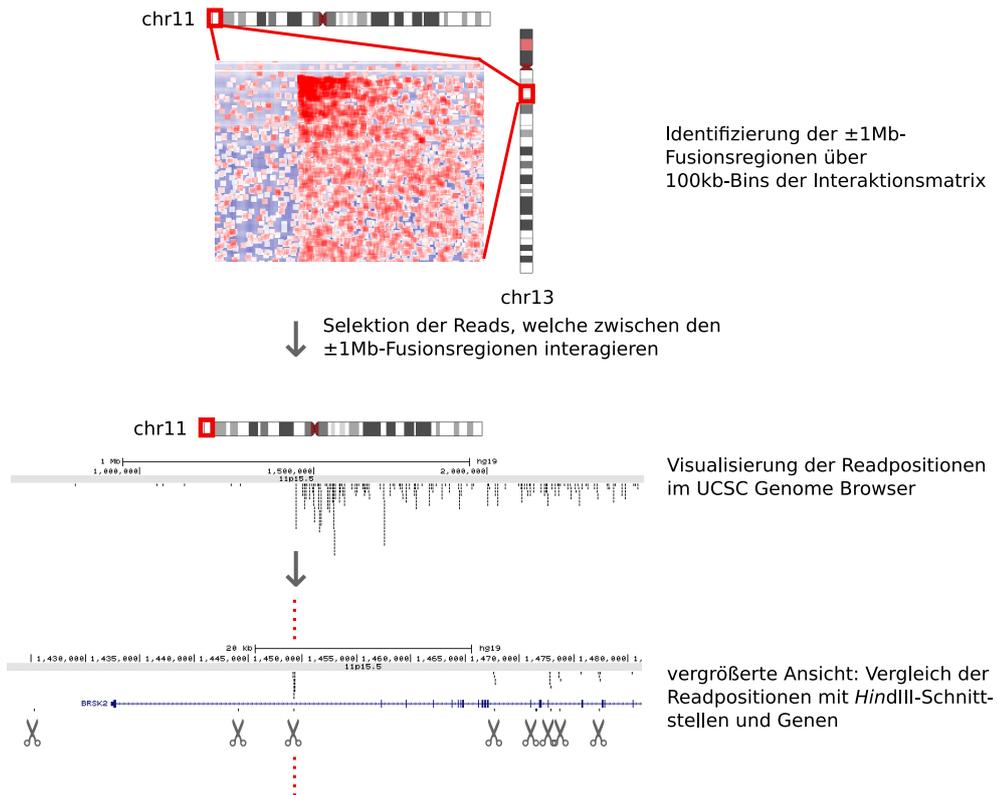


Abbildung 3.18.: Identifizierung der Fusionsregionen.

Fusionsregionen der Translokationen werden über die Interaktionsmatrizen identifiziert, dabei wird die Region auf $\pm 1\text{ Mb}$ eingegrenzt. Die Positionen der Readpaare, welche innerhalb dieser Regionen interagieren, werden über den UCSC Genome Browser dargestellt. *Hind*III-Schnittstellen (in der Abbildung repräsentiert durch Scheren), deren Position mit dem Übergang von wenigen zu vielen Reads übereinstimmen, werden mit RefSeq und UCSC Genen verglichen. Das entsprechende *Hind*III-Fragment umspannt damit den Bruchpunkt des Rearrangements. Gene, deren Position mit dem *Hind*III-Fragment überlappen, gelten als potentielle Fusionsgene.

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

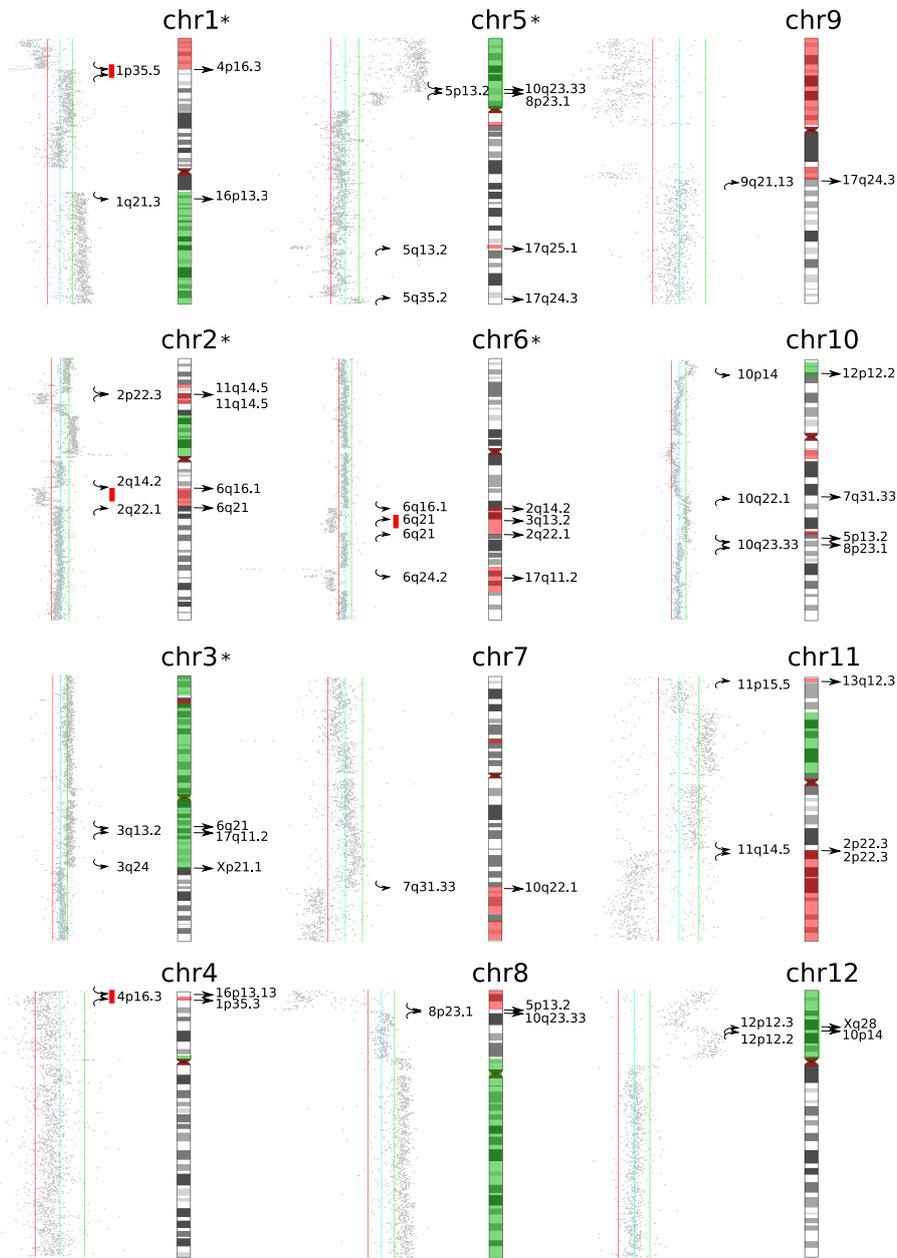


Abbildung 3.19.: Chromosomale Veränderungen der Seax-Zelllinie: Chromosomen 1 bis 12.

Interchromosomale Rearrangements und DNA-Kopienzahländerungen der Seax-Zelllinie. Links der Idiogramme: Ansicht der aCGH-Ergebnisse (links: Deletionen, rechts: Duplikationen; vertikale Linien entsprechen dem \log_2 -Verhältnis von -0,2 bzw. 0,2). Die Idiogramme mit Visualisierung der entsprechenden CBS-Ergebnisse (rot: Deletionen, grün: Duplikationen). Die Fusionsregionen des jeweiligen Chromosoms sind links des Idiogrammes mit Pfeilen entsprechend der Orientierung innerhalb des Rearrangements markiert. Die Partnerregionen sind rechts des Idiogrammes aufgelistet. Benachbarte Rearrangementsbruchpunkte innerhalb eines Chromosoms, welche eine Deletion flankieren, sind mit einem roten Balken markiert (* Chromosom weist zusätzliche, nicht dargestellte, intrachromosomale Rearrangements auf).

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

Tabelle 3.7.: Bruchpunktregionen und potentielle Fusionsgene.

Start- und Endpositionen (in hg19) entsprechen den über Interaktionsmatrizen und UCSC Genome Browser identifizierten Bruchpunkten, repräsentiert durch *HindIII*-Fragmente. Mit * markierte Startpositionen konnten nicht auf *HindIII*-Fragmentebene detektiert werden, sondern nur über die 100-kb-Bin-Interaktionsmatrix (mit Ausnahme von t15). Die identifizierten Gene überlappen mit den Positionen der betroffenen *HindIII*-Fragmente. Potentielle Fusionsgene mit höherer Konfidenz (siehe Text) sind hervorgehoben (Chr.: Chromosom, Orient.: Orientierung).

ID	Partner ID	Chr.	Start	Ende	Orient.	Gene	Information
t1	t1_1p35.3	chr1	28238215	28248287	Head	<i>RPA2</i>	—
	t1_4p16.3	chr4	4050000*	4150000	Tail	<i>BC042823</i>	—
t2	t2_1q21.3	chr1	150769818	150781274	Head	<i>CTSK</i>	mögliches
	t2_16p13.13	chr16	11004475	11019710	Head	<i>CIITA</i>	Fusionsgen
t3	t3_2q14.2	chr2	120270564	120285427	Head	<i>SCTR</i>	—
	t3_6q16.1	chr6	98127299	98143928	Head	<i>AK091365</i>	—
t4	t4_2q22.1	chr2	137227625	137228172	Tail	—	—
	t4_6q21	chr6	114395987	114399043	Tail	<i>HS3ST5</i>	—
t5	t5_11q14.3	chr11	88350000*	88450000	Head	<i>GRM5</i>	reziprok zu
	t5_2p22.3	chr2	32260000*	32360000	Head	mehrere Gene	t6
t6	t6_11q14.3	chr11	88350000*	88450000	Tail	<i>GRM5</i>	reziprok zu
	t6_2p22.3	chr2	32230000*	32330000	Tail	mehrere Gene	t5
t7	t7_3q13.2	chr3	112240000*	112340000	Head	mehrere Gene	—
	t7_6q21	chr6	106400000*	106500000	Tail	—	—
t8	t8_3q13.2	chr3	113435703	113437137	Tail	<i>NAA50</i>	mögliches
	t8_17q11.2	chr17	28807394	28808819	Head	<i>GOSR1</i>	Fusionsgen
t9	t9_3q24	chr3	143214724	143222670	Head	<i>SLC9A9</i>	—
	t9_Xp21.1	chrX	36597000	36600264	Head	—	—
t10	t10_16p13.13	chr16	11954332	11955153	Tail	—	—
	t10_4p16.3	chr4	3315595	3322490	Head	<i>RGS12</i>	—
t11	t11_5p13.2	chr5	37046678	37051545	Tail	<i>NIPBL</i>	—
	t11_8p23.1	chr8	10977335	10981039	Head	<i>XKR6</i>	—
t12	t12_10q23.33	chr10	91799056	91801084	Head	—	—
	t12_5q13.2	chr5	37046678	37051545	Head	<i>NIPBL</i>	—
t13	t13_17q25.1	chr17	70991114	71006897	Tail	<i>SLC39A11</i>	mögliches
	t13_5q31.3	chr5	143656972	143667413	Tail	<i>KCTD16</i>	Fusionsgen
t14	t14_5q35.2	chr5	176114621	176120288	Tail	—	—
	t14_17q24.3	chr17	70537801	70539618	Head	<i>LINC00673</i>	—
t15	t15_6q24.2	chr6	143380000*	143480000	Head	<i>AIG1</i>	reziprok,
	t15_17q11.2	chr17	28770000*	28870000	Tail	<i>CPD, GOSR1</i>	Fusionsgen
t16	t16_10q22.1	chr10	71038330	71050820	Tail	<i>HK1</i>	mögliches
	t16_7q31.33	chr7	126890674	126899921	Head	<i>GRM8</i>	Fusionsgen
t17	t17_10q23.33	chr10	95298787	95314132	Tail	—	—
	t17_8p23.1	chr8	10977335	10981039	Tail	<i>XKR6</i>	—
t18	t18_12p12.2	chr12	20833396	20843983	Tail	<i>PDE3A</i>	—
	t18_10p14	chr10	7970549	7976904	Head	<i>TAF3</i>	—
t19	t19_13q12.3	chr13	31132320	31134497	Tail	<i>HMGB1</i>	mögliches
	t19_11p15.5	chr11	1444072	1449221	Tail	<i>BRSK2</i>	Fusionsgen
t20	t20_12p12.3	chr12	18675961	18680225	Tail	<i>PIK3C2G</i>	—
	t20_Xq28	chrX	147292149	147300244	Tail	—	—
t21	t21_17q25.3	chr17	75511241	75516019	Tail	—	—
	t21_19p13.3	chr19	3737154	3750893	Tail	<i>TJP3</i>	—
t22	t22_17q24.3	chr17	69185554	69188239	Head	<i>BC039327</i>	—
	t22_9q21.13	chr9	74350000*	74600000	Tail	mehrere Gene	—

mit den 3'-Enden des jeweiligen Gens, sodass ein chimäres Transkript unwahrscheinlich ist. Die Bruchpunktregionen der Gene *NIPBL* und *XKR6* sind noch an weiteren Translokationen beteiligt. So transloziert der Bereich stromabwärts des *NIPBL*-Bruchpunktes (chr5, Abb. 3.21 A) und auch die stromaufwärts gelegene Region des *XKR6*-Bruchpunktes (chr8, Abb. 3.21 B) mit der Region 10q23.31. Hierbei fusioniert chr8 mit der Region stromaufwärts und chr5 mit der Region stromabwärts des chr10-Bruchpunktes (Abb. 3.21 C). Die beiden Bruchpunkte auf Chromosom 10 teilen sich jedoch auf den ersten Blick nicht dieselbe Bruchpunktregion. Allerdings ist diese Region zusätzlich von intrachromosomalen Rearrangements betroffen (Abb. 3.21 C und Abb. S14), welche die beiden Bruchpunkte der interchromosomalen Translokationen des Chromosom 10 untereinander verbinden (Abb. S14 D). Somit teilen sich die drei verschiedenen Chromosomen die Bruchpunkte, die durch eine Rearrangementkette untereinander verbunden sind (Abb. 3.23 A und Abb. S14 D). Durch gemeinsame Bruchpunkte sind auch die chromosomalen Veränderungen der Chromosomen chr1, chr4 und chr16 ein Beispiel einer solchen Rearrangementkette (Abb. 3.23 B). Eine Übersicht der 22 aus der Hi-C-Analyse resultierenden Fusionsregionen und der den Translokationen entsprechenden chromosomalen Interaktionen zeigt die Komplexität der chromosomalen Veränderungen in Seax, und dass noch weitere Rearrangementketten möglich sind (Abb. 3.22). Auffällig ist hier auch, dass zwei Rearrangementregionen häufig eine Deletion flankieren (siehe auch Abb. 3.19 und Abb. 3.20).

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

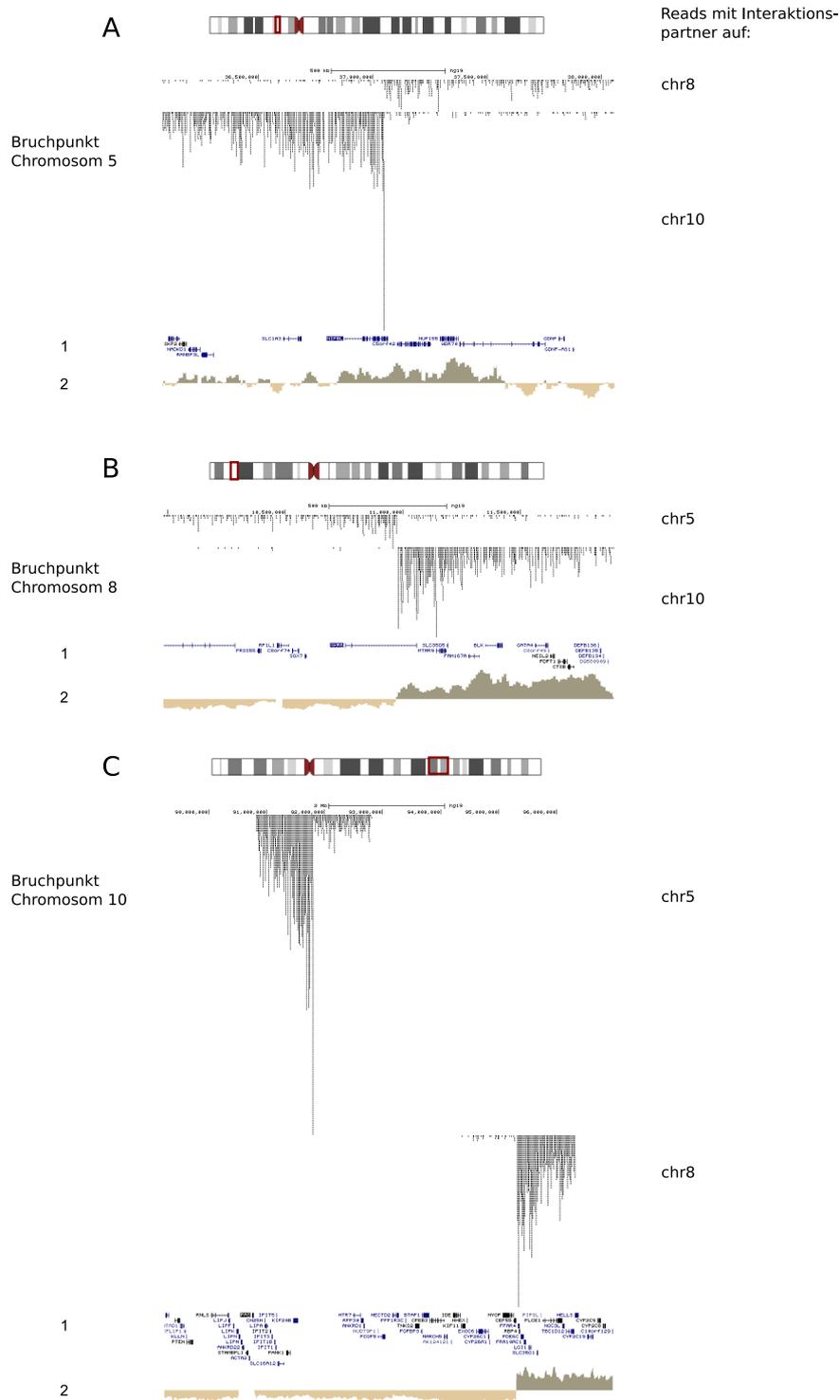


Abbildung 3.21.: Bruchpunktregionen des Rearrangements zwischen den Chromosomen chr5, chr8 und chr10 zeigen eine Rearrangementskette.

Bruchpunktregionen (± 1 Mb) der Rearrangements t11, t12 und t17. Die Translokationen bilden eine Rearrangementskette **A** des Chromosom 5, **B** Chromosom 8 und **C** Chromosom 10. Dargestellt ist die UCSC Genome Browser-Ansicht der Verteilung interagierender Reads zwischen den jeweiligen Partnerchromosomen, sowie 1) Genlokalisierungen und 2) Hauptkomponentenanalyse (*principal component analysis*; *PCA*) der Hi-C-Daten der Seax-Zelllinie.

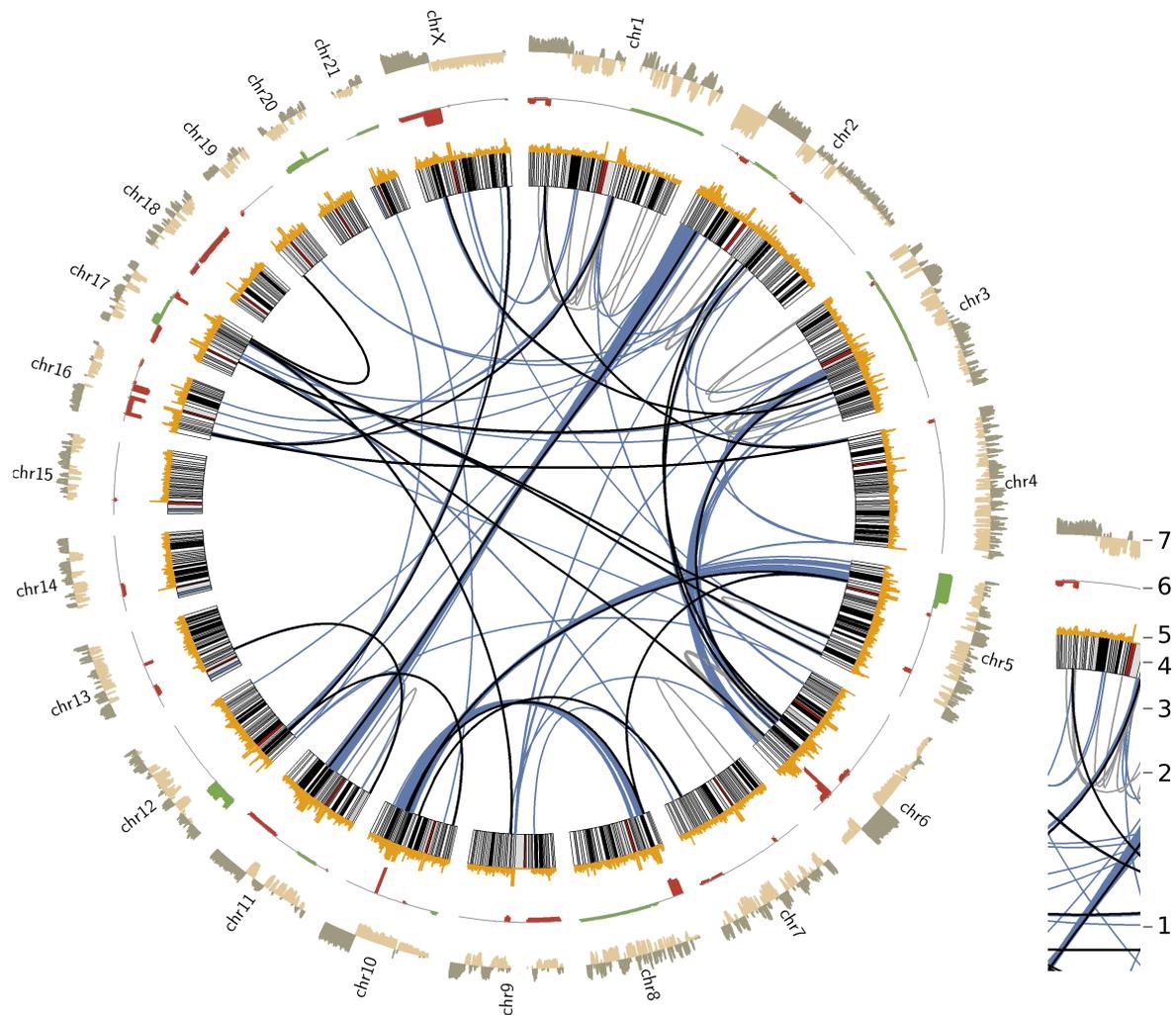


Abbildung 3.22.: Genomweite signifikante intra- und interchromosomale Interaktionen der Zelllinie Seax.

Circos-Plot der signifikanten Interaktionen zwischen 250-kb-Bins der Zelllinie Seax. Zur Reduktion und besseren Übersicht werden nur Interaktionen dargestellt, welche über Interaktionsdistanzen von über 25 Mb verlaufen, eine *False Discovery Rate* (FDR) von mindestens 0,001 aufweisen und einen Abstand von mindestens 3 Mb zum Zentromer einhalten. 1) blaue Linien: signifikante interchromosomale Interaktionen, 2) graue Linien: signifikante intrachromosomale Interaktionen, 3) schwarze Linien: die 22 über Interaktionsmatrizen ermittelten Fusionsregionen, 4) Idiogramm, 5) innerer Ring in gelb: Anzahl der signifikanten Interaktionen pro 50-kb-Bin (ohne Einschränkung auf Interaktionsdistanzen und Abstand zum Zentromer; max=10), 6) mittlerer Ring: Kopienzahlveränderungen der Zelllinie Seax (CBS-Werte der aCGH-Analyse; grün: Duplikationen, rot: Deletionen), 7) äußerer Ring: PCA-Analyse der Zelllinie Seax, entspricht A-Kompartimenten (grau) und B-Kompartimenten (beige) bzw. chromosomalen Rearrangements (z. B. Chromosom 6). Bingröße der Berechnung signifikanter Interaktionen: 250 kb mit der Fenstergröße von 50 kb.

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

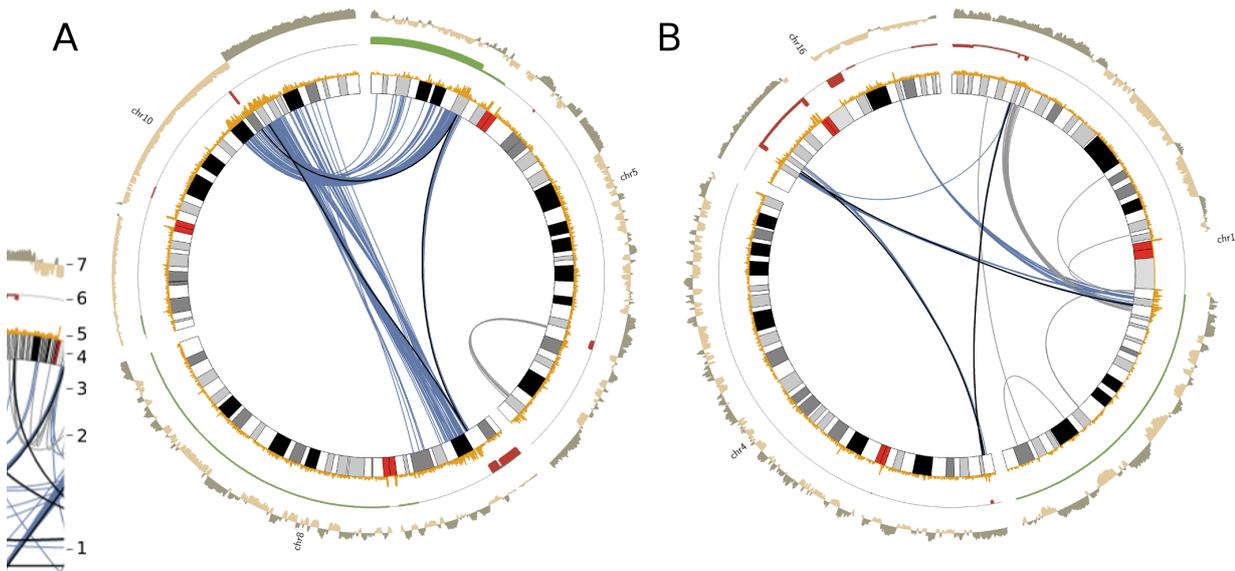


Abbildung 3.23.: Beispiele für Rearrangementsketten des Seax-Genoms.

Circos-Plots der Rearrangementskette zwischen **A** Chromosom chr5, chr8 und chr10 bzw. **B** Chromosom chr1, chr4 und chr16. Zur besseren Übersicht werden nur Interaktionen dargestellt, welche über Interaktionsdistanzen von über 25 Mb verlaufen, eine FDR von mindestens 0,001 aufweisen und einen Abstand von mindestens 3 Mb zum Zentromer einhalten. 1) blaue Linien: signifikante interchromosomale Interaktionen, 2) graue Linien: signifikante intrachromosomale Interaktionen, 3) schwarze Linien: die 22 über Interaktionsmatrizen ermittelten Fusionsregionen, 4) Idiogramm, 5) innerer Ring in gelb: Anzahl der signifikanten Interaktionen pro 50 kb Bin (ohne Einschränkung auf Interaktionsdistanzen und Abstand zum Zentromer; max=10), 6) mittlerer Ring: DNA-Kopienzahlveränderungen der Zelllinie Seax (CBS-Werte der aCGH-Analyse; grün: Duplikationen, rot: Deletionen), 7) äußerer Ring: PCA-Analyse der Zelllinie Seax (entspricht A-Kompartimenten (grau) und B-Kompartimenten (beige) bzw. chromosomalen Rearrangements (z. B. Chromosom 6). Bingröße der Berechnung signifikanter Interaktionen: 250 kb mit der Fenstergröße 50 kb.

3.3.3. Vergleich der identifizierten Fusionsregionen mit publizierten Rearrangements der Zelllinie Seax

Um zu überprüfen, inwieweit die mit Hi-C identifizierten Rearrangements mit bekannten chromosomalen Veränderungen übereinstimmen, wurden diese mit bereits publizierten NGS- und FISH-Daten von [178] verglichen. Izykowska et al. konnten mittels zytogenetischer Methoden in Seax-Zellen bereits ein komplexes Rearrangement zwischen den Chromosomen chr6, chr3, chr9 und chr17 identifizieren, welches zu einem derivaten Chromosom führt:

$der(3)t(3;6;17;9)(3pter \rightarrow 3q13.2::6q21 \rightarrow 6q24::17q11 \rightarrow 17q23?::9q22.3 \rightarrow 9qter)$ [178].

NGS-Daten derselben Studie konnten die Rearrangements durch die Fusionen von 6q24.2::17q11.2, 3q13.2::6q21 und 3q13.2::3q26.1 bestätigen. Beim Vergleich der publizierten Rearrangements mit den hier untersuchten Seax Hi-C-Daten ließen sich diese Bruchpunkte ebenfalls nachweisen (t15 entspricht 6q24.2::17q11.2, t7 entspricht 3q13.2::6q21; Abb. 3.24). Ebenfalls kann ein intrachromosomales Rearrangement der Regionen 3q13.2::3q26.1 mithilfe der Interaktionsmatrizen festgestellt werden (wurde nicht in Tab. 3.7 aufgenommen, da hier nur interchromosomale Rearrangements berücksichtigt wurden). Abweichungen der Rearrangements von [178] zeigen sich im Rearrangement zwischen den Chromosomen 9 und 17. Auch dies konnte über die Interaktionsdaten gefunden werden, allerdings zeigen diese, im Unterschied zu [178], das Rearrangement zwischen den Regionen 17q24.3::9q21.13 (t22) (im Unterschied zu den anderen Rearrangements wurde diese Translokation durch [178] nicht mittels NGS nachgewiesen, sondern bezieht sich auf G-Banding Experimente). Mithilfe der Hi-C-Interaktionsdaten lassen sich somit die Bruchpunkte des derivaten Chromosoms $der(3)t(3;6;17;9)$ bestätigen und ergänzen (Abb. 3.25).

Neben den Fusionsregionen des beschriebenen derivaten Chromosoms sind in den Interaktionsmatrizen weitere Bruchpunktregionen zu erkennen (graue Kästen in Abb. 3.24). Die Chromosomen chr3, chr6 und chr17 sind dementsprechend zusätzlichen intra- und interchromosomalen Rearrangements unterlegen, deren Fusionsregionen jedoch nicht mit weiteren derivaten Chromosomen übereinstimmen, die in [178] mittels zytogenetischer Methoden identifiziert wurden. Die Hi-C-Daten weisen hierbei auf weitere Fusionen zwischen den intrachromosomalen Regionen 6q25.2::6q21, 3q13.2::3q25.1, 3q13.2::3q26.1 und den interchromosomalen Regionen 3q13.2::17q11.2 (t8, Tab. 3.7) hin. Weiterhin lässt sich ver-

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

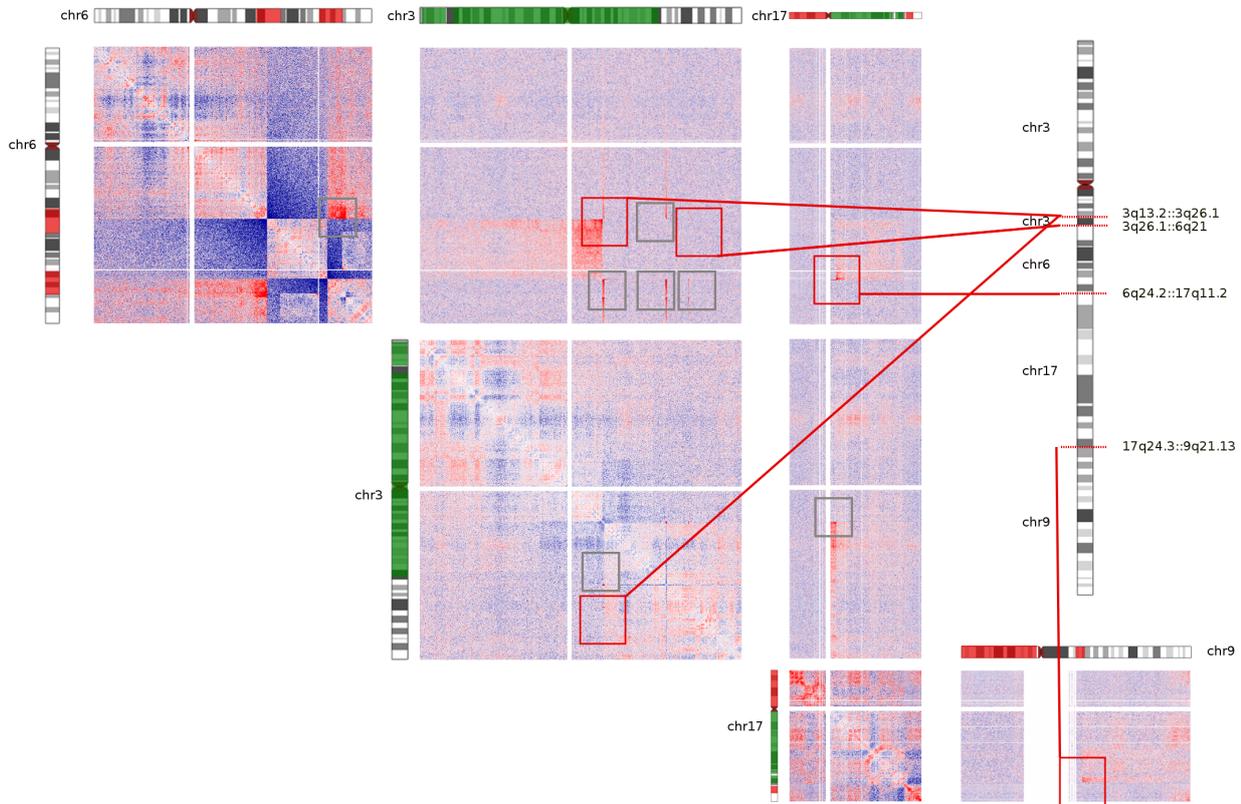


Abbildung 3.24.: Das derivative Chromosom aus Segmenten der Chromosomen chr3, chr6, chr9 und chr17.

Normalisierte Interaktionsmatrizen der intra- und interchromosomalen Interaktionen der Chromosomen chr6, chr3, chr9 und chr17 zeigen Übereinstimmungen mit Fusionsregionen des in [178] beschriebenen derivaten Chromosoms der(3)t(3;6;17;9) (rote Kästen). Fusion 3q26.1::6q21: Zur besseren Übersicht ist dies in Abb. S15 vergrößert. Fusion 3q13.2::3q26.1: Die Fusion ist in der intrachromosomalen Matrix des Chromosom 3 repräsentiert, jedoch auch in der Matrix zwischen Chromosom 3 und Chromosom 6. Durch die geringe Größe von wenigen kb des 3q26.1-Segmentes wird auch der angrenzende 3q13.3-Bruchpunkt im derivaten Chromosom in die Nachbarschaft des chr6-Segmentes gebracht und weist dadurch vermehrte Interaktionen zu chr6 auf. Fusion 6q24.2::17q11.1: Das distale Ende des Chromosom 6-Fragmentes ist mit Chromosom 17 fusioniert. Fusion 17q24.3::9q21.13: Die Regionen der Bruchpunkte der Fusion zwischen Chromosom 17 und 9 konnten ebenfalls bestimmt werden, weicht aber von [178] ab. Zusätzlich zu den in [178] gefundenen Rearrangements gibt es weitere Regionen, die von chromosomalen Rearrangements betroffen sind (graue Kästen). Darunter befinden sich intrachromosomale Rearrangements des Chromosom 6 und Chromosom 3. Die drei Segmente 3q13.2, 3q25.1 und 3q26.1 scheinen dabei in das Chromosom 6 zu integrieren, anstelle des 14 Mb großen Segmentes, welches am derivaten Chromosom beteiligt ist, oder im homologen Chromosom. Weiterhin gibt es eine Fusion zwischen den Chromosomen 3 und 17, welche die Regionen 3q13.2 und 17q11.2 involviert. Chromosomale Aberrationen sind entsprechend ihrer Lokalisation grün (Duplikationen) oder rot (Deletionen) am Idiogramm markiert. Bingröße der Interaktionsmatrizen: 250 kb mit einer Fenstergröße von 50 kb.

muten, dass die verbliebenen chr6-Segmente, die durch das ins derivate Chromosom integrierte 14-Mb-große Fragment separiert werden, unter Aufnahme der drei chr3-Fragmente (mit den Bruchpunkten bei 3q13.2, 3q25.1 und 3q26.1) fusionieren (Abb. 3.24). Möglicherweise ist aber auch ein homologes Chromosom an diesem Rearrangement beteiligt. Vergleiche mit Interaktionsmatrizen weiterer Chromosomen weisen außerdem auf Rearrangements zwischen den Chromosomen chr6 und chr2 hin. Die fusionierenden Bruchpunktregionen sind hierbei flankierende Regionen von deletierten Segmenten (Deletion chr2: 2q14.2-q22.1; Deletion chr6: 6q16.1-q21) und resultieren in den Fusionen 2q14.2::6q16.1 (t3) und 2q22.1::6q21 (t4; Tab. 3.7, Abb. S15 und Abb. 3.25).

Die Bruchpunktregionen der von [178] identifizierten Rearrangements zeigen in den Hi-C Interaktionsmatrizen auch Beispiele von reziproken Translokationen. Reziproke Rearrangements erschweren jedoch die Identifizierung der genauen Bruchpunkte über die Readverteilung innerhalb der fusionierenden Regionen (siehe Abb. 3.18). Das Beispiel des reziproken Rearrangements zwischen den Regionen 6q24.2 und 17q11.2 (t15) und deren entsprechenden Interaktionen zeigt Abb. 3.26. Laut [178] liegen die Bruchpunkte hierbei innerhalb des Intron 2 des *AIG1*-Gens der Region 6q24.2 und des Intron 2 des *GOSR1*-Gens in der 17q11.2-Region und resultieren in *out-of-frame AIG1-GOSR1*-Transkripte. Die Readinteraktionsverteilungen beider Fusionsregionen in Abb. 3.26 verdeutlichen jedoch auch ein reziprokes Rearrangement. Diese ist in Abb. 3.26 A jedoch nicht ersichtlich, da der eigentliche Bruchpunkt innerhalb der Readinteraktionsverteilung verdeckt ist. Interaktionsmatrizen weisen jedoch auf reziproke Rearrangements hin (Abb. S15 C) und auch der Bruchpunkt des *AIG1* Intron 2 aus [178] (roter Pfeil, Abb. 3.26) weicht von der Readinteraktionsverteilung in Abb. 3.26 A ab. Erst bei getrennter Visualisierung der Regionen stromaufwärts bzw. stromabwärts der Bruchpunkte aus [178] wird deutlich, dass zu beiden Seiten des *AIG1* Intron 2 bzw. *GOSR1* Intron 2 vermehrt Interaktionen auftreten, welche jedoch reziproke Fusionsregionen aufweisen (Abb. 3.26 B1-C2).

Somit lassen sich die, von [178] über NGS-Daten ermittelten, Rearrangementbruchpunkte durch die Hi-C Interaktionsdaten auch auf Ebene von *HindIII*-Fragmenten bestätigen. Weiterhin konnten zusätzliche Rearrangements in Form von intra- und interchromosomalen Veränderungen und reziproken Translokationen identifiziert werden.

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

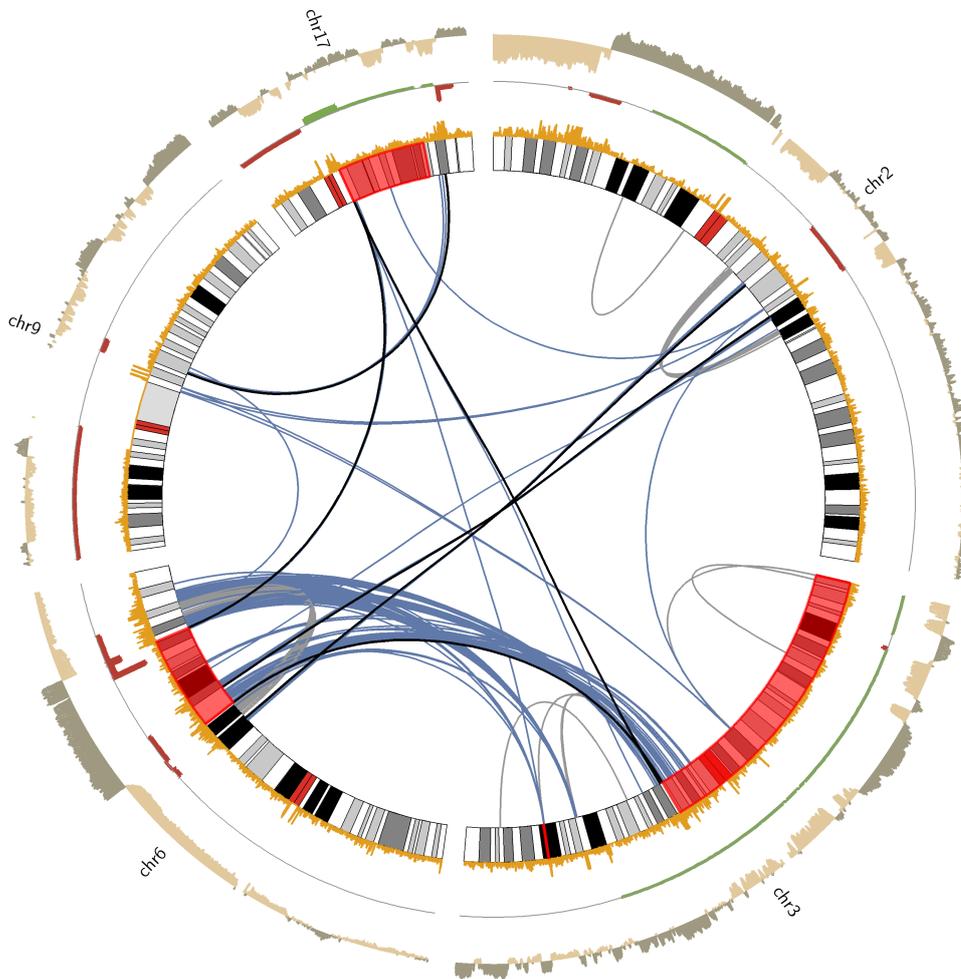


Abbildung 3.25.: Repräsentation des derivaten Chromosoms über chromosomale Interaktionen.

Entsprechend [178] sind die Chromosomen chr3, chr6, chr9 und chr17 am derivaten Chromosom der(3)t(3;6;17;9) beteiligt. Zusätzliche Translokationen zu Chromosom 2 konnten mithilfe der Hi-C-Daten der Zelllinie Seax identifiziert werden. Die Chromosomensegmente des derivaten Chromosoms sind entsprechend [178] innerhalb des Idiogrammes rot markiert. Signifikante chromosomale Interaktionen (Bingröße der Berechnung signifikanter Interaktionen: 250 kb mit der Fenstergröße 50 kb), welche mittels Hi-C identifiziert wurden, sind in blau dargestellt (ohne Interaktionen, die nahe der Zentromerregionen lokalisiert sind oder eine Interaktionsdistanz von unter 25 Mb aufweisen). Interaktionen, welche die Fusionen in Tab. 3.7 repräsentieren, sind in schwarz dargestellt. Innerer Ring: Histogramm der Readinteraktionshäufigkeiten pro 50-kb-Bin in gelb (ohne Einschränkungen der Zentromerregionen und Interaktionsdistanzen). Mittlerer Ring: Ergebnis der aCGH-Analyse von Seax als \log_2 -Intensitäten mit Deletionen (in rot) und Duplikationen (in grün). Äußerer Ring: Hauptkomponentenanalyse der intrachromosomalen Interaktionen der Seax-Zelllinie. Die Domänengrenzen markieren häufig Fusionsregionen der Translokationen.

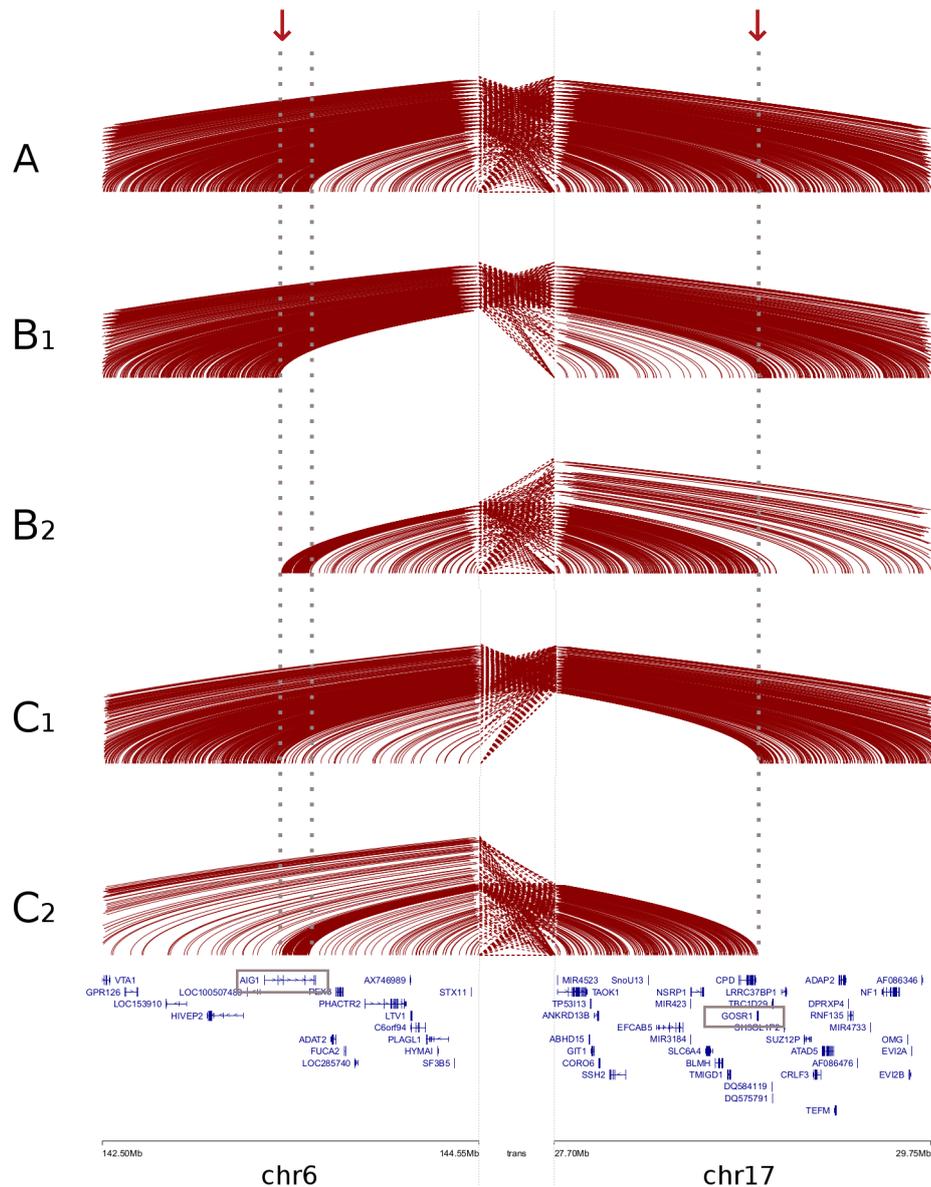


Abbildung 3.26.: Reziproke Rearrangements zwischen den Chromosomen chr6 und chr17.

Bruchpunktkartierung der Translokation $t(6;17)$ (t15). UCSC Genome Browser Ansicht der Readpaarinteraktionen zwischen den Bruchpunktregionen 6q24.2 und 17q11.2 (vertikale gestrichelte Linien: Rearrangementsbruchpunkte laut Seax-Hi-C; rote Pfeile: Bruchpunkte aus [178]). Dabei wurden selektiv die Interaktionen stromaufwärts bzw. stromabwärts eines Bruchpunktes gefiltert, um die Interaktionen nur einer der reziproken Translokationen zu visualisieren **A** Gesamte Readpaarinteraktionen **B₁** Ausgehende Readpaarinteraktionen stromaufwärts des *AIG1* Intron 2 (chr6) **B₂** Ausgehende Interaktionen der Regionen stromabwärts des *AIG1* Intron 2 (chr6) **C₁** Ausgehende Interaktionen stromabwärts des *GOSR1* Intron 2 (chr17) **C₂** Ausgehende Interaktionen stromaufwärts des *GOSR1* Intron 2 (chr17).

3.3.4. Genom- und Chromatineigenschaften der Translokationsbruchpunkte

Für ein besseres Verständnis des Einflusses der Chromatinorganisation auf die Entstehung von Rearrangements wurden Genom- und Chromatineigenschaften der Fusionsregionen charakterisiert. Dabei wurden 1 Mb stromaufwärts und stromabwärts der Rearrangementbruchpunkte publizierte und eigene Daten verschiedener Chromatincharakteristika analysiert. Zunächst wurden sowohl der Durchschnitt des DNA-Kopienzahlstatus aller Rearrangementregionen als auch der Durchschnitt der Intensitäten epigenetischer Modifikationen der Seax-Zelllinie mit denen der CTCL-Zelllinien HH, Hut78, Myla und anderen Krebszelllinien wie SkBr3, SH-SY5Y, Mewo und Hey verglichen (Abb. 3.27). Bei der Betrachtung der DNA-Kopienzahl zeigt sich ausschließlich in Seax eine Assoziation der nicht-translozierten Bereiche der Rearrangementbruchpunkte mit Deletionen der Zelllinie. In einem Abstand von 500 kb zum Bruchpunkt zeigen 18 der 35 untersuchten Bruchpunktregionen des Seax-Genoms nicht-translozierte Regionen mit einem geringeren DNA-Kopienzahlstatus als die translozierte Region (Abb. 3.27). Davon sind 14 nicht-translozierte Regionen in diesem Bereich mit einer Deletion assoziiert. Dementsprechend werden die untersuchten Rearrangements von Deletionen begleitet, die ausschließlich in der Seax-Zelllinie vorliegen. Im Gegensatz dazu zeigen die Muster epigenetischer Modifikationen aller Zelllinien einen ähnlichen Verlauf. Ein allgemeines Muster der Modifikationen ist dabei besonders deutlich am Beispiel der H3K4-Methylierung zu erkennen. Hier zeigt sich ein sinusförmiger Verlauf mit lokalen Maxima bei ca. -800 kb, -200 kb, 300 kb und 800 kb. Dieses Muster ist auch bei der DNA-Methylierung und der aH4-Methylierung zu erkennen, wobei einige Zelllinien statt der zwei Peaks im nicht translozierten Bereich nur einen Peak zwischen -500 kb und -300 kb aufweisen. Der Peak mit den höchsten assoziierten Intensitäten liegt dabei ca. -200 kb entfernt von der Bruchpunktregion im nicht-translozierten Bereich. Die Ausnahme dazu bilden die wenig ausgeprägten Peaks der H3K9-Methylierung. Hier kommt es durch eine Verschiebung des Musters zu einer antagonistischen Peakposition. Des Weiteren ist auffällig, dass die Intensitäten der Modifikationen im nicht-translozierten Bereich, mit Ausnahmen der aH4-Modifikation, in der Seax-Zelllinie am höchsten ist.

Als nächstes wurden verschiedene genomische Eigenschaften der Seax-Bruchpunktregionen untersucht (Abb. 3.28 und Abb. S16). Um zu untersuchen, ob die Regionen im Allgemeinen von chromosomalen Veränderungen betroffen sind, wurden diese mit der

DGV-Datenbank nicht-pathogener struktureller Variationen und mit *Genomic Disorders*-assoziierten Regionen (entsprechend dem Syndromreport von DECIPHER) abgeglichen. Dies ergab jedoch keine Hinweise auf eine Anhäufung von nicht-pathogenen strukturellen Variationen in direkter Umgebung der Bruchpunkte. Allerdings liegen vier Bruchpunktregionen in der Nähe zu *Genomic Disorder*-assoziierten Regionen (t17_8p23.1 und t11_8p23.1: 8p23.1 Duplikations/Deletions Syndrom, t14_5q35.2: Sotos Syndrom und ca. 300 kb stromabwärts der Region des t8_17q11.2: NF1-Mikrodeletionssyndrom). Weiterhin wurden die Seax-Bruchpunkte mit Deletionen und Duplikationen in den Tumorzellgenomen der Sézary-Patienten verglichen. Dies ergab, dass 16 der Bruchpunkte in mindestens einer Tumorzellprobe Teil von Deletionen sind, während 13 Bruchpunktregionen Teil von Duplikationen sind. Hingegen wiesen 11 Bruchpunktregionen keine Assoziationen mit Kopienzahlveränderungen in Sézary-Zellen der Patienten auf. Die Bruchpunktregionen, welche in den Patientenproben am häufigsten von Deletionen bzw. Duplikationen betroffen sind, werden durch t21_19p13.3 (10 Patienten mit Deletion) bzw. t22_17q24.3 (9 Patienten mit Duplikation) repräsentiert. Weiterhin zeigen 10 der Bruchpunktregionen Überlappungen mit evolutionären Bruchpunkten, wobei jedoch nur der Bruchpunkt der Region t16_10q22.1 eine direkte Übereinstimmung mit einem evolutionären Bruchpunkt aufweist. Zusammenfassend gibt es keinen eindeutigen Zusammenhang zwischen den Positionen der Rearrangementbruchpunkte in Seax und einer allgemeinen Anfälligkeit dieser Region für chromosomale Veränderungen. Als weitere genomische Eigenschaften der Bruchpunkte wurde ihre mögliche Assoziation mit LADs untersucht. Dies ergab, dass acht der Bruchpunkte innerhalb einer LAD (definiert in Tig3-Zellen) liegen, und alle weiteren sich in inter-LAD-Regionen befinden. Weiterhin sind 23 Bruchpunkte innerhalb von Sub-TADs (der Zelllinie GM12878) lokalisiert, wobei eine GO-Analyse der Gene (mittels PANTHER) innerhalb der bruchpunktüberlappenden GM12878-TADs, keine Anreicherung von Genen gleicher molekularer Funktion oder molekularer Prozesse ergab. Die Untersuchung des Chromatinstatus (entsprechend der Zelllinie GM12878) zeigt, dass 20 Bruchpunkte eher mit aktivem, 13 Bruchpunkte eher mit inaktivem Chromatin assoziiert sind, wobei 7 der 16 untersuchten Bruchpunktpaare den gleichen Aktivitätsstatus zur korrespondierenden Fusionsregion aufweisen. Da Chromosomenbrüche im Genom von Krebszellen häufig innerhalb von *fragile sites* auftreten, wurde die Position der Bruchpunktregionen mit denen von bekannten *fragile sites* verglichen. Dabei zeigt sich, dass vier der hier untersuchten Bruchpunktregionen mit *common fragile sites* assoziiert sind. So wird die Bruchpunktregion t2_1q21.3 von der *common fragile site* FRA1F umspannt und drei weitere Bruchpunkte (t16_10q22.1, t4_6q21 und t4_2q22.1) befinden sich inner-

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

halb des analysierten 2-Mb-Fensters in der Nähe von *common fragile sites*. Des Weiteren ist an dem Rearrangement t20_Xq28 ein Bruchpunkt zwischen den Genen *FMR1* und *AFF2* beteiligt, welche mit den raren *fragile sites* FRAXA bzw. FRAXE assoziiert sind (nicht unter Abb. 3.28 abgebildet).

Weiterhin wurden Sequenzeigenschaften untersucht, welche laut Literatur mit Chromosomenbrüchen assoziiert sind. Dies ergab jedoch keine eindeutige Anhäufung oder auffällige Verteilungsmuster von *Mirror Repeats*, *Alu*-, G4-Sequenzen und zDNA innerhalb der Bruchpunktregionen (Abb. S16; hier wurde ein kleinerer Ausschnitt der Bruchpunktregion von ± 50 kb gewählt). Als weitere destabilisierende Sequenzeigenschaft des Genoms wurden SDs untersucht. Der Vergleich zeigt zwar einige bruchpunktnahe Regionen mit ausgedehnten SD-Sequenzen, jedoch sind diese nicht in direkter Nähe zu den Bruchpunkten lokalisiert und treten nur in vereinzelt Bruchpunktregionen auf. Des Weiteren zeigt der Circos-Plot in Abb. 3.29, dass nur wenige paraloge SD-Paare innerhalb dieser Bruchpunktregionen lokalisiert sind. Lediglich die Regionen Xq28, 16p13.13 und 5q31.3 bzw. 3q13.3, 4p16.3 und 8p23.1 weisen untereinander paraloge SDs auf. Jedoch sind die betroffenen Regionen keine Translokationspartner. Nicht nur über SD-Clusterregionen und paraloge SDs kann homologiebasierte genomische Instabilität erzeugt werden, auch kurze LCRs, sogenannte *self chains* können über destabilisierende Sekundärstrukturen wirken [457]. Ein Vergleich der Bruchpunktregionen mit der *self chain*-Datenbank des UCSC-Browsers (*Human Chained Self Alignments*) [73, 194, 362] [195] zeigt hierbei zwar fünf verbundene *self chain*-Paare, jedoch kolokalisieren diese nicht zwischen Partnerregionen der Translokationen (Abb. 3.29).

In der Literatur wird beschrieben, dass Translokationsregionen auch vor dem eigentlichen Rearrangement eine höhere Kontaktwahrscheinlichkeit innerhalb der Kernorganisation haben, möglicherweise unter Einwirkung zelltypspezifischer Transkriptionsfaktoren [361]. Ein Vergleich der Bruchpunktregionen mit der chromosomalen 3D-Organisation verschiedener eukaryotischer Zelltypen (4DGenome-Datenbank [406]) und der B-Lymphoblastenzelllinie GM12878, ergab jedoch keine signifikanten chromosomalen Interaktionen zwischen fusionierten Bruchpunktregionen (und sind daher in Abb. 3.29 nicht dargestellt). Ebenfalls kann auch der Vergleich mit Loci von Genen, die mit den GO-Annotation „Lymphozyten“ (sechs Bruchpunktregionen weisen Gene dieser GO-Annotation auf) bzw. „Transkriptionsfaktoren der T-Zell-Entwicklung“ (eine Bruchpunktregion weist ein Gen auf, welches im Zusammenhang mit Th₂-Differenzierung steht) assoziiert sind keine gemeinsame Eigenschaft dieser Rearrangementbruchpunktregionen zeigen.

Die Bruchpunktregionen der Translokationen der Seax-Zelllinie gehen mit Deletionen des nicht-translozierten Bereiches einher. Weiterhin kolokalisieren, laut der epigenetischen Histonmodifikationen der Seaxzelllinie und des Chromatinstatus dieser Regionen in der Zelllinie GM12878, die Bruchpunkte eher mit aktiven Chromatinbereichen. Die weiteren untersuchten genomischen Eigenschaften der Chromatin- und Sequenzstruktur der Bruchpunktregionen ergaben keine eindeutigen gemeinsamen Charakteristika.

Abbildung 3.27. : (Nächste Seite) Chromatineigenschaften der Translokationsbruchpunkte.

Genom- und Chromatineigenschaften der Regionen -1 Mb (nicht-translozierte Region) und +1 Mb (translozierte Region) der Seax-Rearrangementbruchpunkte. Die folgenden Genom- und Chromatineigenschaften wurden untersucht: 1) DNA-Kopienzahl, 2) DNA-Methylierung, 3) H3K9me2, 4) H3K4me3 und 5) H4-Acetylierung. Linke Seite: Durchschnittliches DNA- bzw. Modifikationsprofil aller Seax-Bruchpunktregionen für die Zelllinien HH, Hey, Hut78, MeWo, Myla, SH-SY5Y, Seax und SkBr3. Die durchschnittlichen log-Intensitäten der 20-kb-Bins sind dabei als geglättete (*smoothed*) Kurven dargestellt. Rechte Seite: Heatmap der 20-kb-gebinteten aCGH(CBS)-, MeDIP- bzw. ChIP-chip-Resultate der Zelllinie Seax für jede einzelne Bruchpunktregion. Genomische DNA (grün: Duplikation, rot: Deletion), DNA-Methylierung (rot: DNA-Hypermethylierung, blau: DNA-Hypomethylierung), Histonmethylierung (rot: Histonhypermethylierung, blau: Histonhypomethylierung), Histonacetylierung (rot: Histonhyperacetylierung, blau: Histonhypoacetylierung).

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

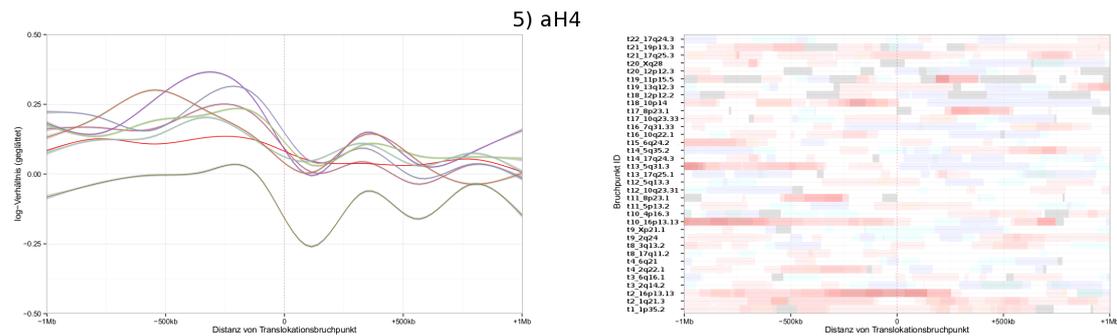
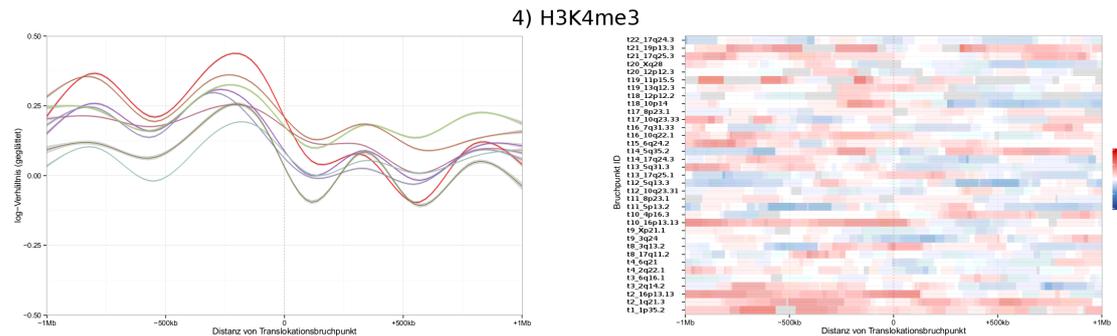
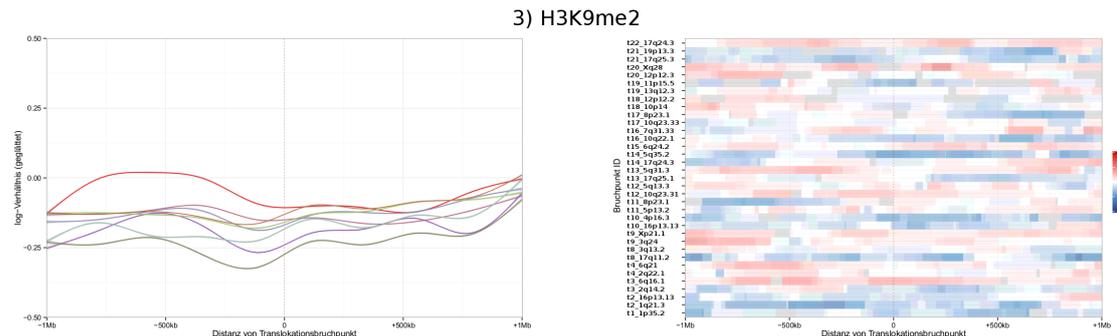
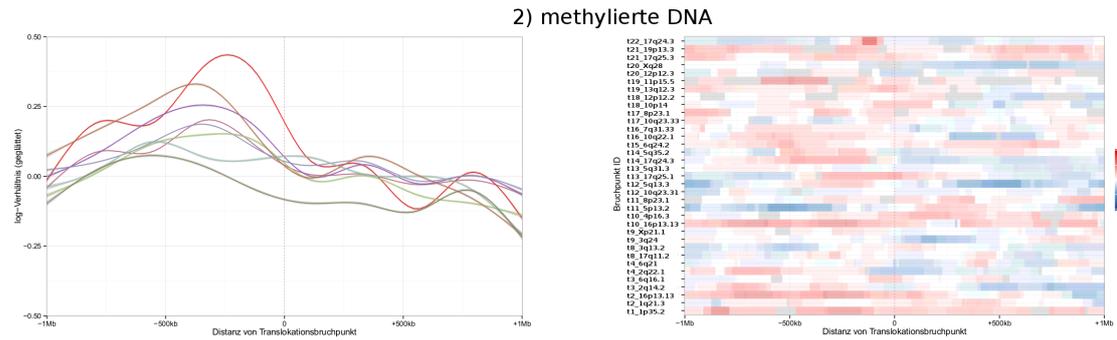
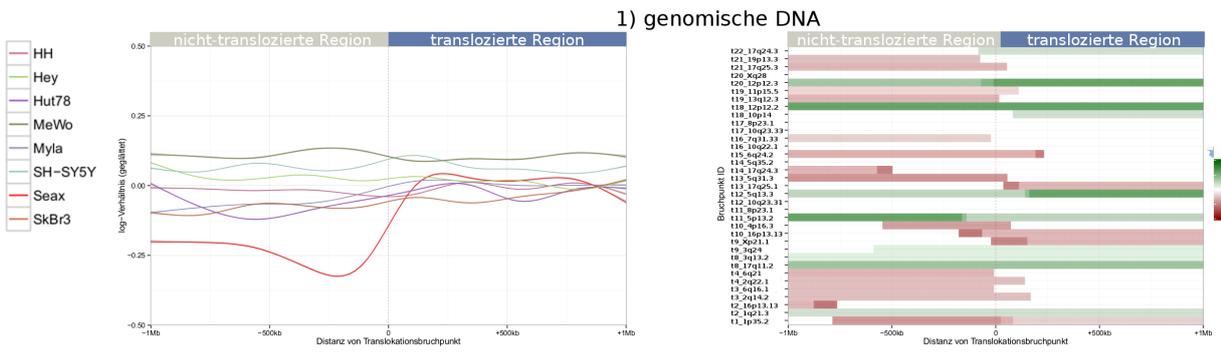
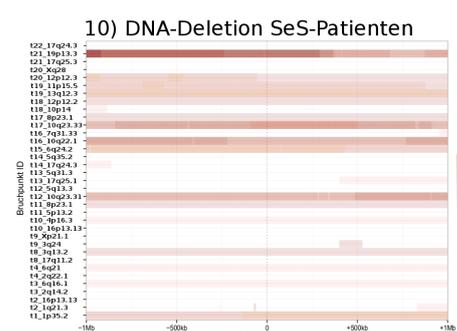
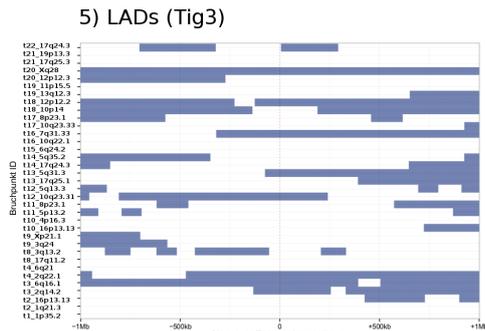
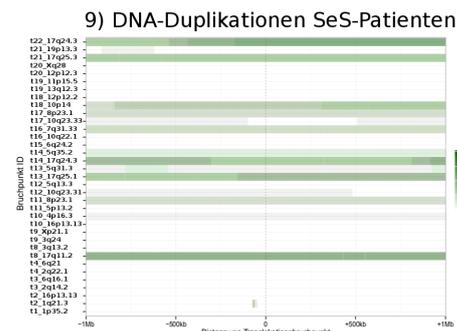
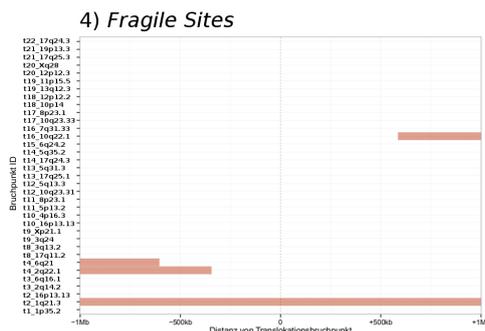
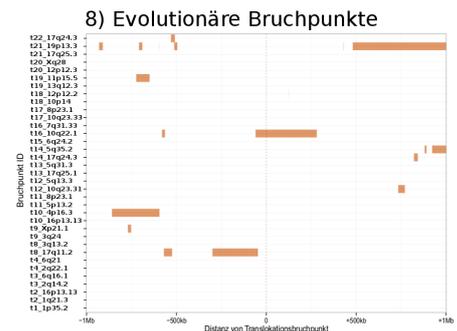
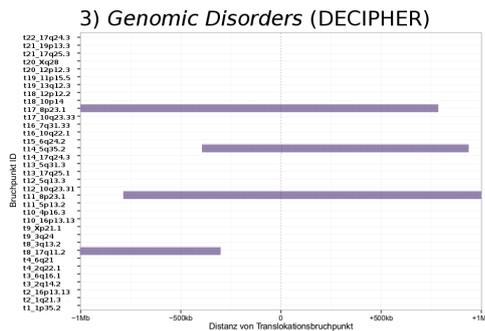
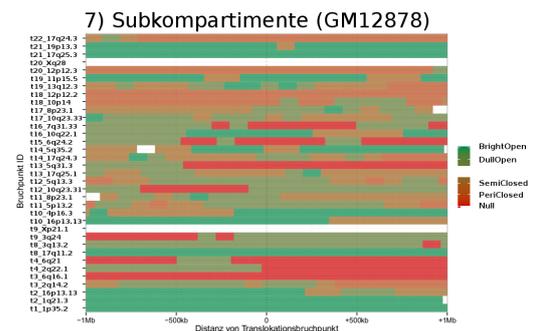
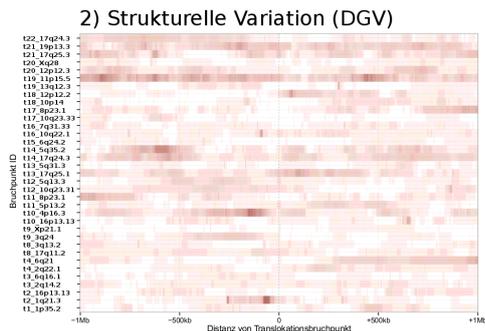
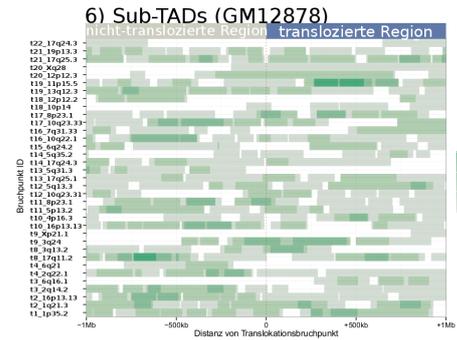
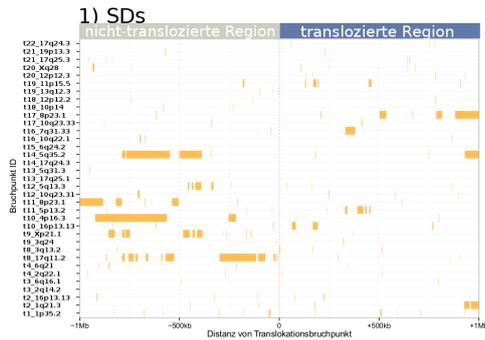


Abbildung 3.28. : (Nächste Seite) Genomeigenschaften der Translokationsbruchpunkte.

Genomeigenschaften der Regionen -1 Mb (nicht-translozierte Region) und +1 Mb (translozierte Region) der Seax-Rearrangementbruchpunkte. 1) Segmentale Duplikationen, 2) Strukturelle Variationen (SV) der DGV-Datenbank (Intensität entspricht der Anzahl an SVs), 3) *Genomic Disorders* der DECIPHER-Datenbank, 4) *Common fragile sites*, 5) LADs der Zelllinie Tig3, 6) Sub-TADs (Intensität entspricht der Anzahl der ineinander geschachtelten Sub-TADs), 7) Subkompartimente assoziiert mit dem Chromatinstatus *BrightOpen*, *DullOpen*, *SemiClosed*, *PeriClosed*, *Null* der Zelllinie GM12878 (Intensität entspricht dem assoziierten Aktivitätsstatus der Region mit aktiven Regionen in grün und inaktiven Regionen in rot), 8) Evolutionäre Bruchpunktregionen, 9) DNA-Duplikationen in Tumorzellen der SeS-Patienten (Intensität entspricht der Anzahl der Patienten mit DNA-Duplikation), 10) DNA-Deletion in Tumorzellen der SeS-Patienten (Intensität entspricht der Anzahl der Patienten mit DNA-Deletion)

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax



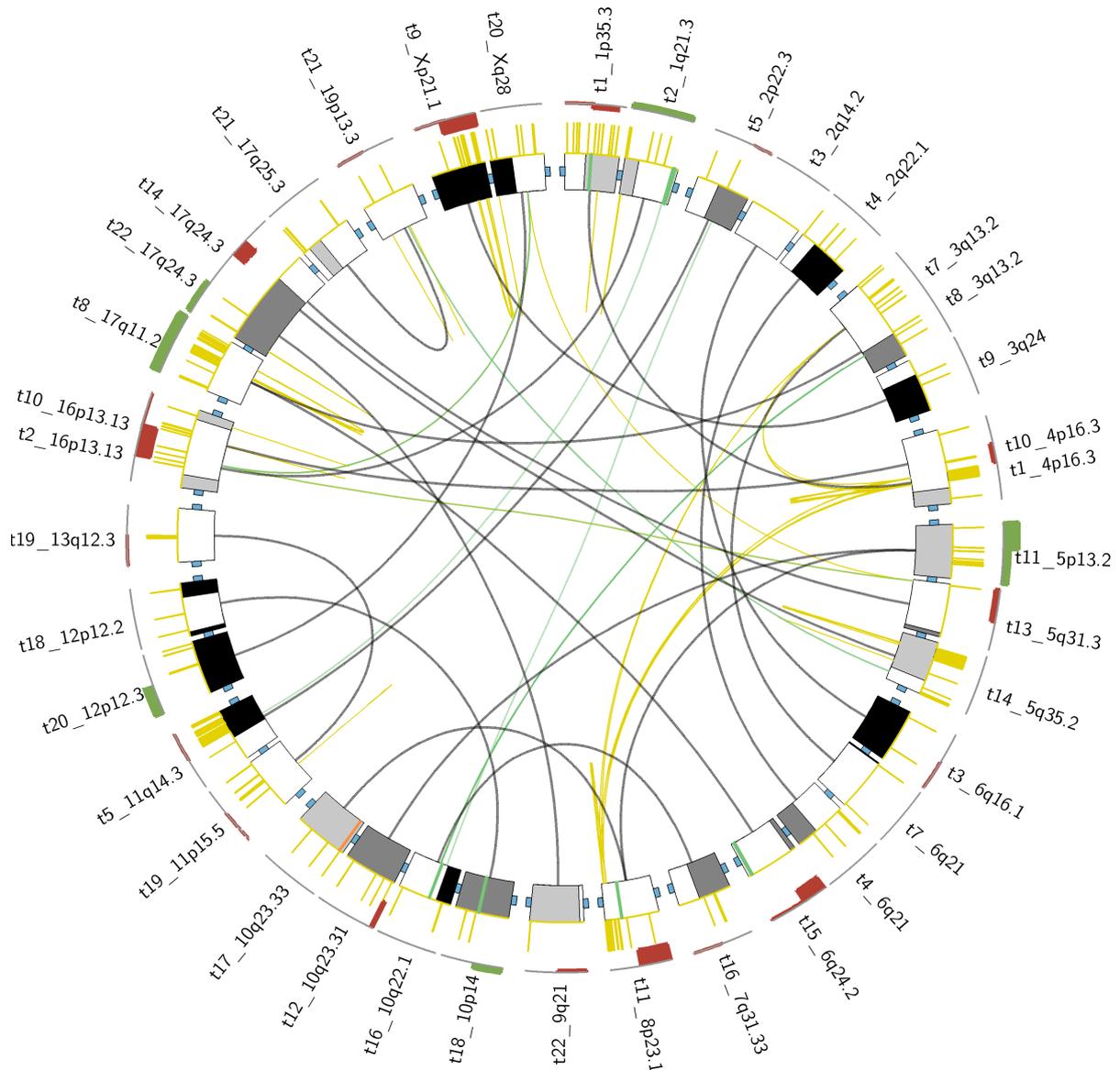


Abbildung 3.29.: Bruchpunktregionen der Rearrangements im Vergleich zu SD- bzw. *self chain*-Paaren, ausgesuchten Genlisten und dem DNA-Kopienzahlstatus in Seax.

Circos-Plot der 2-Mb-Bruchpunktregionen mit den darin lokalisierten SD-Paralogen (gelbe Linie; ohne Einschränkung der Sequenzähnlichkeit) und *self chain*-Paaren (grüne Linie). Die unter Abschnitt 3.3.2 identifizierten Fusionen sind als schwarze Linien hervorgehoben. Der innere Ring repräsentiert die Positionen der SD-Sequenzen, während der äußere Ring die DNA-Kopienzahländerung der Zelllinie Seax darstellt (rot-Deletionen, grün-Duplikationen). Die Positionen von Genen, welche mit ausgesuchten Annotationen assoziiert sind, sind innerhalb des Idiogramms in grün (GO-Annotation: „Lymphozyten“) und orange (bekannt als „Transkriptionsfaktoren der T-Zell-Entwicklung“) hervorgehoben.

3.3.5. Die 19p13.3-Deletion und chromosomale Interaktionen

Zur Analyse der chromosomalen Interaktionen mit der unter Abschnitt 3.2.2 beschriebenen tumorrelevanten 19p13.3-Deletion wurden die Bruchpunktintervalle der 20 untersuchten Tumorzellproben und der Seax-Zelllinie mit den Seax-Hi-C-Daten verglichen (Abb. 3.30). Interaktionshäufigkeiten der Binpaare wurden unter Abb. 3.30 als Heatmap dargestellt. Die 19p13.3-Deletion der Seax-Zelllinie spiegelt sich dabei in einer im Vergleich zu den angrenzenden Bereichen erhöhten Interaktionshäufigkeit innerhalb der deletierten Region wider (siehe Interaktionsmatrix Abb. 3.30). Durch die intrachromosomalen Interaktionen des Intervalls der Seax 19p13.3-Deletion werden drei topologische Domänen von dem TAD-Detektionsprogramm TADBit identifiziert (graue vertikale Linien in Abb. 3.30). Die zugrundeliegenden Hi-C-Daten der distalen TAD sind jedoch von nicht-annotierten Bereichen durchsetzt und erschweren so die Interpretation der Daten.

Der proximale Bruchpunkt der Seax 19p13.3-Deletion fällt mit einer der Seax-TAD-Grenzen zusammen. Vergleicht man die hier aufgrund der Seax-Hi-C generierten TADs mit publizierten Sub-TADs, zeigt sich eine Übereinstimmung der Grenzbereiche, auch wenn wie erwartet, sich die Dimensionen der TADs stark unterscheiden.

Ein Vergleich dieser TAD-Grenzen mit in der Literatur beschriebenen Sub-TADs verschiedener Zelllinien [333] zeigt eine Übereinstimmung der detektierten Seax-TAD-Grenzen mit Sub-TAD-Grenzen, wohingegen die Dimensionen der TADs mit den publizierten Sub-TADs, wie auch erwartet, nicht übereinstimmen. Ein Vergleich der TADs mit den Deletionen aus den Patientenproben und deren kleinste gemeinsam deletierte Region (*smallest region of overlap*, SRO) zeigen nur in einem Fall (Patient 3) eine Übereinstimmung mit den für Seax identifizierten TAD-Grenzen. Auch der visuelle Vergleich dieser Deletionsgrenzen mit den Sub-TAD-Grenzen der verschiedenen Zelllinien zeigt kaum Übereinstimmungen, sondern eine unabhängige Positionierung der Deletionsbruchpunkte von den Sub-TAD-Grenzen.

Die Region der Seax-19p13.3-Deletion zeigt keine signifikanten intra- oder interchromosomalen Interaktionen. Jedoch weist die Region stromabwärts des proximalen Bruchpunktes der 19p13.3-Deletion der Seax-Zelllinie interchromosomale Interaktionen zu Chromosom 17 auf und fusioniert mit dem distalen Bereich der 17q25.3-Region (Rearrangement t21, siehe Tab. 3.7). Ein mögliches Fusionsgen konnte dabei nicht identifiziert werden.

Da E2A, wie unter Abschnitt 3.2.3 behandelt, ein tumorrelevanter Transkriptionsfaktor

ist wurde untersucht, ob innerhalb der 2-Mb-Bruchpunktregionen durch E2A regulierte oder für Th₂-Lymphozyten relevanten Gene lokalisiert sind. Jedoch ist keines der durch E2A regulierten Gene (aus Tab. S5) innerhalb der Bruchpunktregionen lokalisiert.

Der *E2A*-Locus ist Teil der 19p13.3-Deletion und ist an der Grenze der mittleren Seax-TAD lokalisiert. Auch im Vergleich zu den Sub-TADs der weiteren Zelllinien ist *E2A* an der Grenze einer Sub-TAD positioniert. Eine Untersuchung, ob durch E2A deregulierte Gene (siehe Abschnitt 3.2.3) innerhalb der gesamtgenomischen signifikanten Interaktionen (identifiziert über Bins für die Fenstergröße 250 kb) vorliegen, konnte keine signifikante Anreicherung der 222 E2A-deregulierten Gene in den interagierenden Regionen feststellen (E2A-deregulierte Gene in Interaktionsregionen: 10 Gene, E2A-deregulierte Gene in Interaktionsregionen der randomisierten Datensätze: durchschnittlich 11 Gene, p -Wert = 0,7458; bei 10.000 Randomisierungen).

3.3. Chromosomale Interaktionen und Rearrangements der Tumorzelllinie Seax

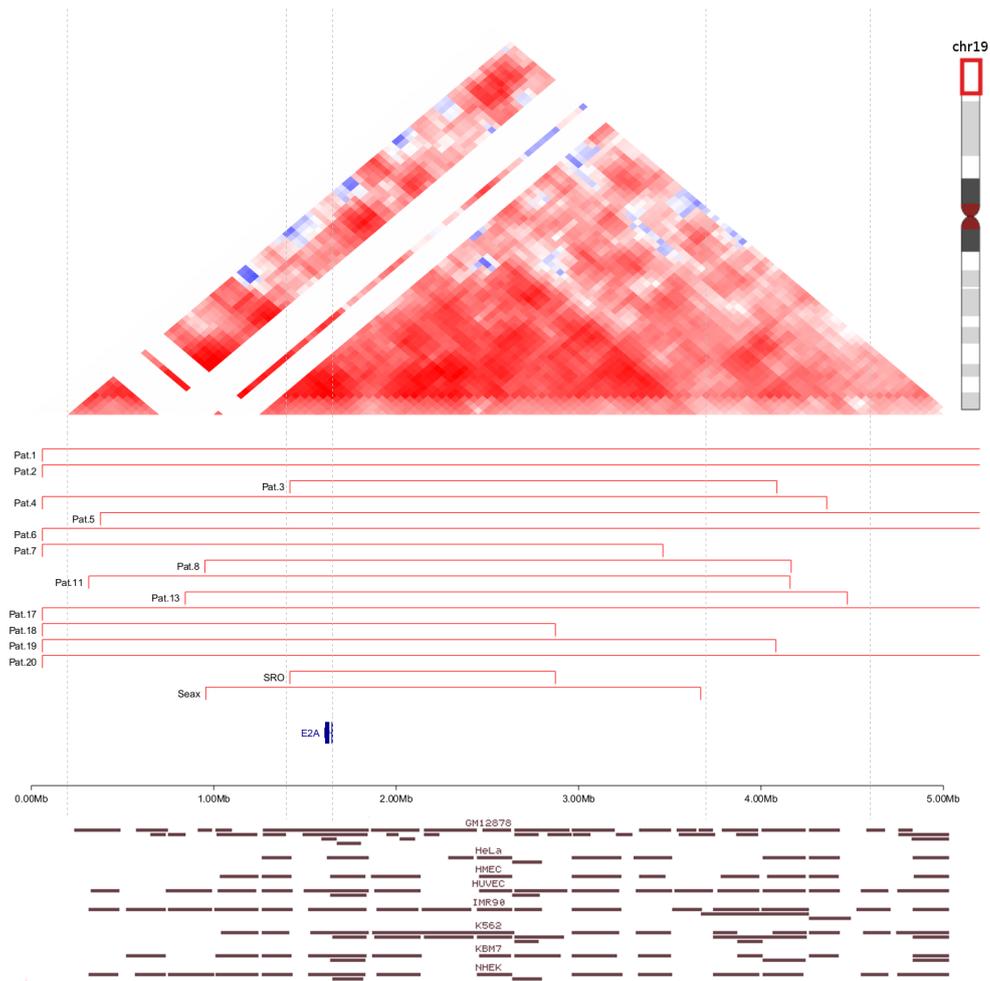


Abbildung 3.30.: TADs der 19p13.3-Deletion in Seax und weiteren Zelllinien.

Grenzen, der mit TADBit identifizierten TADs der Zelllinie Seax, sind als vertikale grau gestrichelte Linien repräsentiert. Oben: Normalisierte Heatmap der intrachromosomalen Interaktionen für die Region 19p13.3 (chr19: 0-5 Mb; hg19) der Zelllinie Seax. Nicht-annotierte Regionen in weiß. Mitte: Die Intervalle der Deletionen der Sézary-Tumorproben der Patienten 1 bis 20, deren Region der kleinsten gemeinsamen Deletion (*smallest region of overlap*, SRO) und die Seax 19p13.3-Deletionsregion. Unten: Die Intervalle der Sub-TADs der Zelllinien GM12878, HeLa, HMEC, HUVEC, IMR90, K562, KBM7 und NHEK aus [333].

4. Diskussion

Die Wahrscheinlichkeit für das Auftreten chromosomaler Aberrationen variiert zwischen genomischen Regionen. Einige Segmente weisen eine Häufung struktureller Variationen auf und zeigen damit eine verminderte chromosomale Stabilität [3]. Hierbei können die strukturellen Chromosomenveränderungen durch verschiedene Mechanismen entstehen. Strukturelle Variationen der Keimbahn, welche u. a. *Genomic Disorders* und evolutionäre Rearrangements hervorbringen, resultieren vermehrt über die nicht-allelische Rekombination homologer Sequenzen, während das Auftreten somatischer Veränderungen von Tumorzellen eher rekombinationsunabhängig erfolgt [114] [45].

4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen

Intrachromosomale Interaktionen des Chromosom 7 zeigen räumliche Syntenie

Strukturelle Veränderungen haben im Laufe der Evolution Chromosomen rearrangiert und zu Unterschieden im Genom verschiedener Spezies geführt. Von solchen Rearrangements war auch das Chromosom 7 während der Hominoidenevolution betroffen [287]. Hierbei haben zwei aufeinanderfolgende evolutionäre Inversionen die Organisation des Chromosom 7 verändert (Abb. 4.1) [287] [453]. Mittels integrativer Analysen eigener und publizierter Daten wurde die Genom- und Chromatinorganisation der evolutionär rearrangierten Segmente auf gemeinsame Charakteristika untersucht. Wie man in der zirkulären Ansicht von Chromosom 7 deutlich erkennen kann, zeigen die chromosomalen Interaktionen ein trianguläres Muster. Dieses Muster ist unabhängig von den gewählten Schwellenwerten verschiedener Filterkriterien und reflektiert jene chromosomale Regionen, welche über eine

4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen

Distanz von 10 oder mehr Mb interagieren. Durch die *in-silico*-Revertierung der zwei evolutionären Inversionen des Chromosom 7 wird die vermutete chromosomale Organisation der letzten gemeinsamen Vorfahren zwischen Mensch und Gorilla (perizentrische Inversion) bzw. Vorfahren von Mensch und Schimpanse nach Abspaltung des Gorillas (parazentrische Inversion) simuliert. Die im Zuge dieser *in-silico*-Revertierung re-positionierten Interaktionen, welche zuvor als trianguläres Muster angeordnet waren, konzentrieren sich danach innerhalb eines zusammenhängenden Segmentes (Abb. 3.4 A-C). Ein Vergleich mit Sequenzen von Neuweltaffen, deren letzter gemeinsamer Vorfahre mit Menschen vor den beiden Inversionen auftrat, zeigt, dass die drei beteiligten Segmentblöcke tatsächlich aus einer ursprünglich zusammenhängenden Region hervorgegangen sind. Im Beispiel des Weißbüschelaffen *C. jaccus* umfasst diese Region 17,9 Mb des Chromosom 2 (Abb. 3.4 D). Die Anhäufung legt nahe, dass dieses vormalig zusammenhängende Segment, trotz Auftrennung in mehrere Segmentblöcke im Laufe der Evolution, die chromosomalen Interaktionen untereinander beibehalten hat. Eine in der Evolution konservierte 3D-Organisation des Interphasekerns, trotz chromosomaler Rearrangements, wurde bereits in anderen Studien beobachtet. So wurde unter anderem gezeigt, dass die Genome verschiedener Spezies auch nach genomischer Umstrukturierung eine von der Gendichte abhängige 3D-Organisation einnehmen [292] und ursprünglich linear benachbarte Gene ihre räumliche Nähe im Zellkern beibehalten [84]. Ebenfalls räumlich nahe bleiben auch die Regionen der evolutionären Bruchpunkte innerhalb des Nukleus. Die daraus resultierende konservierte räumliche Nähe der evolutionär getrennten Segmente im Zellkern wird auch als räumliche Syntenie (*spatial synteny*) bezeichnet, was einer Aufrechterhaltung der chromosomalen Interaktionen zwischen funktionell zusammenhängenden Genen und deren regulatorischen Elementen dienen könnte [424]. Die evolutionären Bruchpunkte treten dabei bevorzugt in Regionen hoher transkriptioneller Aktivität auf [32]. Auch die interagierenden Regionen des untersuchten IMR90-Genoms zeigen eine Assoziation mit Eigenschaften von aktiven, euchromatischen Bereichen der DNA (Abb. 3.2). Es wird vermutet, dass die offene, dekondensierte Chromatinstruktur in einer erhöhten Instabilität der genomischen Region resultiert und diese damit anfälliger für chromosomale Brüche wird [225]. Die zusätzliche räumliche Nähe dieser Bereiche erhöht die Wahrscheinlichkeit eines nachfolgenden Rearrangements untereinander [32].

Ein möglicher Zusammenhang zwischen segmentalen Duplikationen und der Aufrechterhaltung der räumlichen Syntenie

Regionen evolutionärer Bruchpunkte können auch mit ähnlichen Sequenzeigenschaften assoziiert sein. So weisen Bruchpunktregionen verschiedener Spezies eine Anreicherung von SDs und repetitiver Sequenzen auf [238] [51] [69] [16] [192] [456] [114]. Eine Untersuchung der Position der SD-Sequenzen im Chromosom 7 ergab eine signifikante Anreicherung dieser Sequenzen in den intrachromosomal interagierenden Regionen. Entsprechend der allgemeinen Anreicherung an SD-Sequenzen in den gesamten intrachromosomal interagierenden Regionen des Chromosom 7, konnte diese Anreicherung auch innerhalb der Regionen festgestellt werden, welche zusätzlich zu *C. jaccus* syntenisch sind (Abb. 3.4 E). Weiterhin zeigt sich hier eine Anreicherung von *Alu*- und G4-Sequenzen (Abb. 3.4 F und G). Ein vermehrtes Vorliegen dieser Sequenzen innerhalb der räumlich konservierten chromosomalen Regionen kann entweder durch eine erleichterte Ausbreitung der SD- und repetitiven Sequenzen durch die räumliche Nähe erklärt werden oder durch eine mögliche Beteiligung dieser Elemente selbst an der Aufrechterhaltung der 3D-Chromosomenorganisation.

Dass repetitive DNA nicht einfach nur „parasitär“ und funktionslos vorliegt, sondern einen Einfluss auf die Genregulation und Chromatinorganisation ausübt, konnte in den letzten Jahren gezeigt werden (zusammengefasst in [212] [366]). Die genaue Bedeutung für die Chromatinorganisation ist dabei jedoch noch unklar. Heute wird von einem Modell der Selbstorganisation des Interphasechromosoms im dreidimensionalen Raum ausgegangen, wobei hauptsächlich die DNA-Sequenz die Konformation determiniert. Obwohl einige Proteine des Chromatins und deren Modifizierungen mit der Chromosomenorganisation assoziiert werden können, weist doch vieles auf eine direkte Beteiligung der DNA-Sequenzen hin [92] [33] [277] [279] [229] [78] [24] [79] [300]. So scheinen sich vor allem jene Chromosomenregionen im Nukleus aneinander zu lagern, die reich an repetitiven Elementen, insbesondere *Alu*-Repeats, sind. Dies konnte z. B. in humanen Lymphozyten, Fibroblasten und auch in den Retinazellen von nachtaktiven Tieren nachgewiesen werden [434] [380] [38]. Die Aneinanderlagerung von *Alu*-Sequenzen wurde hierbei bereits als Mechanismus der 3D-Organisation von Interphasechromosomen vorgeschlagen [202] [403] [80]. Anlagerungen zwischen den Sequenzen könnten dabei über Proteine realisiert werden, die repetitive Sequenzen binden wie z. B. CTCF oder *Transcription Factories* ermöglicht werden [402] [269] [80] [239]. Eine proteinunabhängige homologe Paarung von Sequenzen ist jedoch auch möglich und wurde bereits *in vitro* nachgewiesen [174] [86] [21] [207]. Die

4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen

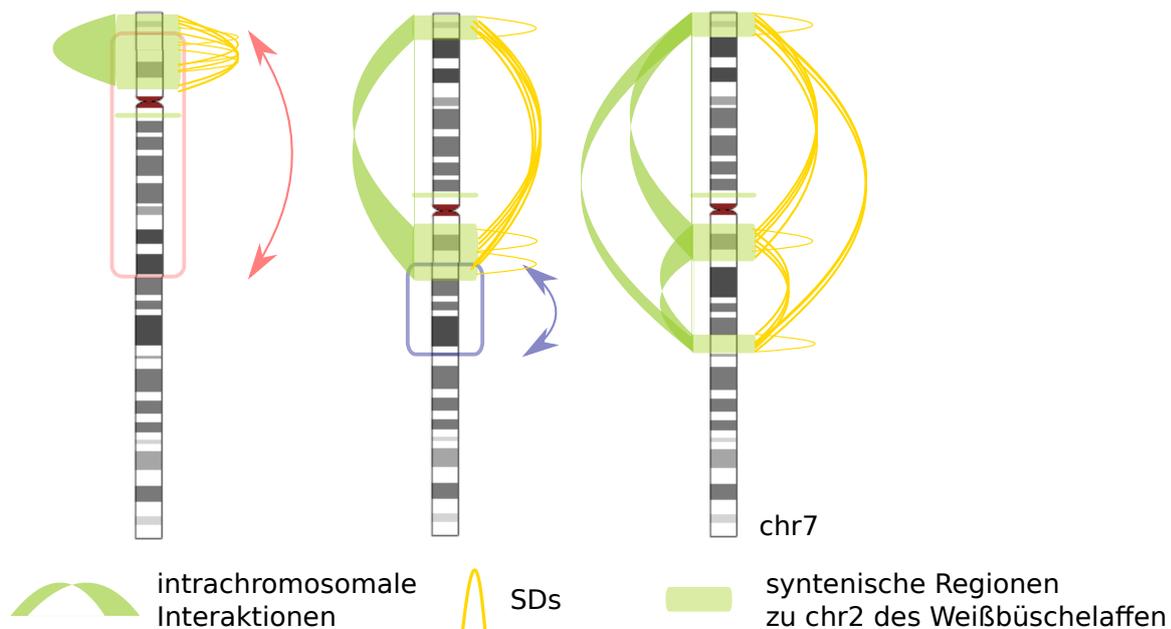


Abbildung 4.1.: Schema der evolutionären Inversionen des Chromosom 7 und der Folgen für die Chromosomenorganisation.

Schematische Darstellung der Organisation des Chromosom 7 vor den Inversionen (linke Seite), nach der perizentrischen Inversion (Mitte) und nach der parazentrischen Inversion (rechte Seite). Die Regionen der Inversionen werden durch einen roten (perizentrische Inversion) und blauen (parazentrische Inversion) Kasten hervorgehoben. Das mit *C. jacchus* chr2 syntenische Segment (linke Seite, grün markierte Bereiche des Idiogrammes) wird durch die Inversionen auf die Regionen 7p22, 7q11 und 7q22 verteilt. Über intrachromosomale Interaktionen bleiben diese Bereiche in räumlicher Nähe zueinander (räumliche Syntenie). Durch die Rearrangements werden ebenfalls die SDs dieses Segmentes innerhalb des Chromosoms aufgeteilt. In dieser Darstellung sind ausschließlich die entsprechenden intrachromosomalen Interaktionen und SDs schematisch wiedergegeben, welche in dem zu *C. jacchus* chr2 syntenischem Segment vorliegen. Ob die SDs tatsächlich bereits vor den Inversionen vorlagen, oder ob diese erst später inserierten, ist jedoch unbekannt.

bekannteste Paarung homologer Doppelstrangmoleküle tritt während der Schwesterchromatidpaarung der Meiose ein (zusammengefasst in [409]), aber auch in somatischen Zellen sind solche Phänomene bekannt und haben weitreichende funktionelle Konsequenzen z. B. in perizentrischen Regionen, bei der Telomerorganisation oder bei der homologen Rekombination [7] [146] [26] [393] [357]. Die Stabilität dieser DNA-Paarungskomplexe kann durch Proteine, aber auch über spezifische Wasserstoffbrückenbindungen, den Hoogsten-Basenpaarungen, vermittelt werden. Hoogsten-Basenpaarungen sind auch die Grundlage der Formierung von Quadruplexstrukturen der G4-Motive. Diese können nicht nur kurze intramolekulare Schleifen ausbilden, sondern über intermolekulare Bindungen transiente viersträngige Strukturen generieren. So wurde für Quadruplexstrukturen eine Rolle bei der Paarung der Schwesterchromatiden während der Meiose postuliert, wobei sich G4-Motive im Reißverschlussprinzip entlang der Chromosomen aneinander lagern [364] [35] [159].

Dieser in der Literatur beschriebene Zusammenhang zwischen der chromosomalen 3D-Organisation und *Alu*-Repeats bzw. G4-Motiven könnte darauf hinweisen, dass die homologe Paarung dieser Sequenzen auch zur Aufrechterhaltung der räumlichen Nähe von Regionen evolutionärer Rearrangements beiträgt. Unterstützt wird diese These durch die Veröffentlichung von Corurnac et al. [80], wo gezeigt wurde, dass Regionen mit über lange Distanzen hinweg wirkenden Interaktionen eine Anreicherung von evolutionär nah verwandten Retrotransposons, und Syntenieblöcke zwischen Mensch und Maus eine Kollokalisierung der repetitiven Elemente aufweisen [80]. Alternativ zu der Annahme, dass die repetitive DNA einen Anteil an der Chromosomenorganisation hat, könnte die Anreicherung auch die passive Folge einer höheren Insertionsrate in diesen Regionen sein. Diesen Überlegungen steht aber entgegen, dass *Alu*-Sequenzen sich bei der Verbreitung gleichmäßig im Genom verteilen und anschließend zügig aus GC-reichen Sequenzen entfernt werden. Dieser selektive Druck auf die Verteilung der repetitiven Sequenzen legt nahe, dass diesen auch eine Funktion zugeordnet werden kann [266] [28] [37] [147] [182].

Auch wenn repetitive Elemente auf diese Weise eine Funktion bei der Aufrechterhaltung der räumlichen Syntenie übernehmen könnten, bleibt offen, wie eine solche Kernorganisation ausreichend Sequenzspezifität erlangen könnte, da *Alu*-Elemente sehr häufig im humanen Genom vorkommen (entspricht 10 % des Genoms [429]) und G4-Motive kaum Sequenzdiversität aufweisen. Daher muss davon ausgegangen werden, dass eine Erhöhung der Sequenzspezifität der chromosomalen Interaktionen über eine Beteiligung weiterer DNA-Sequenzen, stabilisierender Proteine oder räumliche Einschränkungen der Chromo-

4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen

somenorganisation erfolgt [403].

Einen möglichen Ansatzpunkt für die Sequenzspezifität liefert die in dieser Arbeit gezeigte signifikante Anreicherung der SDs in intrachromosomal interagierenden Regionen. Im Vergleich zu *Alu*-Sequenzen weisen SDs eine deutlich geringere Verbreitung auf (5 % des humanen Genoms [19]). Daher könnten paraloge SDs eine kontrollierte und gerichtete chromosomale Partnerselektion während der Ausbildung der 3D-Organisation des Nukleus vermitteln. Das steht auch im Einklang mit der in der Literatur beschriebenen Beobachtung, dass sich SD-reiche Sequenzen ebenfalls im Nukleus zusammenlagern z. B. die subtelomerischen und perizentromerischen Regionen [233] [16] [393]. Weiterhin weisen paraloge SDs eine hohe Rate an Genkonversion untereinander auf (*interlocus gene conversion*) [97]. Dieser häufige Sequenzaustausch impliziert eine räumliche Nähe der paralogen SDs in der 3D-Organisation des Nukleus, da dies die Anlagerung und den Austausch der Sequenzen fördert. Unsere Analysen deuten darauf hin, dass die SDs nicht nur in intrachromosomal interagierenden Regionen des Chromosom 7 angereichert vorliegen, sondern dies auch gesamtgenomisch gilt (Abb. 3.3 und Tab. 3.1). Jedoch sind hierzu weitere Analysen nötig, da die angewandten Schwellenwerte hauptsächlich für Chromosom 7 überprüft worden sind.

Zusammenfassend legt die nichtzufällige Verteilung nahe, dass die Anreicherung der SD-, *Alu*- und G4-Sequenzen in den, trotz evolutionärer Rearrangements, räumlich nahen Segmenten des Chromosom 7 im Zusammenhang mit einer Funktion dieser Elemente bei der Aufrechterhaltung der räumlichen Syntenie steht. Die nicht zufällige Verteilung von SDs ist nicht nur aus der Perspektive der Chromosomenevolution interessant. Die SD-vermittelte Genominstabilität kann auch zu *Genomic Disorders* führen, die in der Humangenetik große Bedeutung haben [242]. Die Anreicherung von SDs in distinkten Regionen trägt somit auch zur Erhöhung der genomischen Instabilität dieser Regionen bei. Eine Funktion bei der 3D-Konformation der Chromosomen könnte eine solche destabilisierende Wirkung der SDs kompensieren.

Die für das Williams-Beuren-Syndrom kritische Region verfügt über eine distinkte Chromatinkonformation

Exemplarisch für die Untersuchung von chromosomalen Interaktionen im Zusammenhang mit *Genomic Disorders* wurde die Chromatinorganisation der 7q11.23-Region analysiert. Diese Region umspannt ein Segment (7q11-Segment), welches die Williams-Beuren-

Syndrom-assoziierte Region und die angrenzende 7q11.23-Deletionssyndrom-assoziierte Region beinhaltet. Dieses Segment wird durch drei große SD-Cluster geprägt, welche die CNV-Regionen flankieren, und deren nicht-allelische homologe Rekombinationen für WBS bzw. 7q11.23-Deletionssyndrom verursachende Deletionen verantwortlich sind. Die Chromatinarchitektur des 7q11-Segments wurde über eine integrative Analyse mit publizierten Daten zu u. a. den intrachromosomalen Interaktionen der Zelllinie IMR90 [92] und eigenen Daten der Chromatinorganisation untersucht. Dabei konnte demonstriert werden, dass innerhalb der flankierenden SDs eine spezifische distinkte Chromatinorganisation, eine Chromatindomäne, vorliegt, die sich von den angrenzenden Regionen außerhalb des 7q11-Segments unterscheidet. Das 7q11-Segment hat Gemeinsamkeiten mit *regions of increased gene expression* (RIDGEs) [52] [425]. Dabei handelt es sich um Cluster von transkriptionell hochaktiven Genen mit u. a. kurzen Introns und einem hohen GC- bzw. *Alu*-Anteil [425]. Publierte Analysen der RIDGEs schließen auf eine weniger kondensierte Chromatinkonformation dieser Regionen, was zu der Ausbildung einer subnuklearen Umgebung einer erhöhten Transkription führen könnte [259] [138]. Dies wird durch Untersuchungen unterstützt, dass unterwundene Regionen ebenfalls die Charakteristika, wie einen hohen GC-Anteil und transkriptionelle Aktivität, aufweisen [290]. Die Aufrechterhaltung einer solchen Konformation kann laut Literatur durch G4-Motive erreicht werden [96], die auch in diesem Segment angereichert vorliegen. Das untersuchte 7q11-Intervall wird proximal begrenzt durch den Übergang von Hetero- zu Euchromatin (Abb. 3.6 und Abb. S5) und distal durch eine potentielle Assoziation dieser Region mit der Kernlamina. Letzteres wird durch die erhöhten Lamin-B1-Signale der distalen Grenzregion angenommen, was jedoch auch das Resultat technischer Artefakte aufgrund des Auftretens von SD-Clustern in dieser Region sein könnte (Abb. 3.6 M). Dass die WBS-assoziierte Region als eine abgegrenzte Chromatindomäne vorliegt, zeigt sich auch in den ausgewerteten Hi-C-Daten der intrachromosomalen Interaktionen, welche die Existenz einer TAD nahelegen, die an den Grenzen von SDs flankiert ist. Ein Vergleich mit den bereits publizierten genomweiten Daten von TAD-Regionen [92] bestätigt die hier identifizierten TAD-Grenzen. Diese Domänengrenzen werden durch die SD-Cluster flankiert. Da die Definition von TADs durch die hohen Sequenzsimilaritäten der SD-Cluster verfälscht sein könnte, musste nach Möglichkeiten gesucht werden, technische Artefakte auszuschließen. Deshalb wurde der Umstand genutzt, dass TAD-Grenzen in Maus und Mensch hoch konserviert sind und dass dem murinen Genom in der korrespondierenden Region diese SD-Cluster fehlen, welche im humanen Genom zu Alignmentproblemen führen könnten [92] [9] [18] [15]. Tatsächlich stimmen auch im untersuchten 7q11-Segment die TAD-Grenzen mit denen der

4.1. Strukturelle Chromosomenveränderungen der Keimbahn und chromosomale Interaktionen

entsprechenden syntenischen Region der Maus überein (Abb. 3.7). Somit konnten die hier identifizierten TAD-Grenzen in einem Organismus bestätigt werden, in dem Alignmentprobleme aufgrund der SD-Cluster während der Datenauswertung ausgeschlossen werden können. Gleichzeitig zeigte sich durch diesen artübergreifenden Vergleich auch, dass die SD-Cluster sich im Laufe der Evolution an den der Maus homologen TAD-Grenzen inseriert haben. Auch bei *Drosophila* konnte eine Anreicherung von transgenen P-Elementen innerhalb TAD-flankierender Regionen bereits gezeigt werden [161]. Dies könnte ein Hinweis auf eine höhere Vulnerabilität von TAD-Grenzen und den Sequenzen zwischen zwei Domänen sein. Das legt die Vermutung nahe, dass auch im Fall des 7q11-Segments die topologische Organisation der DNA zur präferentiellen Insertion dieses klinisch relevanten SD-Clusters beigetragen hat.

Der mögliche Einfluss von SDs auf die lokale Chromatinorganisation

Auffällig an den flankierenden SDs der WBS-Region ist ihre hohe Sequenzähnlichkeit. Durch die Sequenzähnlichkeit steigt auch die Wahrscheinlichkeit eines NAHR-Rekombinationsereignisses, was zu der Instabilität der Region beiträgt. Diese Eigenschaft teilen sie mit weiteren *Genomic-Disorder*-assoziierten SDs (z. B. DiGeorge-Syndrom, Sotos Syndrom), während die Sequenzen vieler anderer SDs nach ihrer Insertion relativ schnell divergieren [10] [166] [313] [168] [115] [149] [155] [98] [360]. Diese Konservierung der Sequenzähnlichkeit entsteht durch eine hohe Rate an Genkonversion zwischen den flankierenden SDs und spricht für eine räumliche Nähe der Sequenzen, da diese den Sequenzaustausch begünstigt [166] [97] [115]. SDs könnten damit die Basis für genomische Interaktionen in der Chromatinorganisation bilden, was z. B. durch die intramolekulare somatische Paarung realisiert werden könnte [450]. Eine Folge der Insertion von SDs zwischen den TADs könnte auch die Einbringung oder Positionsveränderung eines Insulatorelementes sein. Wie auch die SDs des 7q11-Segments sind SDs allgemein häufig an den Übergängen von Eu- zu Heterochromatin lokalisiert und markieren dabei den Wechsel der Chromatinstruktur und -funktion [87] [141] [200] [159]. Eine Funktion der SDs als Trennelement zwischen Regionen unterschiedlicher Aktivität wurde bereits für SDs diskutiert, die in perizentromerischen Regionen vorliegen und am Übergang zwischen heterochromatischen Zentromerregionen und euchromatischen, kodierenden Regionen lokalisiert sind [102] [136]. Die funktionelle Relevanz von insulatorischen Elementen an den TAD-Grenzen zeigen Deletionsexperimente, bei denen es nach deren Entfernung zu einer aberranten Genexpression kommt. Diese Veränderungen können sogar klinische Relevanz haben [299] [240] [171].

Auswirkungen von Deletionen der WBS-Region auf die Organisation benachbarter TADs wurden bereits ermittelt, wobei der Verlust der WBS-Region zu einer Änderung angrenzender Chromatininteraktionen führte [135].

Ein gezieltes Entfernen der flankierenden SD-Sequenzen mittels z. B. der CRISPR/Cas-Technik [132] in Kombinationen mit 3C-basierten Methoden (wie auch schon im Zusammenhang mit anderen Loci beschrieben worden ist [240]) könnte weitere Aufschlüsse über den Einfluss der SDs auf die WBS-Region und die Änderungen der Chromatinorganisation geben.

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

Die unter Abschnitt 4.1 diskutierten rekurrenten strukturellen Chromosomenveränderungen der Keimbahn, wie die pathogenen Kopienzahlveränderungen der *Genomic Disorders*, beruhen hauptsächlich auf homologie-basierte Rekombinationssmechanismen [384]. Hingegen werden die erworbenen Veränderungen der somatischen Zellen während der Tumorgenese vorwiegend über mikrohomologie-basierte oder rekombinations-unabhängige Prozesse wie NHEJ verursacht [144] [307]. Die resultierenden Muster an chromosomalen Veränderungen sind jedoch nicht zufällig im Genom verteilt. Auch Tumorkaryotypen zeigen rekurrente Veränderungen, deren Identifikation und Feinkartierung Einsichten in die Entstehung chromosomaler Veränderungen bringen und ein wichtiger Startpunkt für die Suche nach Tumorgenen ist [296] [347]. Die klinische Relevanz solcher Untersuchungen zeigt sich anhand der ersten identifizierten rekurrenten Chromosomenveränderung, dem Philadelphiachromosom. Es konnte hierbei gezeigt werden, dass Patienten mit chronischer myeloischer Leukämie (CML), die diese Translokation aufweisen, einer effektiven Chemotherapie mit Imatinib unterzogen werden können, was der aberranten Funktion des Fusionsgens entgegenwirkt [281]. Diese chromosomale Veränderung ist ein Paradebeispiel wie solche Analysen zu einer verbesserten Diagnose und zur Entwicklung neuer Therapien führen kann. Methodisch bedingt ist jedoch besonders die Feinkartierung von Translokationen schwierig. Heute werden Translokationen vornehmlich über NGS-basierte Methoden identifiziert, bei der die vom Referenzgenom abweichende Positionierung von Readpaaren auf chromosomale Rearrangements hinweist [68] [66]. Nachteil der bisher verwendeten Methoden ist, dass die Detektion der Bruchpunktregionen durch z. B. re-

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

petitive Sequenzelemente erschwert oder sogar verhindert wird. Da das humane Genom jedoch zu fast der Hälfte aus repetitiven Sequenzen besteht, zeigen diese Methoden eine hohe Fehleranfälligkeit [410]. Eine alternative Methode zur Detektion von chromosomalen Rearrangements sind Abwandlungen der 3C-Technik. Diese können über die räumliche Nähe von rearrangierten chromosomalen Regionen strukturelle Chromosomenveränderungen wie Translokationen oder Inversionen identifizieren [47] [125] [25] [108] [215] [333]. Am Beispiel der Seax-Zelllinie soll nun gezeigt werden, dass auch komplexe Karyotypen mittels der Hi-C-Technik studiert werden können.

Die Identifizierung von strukturellen Veränderungen über chromosomale Interaktionen

Die Analyse der Translokationen der Zelllinie Seax erfolgte in dieser Arbeit über Hi-C, einer Methode zur Identifizierung der Chromosomenkonformation [229]. Die Hi-C-Technik diente ursprünglich der Aufklärung der 3D-Organisation der Chromosomen im Nukleus. Bei der Methodik wird davon ausgegangen, dass die Interaktionsfrequenz zwischen zwei gegebenen Regionen abhängig von der genomischen Distanz zueinander ist [229]. In einem nicht von Rearrangements betroffenen Karyotyp ist daher eine chromosomale Interaktion zwischen Regionen verschiedener Chromosomen im Vergleich zu der Menge an intrachromosomalen Interaktionen selten. Zwei ursprünglich auf verschiedenen Chromosomen lokalisierte Segmente, die durch ein Rearrangement fusionieren liegen nach der Fusion aneinander angrenzend intrachromosomal vor, werden jedoch durch die Analyse entsprechend dem Referenzgenom als interchromosomal bewertet. Diese Nähe wird durch hohe Interaktionshäufigkeiten der beteiligten Regionen repräsentiert. Über die resultierenden hohen Interaktionshäufigkeiten zwischen den beiden Fusionsregionen können nun auch die Bruchpunkte der Rearrangements bestimmt werden. Bisher wurden bereits 3C-basierte Methoden zur Bestimmung von strukturellen Chromosomenveränderungen eingesetzt [373]. Allerdings waren diese Methoden auf die Identifizierung einiger Translokationspartner beschränkt. Hingegen bietet Hi-C eine genomweite Technik und wurde daher schon als Alternative zu bisherigen Methoden der Detektion von strukturellen Chromosomenveränderungen besprochen [47] [125] [25] und angewendet [108].

Für die Analyse des komplexen Seax-Karyotyps wurden mithilfe der Hi-C-Technik zunächst jene genomischen Regionen bestimmt, die hohe interchromosomale Interaktionshäufigkeiten aufweisen. Dabei zeigen rearrangierte Regionen in den Heatmapdarstellungen

der Interaktionsmatrizen ein spezifisches Muster. Über die Position und Orientierung dieses Musters lassen sich die von den Rearrangements betroffenen Regionen bestimmen. Durch die Readverteilung innerhalb dieser Regionen können dann die Bruchpunkte mit einer Genauigkeit von wenigen Kilobasen kartiert werden. Insgesamt konnten 22 interchromosomale Fusionen der Seax-Zelllinie identifiziert werden. Durch das hohe Auflösungsvermögen konnten auch jene Gene bestimmt werden, die direkt von der Fusion betroffen sind. Zur Verifizierung der Hi-C-Daten wurden diese mit dem durch Iżykowska et al. beschriebenen derivaten Chromosom der(3)t(3;6;17;9) verglichen. Dieses komplex rearrangierte Chromosom der Zelllinie Seax beinhaltet unter anderem ein Segment des Chromosom 6q, welches auch in Tumorgenomen von Sézary-Patienten strukturellen Chromosomenveränderungen unterworfen ist [178]. Diese über die *paired-end whole genome*-Sequenzierung ermittelten Bruchpunktregionen des beschriebenen derivaten Chromosoms konnten über die Interaktionsmatrizen des Hi-C-Experiments nachvollzogen werden. Neben den Rearrangements zwischen den Chromosomen chr3, chr6, chr17 und chr9, konnte über die Hi-C-Interaktionsmatrizen eine Reihe weiterer Fusionen detektiert werden.

Auf Ebene der Fusionsgene wurde in [178] das *AIG1-GOSR1* beschrieben, welches von den Autoren auch als Fusionstranskript in Seax nachgewiesen werden konnte. Anhand dieses Fusionsgens soll die Überprüfung der Hi-C-Daten erfolgen. Diese Fusion vereint die Regionen 6q24.2 und 17q11.2 und ist Teil des beschriebenen derivaten Chromosoms. Über Readverteilungen innerhalb dieser Regionen konnte mittels des Hi-C-Experimentes ein Fusionsgen *AIG1-GOSR1* auf genomischer Ebene bestätigt werden. Die Hi-C-Daten konnten über Interaktionsmatrizen und Readinteraktionsverteilungen jedoch noch ein zusätzliches Rearrangement der beiden Gene identifizieren. So liegt hier eine reziproke Translokation vor, die auch ein *GOSR1-AIG1*-Fusionsgen möglich macht (Abb. 4.2). Dabei könnte, im Gegensatz zu dem unter [178] beschriebenen Fusionsgen, *GOSR1* die 5'-Region stellen und mit einem intragenischen Segment des *AIG1*-Gens fusionieren.

Die Daten des Hi-C-Experimentes sind also in der Lage zusätzliche Informationen über chromosomale Rearrangements zu liefern. Das ist möglich, weil eine Fusionsregion durch Reads des gesamten translozierten Segmentes gestützt wird. Hingegen identifizieren Ansätze über *whole-genome*-Sequenzierung oder RNA-Seq die Translokationen vornehmlich über die abweichende Position von *paired-end*-Readpaaren. Erwartet werden hier Readpaare, welche einer zusammenhängenden chromosomalen Region zugeordnet werden können. Durch das Rearrangement treten jedoch chimäre Readpaare aus unterschiedlichen Chromosomen auf, deren Distanz und Orientierung zueinander Rückschlüsse auf das Vor-

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

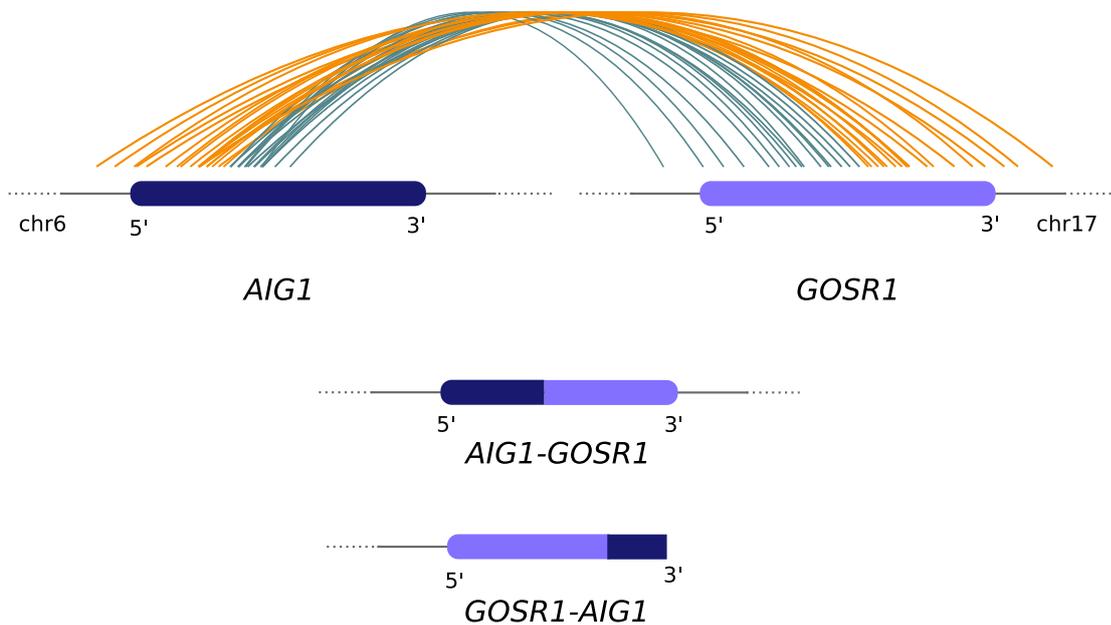


Abbildung 4.2.: Reziproke Translokation zwischen *AIG1* und *GOSR1*.

Schema der interchromosomalen Interaktionen zwischen den Loci der Gene *AIG1* und *GOSR1*. Die Hi-C-Daten lassen eine reziproke Translokation vermuten. Dabei fusioniert, wie auch in [178] beschrieben, 5'-*AIG1* mit 3'-*GOSR1* (repräsentiert durch gelbe Interaktionen). Zusätzlich fusioniert 5'-*GOSR1* mit einem intragenischen *AIG1*-Segment (repräsentiert durch blaue Interaktionen). Das *GOSR1-AIG1* Fusionsgen liegt also verkürzt vor.

handensein chromosomaler Rearrangements erlaubt [303]. Ein großer Nachteil hierbei ist, dass die eigentliche Fusionsregion von nur wenigen Readpaaren umspannt wird. Repetitive Elemente oder eine zu geringe Sequenziertiefe innerhalb der Bruchpunktregionen können dabei die Identifizierung erschweren und zu einer fehlerhaften Definition von Fusionsregionen führen [312] [1]. Strukturelle Chromosomenveränderungen sind jedoch häufig nahe dieser repetitiven Sequenzen positioniert [8]. Dieses Problem versucht man über die Anwendung verschiedener Algorithmen bei der Auswertung der *whole-genome*-Sequenzierung auszugleichen (u. a. [222]).

Die Nutzung von Hi-C zur Bestimmung der Bruchpunkte könnte eine Alternative zu den heute vornehmlich genutzten Methoden sein. Vorteil dabei ist, dass eine im Vergleich zu der *whole-genome*-Sequenzierung geringe Readabdeckung ausreicht um Bruchpunkte zu identifizieren. Ein weiterer Vorteil ist, dass Bruchpunktregionen mit repetitiven Elementen durch Readinformationen der angrenzenden Regionen unterstützt werden.

Nachteil der Hi-C-Technik ist, dass Bruchpunkte nicht basenpaargenau identifiziert werden, sondern auf Ebene von *HindIII*-Fragmenten (ca. 4 kb). Daher müssen die genauen Bruchpunkte mit weiteren Methoden bestimmt bzw. verifiziert werden: z. B. über Bruchpunkt-umspannende PCR. Weiterhin kann über diese Hi-C-Analyse, ähnlich der Alternativmethoden, keine Unterscheidung zwischen den homologen Chromosomen oder einzelnen Zellen der Zellpopulation gemacht werden. Entsprechende Ansätze über Einzelzell-Analysen oder Haplotypisierungen über die Hi-C-Methodik sind jedoch schon beschrieben worden [363] [289]. Weiterhin ist es möglich, dass nicht alle detektierten Interaktionen auf direkte Rearrangements der häufig miteinander interagierenden Segmente zurückzuführen sind, sondern diese können auch durch kleine Zwischensegmente voneinander getrennt vorliegen. So könnten Bruchpunkte, die in der Interaktionsmatrix mit mehreren Regionen häufig interagieren nicht direkte Fusionspartner, aber benachbarte Segmente eines derivativen Chromosoms repräsentieren. Daher ist es notwendig eine zusätzliche Kartierung der Segmentanordnung vorzunehmen, möglicherweise generiert über den direkten Vergleich der Readhäufigkeiten, was jedoch eine Erhöhung der Sequenziertiefe erfordert.

Translokationsbruchpunkte und deren potentielle Fusionsgene

Es wird davon ausgegangen, dass die Wahrscheinlichkeit einer Translokation zwischen zwei Regionen mit ihrer räumlichen Nähe im Nukleus steigt. Dies ist z. B. in *Transcription Factories* gegeben, in denen Gene verschiedener Loci zusammenkommen und gemeinsam

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

transkribiert werden. Dementsprechend könnte die Translokation zwischen Genen auf eine Koregulation und damit auf eine mögliche gemeinsame Funktion hinweisen [341]. Die Gene, welche entsprechend der Hi-C-Ergebnisse direkt von Translokationsbruchpunkten betroffen sind (siehe Tab. 3.7), zeigen in der durchgeführten GO-Analyse jedoch keine signifikante Anreicherung gemeinsamer molekularer Funktionen oder biologischer Prozesse. Auch eine entsprechende Untersuchung aller Gene, die innerhalb der von den Translokationen betroffenen TADs vorliegen, konnte keine gemeinsame Assoziation mit molekularbiologischen Prozessen detektieren. Somit ergibt die GO-Analyse keinen Hinweis auf eine Organisation der von den Translokationen betroffenen Genen in gleichen biologischen Prozessen.

In der Literatur sind die Funktionen der hier identifizierten Bruchpunktgene in Zusammenhang mit tumorrelevanten Prozessen wie der Chromosomenstabilität bzw. DNA-Schadensreparatur (*NIPBL*, *RPA2*, *HMGB1* [23] [106] [219] [235]), Th₂-Differenzierung (*CIITA* [308] [58]), Inflammation (*HK1* [283]), Transkription (*TAF3* [221]), ECM bzw. Zell-Adhäsion (*CTSK*, *TJP3* [422] [324]) und Zellzyklusprogression (*NAA50*, *SCTR* [185] [189]) beschrieben worden.

Die Genprodukte von *HMGB1* (*high-mobility group box-1*) und *RPA2* (aus unabhängigen Translokationen) formieren laut Literatur direkt einen Komplex während der DNA-Reparatur. Dieser HMGB1-RPA2-Komplex erkennt Läsionen in der DNA, welche z. B. über Quervernetzungen durch Psoralen in die DNA eingebracht werden. Psoralen ist ein DNA-interkalierender sekundärer Pflanzenstoff, welcher in Kombination mit UVA-Bestrahlung (PUVA) auch in der CTCL-Therapie angewandt wird [335] [235]. Der HMGB1-RPA-Komplex ist in der Lage, die resultierenden DNA-Strukturen zu identifizieren und anschließend Proteine der *Nucleotide excision repair* (NER)-Reparatur zu rekrutieren [219] [235]. Dementsprechend weisen HMGB1-defiziente Zelllinien eine hohe UV-vermittelte Photo-Sensitivität und geringere Zellüberlebensrate bei der Behandlung mit Psoralen auf [218]. Defizienzen von HMGB1 oder RPA2 werden daher auch mit Chromosomeninstabilität und Chromothripsis assoziiert [408] [137]. Eine direkte Assoziation von HMGB1 mit T-Zell-Lymphomen konnte bereits nachgewiesen werden [251] [126].

Auch die Expression benachbarter Gene kann von Translokationen durch resultierende dislozierte regulatorische Elemente oder veränderter 3D-Konformation beeinflusst werden [153]. So sind essentielle Gene der Lymphozytenmigration (*CXCR4*) [309], T-Zell-Differenzierung (*PRDM1* und *GATA3*) [36] [428] und T-Zellintegrität/Zytokinsignaling

in der Immunantwort (*HDAC2*) [104] [93] [208] nur wenige hundert kb von Bruchpunkten lokalisiert.

Insgesamt konnten sechs potentielle Fusionsgene mit 5'-3'-Orientierung durch die *in silico* Analyse der Hi-C-Daten auf Translokationen identifiziert werden. Die möglichen Fusionsgene mit einer 5'-3'-Orientierung sind *CIITA-CTSK* (t2), *NAA50-GOSR1* (t8), *SLC39A11-KCTD16* (t13), *AIG1-GOSR1* bzw. *GOSR1-AIG1* (t15), *GRM8-HK1* (t16) und *HMGB1-BRSK2* (t19). Zum Beispiel fusioniert das Rearrangement t19 zwischen den Regionen 13q13.3 und 11p15.5 möglicherweise die Gene *HMGB1* und *BRSK2* in einer 5'-3'-Orientierung miteinander. Die 5'-Region wird dabei durch das *HMGB1*-Gen gestellt, welches einen der am höchsten exprimierenden Promotoren aufweist [103]. *HMGB1* bricht im Intron 1 des *HMGB1-012* Transkriptes (ENST00000405805) und würde durch eine Translokation somit sämtliche Eigenschaften seines Proteins verlieren. Die *HMGB1*-5'-UTR-Sequenzen fusionieren mit Sequenzen des ersten Exons und des restlichen 3'-Bereiches des *BRSK2*-Gens, dessen Genprodukt eine AMPK-verwandte Kinase darstellt. Durch die möglicherweise ektopische Expression der fusionierten Kinase könnte es zu einer Deregulation der assoziierten Signalwege kommen.

Da keine Beschreibung der identifizierten Fusionsgene in der Literatur vorliegt (bis auf *AIG1-GOSR1* in Seax-Zellen durch [178]), scheinen diese nicht rekurrent in Tumorentitäten vorzukommen. Der MHC II-Masterregulator *CIITA* ist jedoch ein häufiger Translokationspartner in lymphatischen Tumoren, wobei 3'-Partnergene überexprimiert werden, und das Auftreten eines *CIITA*-Rearrangements einen prognostischen Indikator für eine niedrigere Überlebensrate repräsentiert [388] [413]. In der Literatur finden sich somit Hinweise auf die Assoziation der von den Bruchpunkten direkt betroffenen Gene mit tumorrelevanten Funktionen, wie der DNA-Schadensreparatur. Inwieweit diese und die identifizierten potentiellen Fusionsgene tatsächlich an der Tumorgenese beteiligt sind, müssen weiterführende funktionelle Analysen klären.

Genom- und Chromatineigenschaften der Bruchpunktregionen

Die Bruchpunktregionen der Rearrangements wurden mit allgemeinen genomischen Eigenschaften, sowie Histonmodifikationen und dem DNA-Kopienzahlstatus der Seax-Zelllinie verglichen. Dabei zeigt sich, dass die Bruchpunktregionen häufig benachbart zu Deletionen auftreten. Studien haben bereits nachgewiesen, dass Translokationen oft Teil komplexer Rearrangements sind und von Deletionen oder Duplikationen begleitet werden, die mehrere Mb umspannen können [204] [76] [162]. Besonders in hämatologischen Tu-

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

moren werden Translokationen von submikroskopischen Deletionen flankiert, z. B. bei der *BCR-ABL*-Translokation des Philadelphiachromosoms, was sich in einigen Studien negativ auf den Krankheitsverlauf und Einfluss der Therapie der Patienten ausgewirkt hat [374] [204] [165] [11] [14] [282]. Innerhalb der Translokations-begleitenden Deletionen konnte z. B. eine hohe Dichte von *Alu*-Sequenzen detektiert werden, was darauf hinweist, dass eher genreiche Segmente von den Deletionen betroffen sind [204] [205]. Über einige dieser flankierenden Deletionen werden die Translokationen der Seax-Zelllinie miteinander verbunden. So sind die Bereiche zwischen den Translokationen t1_4p16.3/t10_4p16.3, t2_16p13.13/t10_16p13.13, t3_2q14.2/t4_2q22.1, t3_6q16.1/t4_6q21 (oder t7_6q21) und t14_17q24.3/t13_17q25.1 (oder t21_17q25.3) deletiert. Einige Translokationsloci sind auch an Duplikationsbruchpunkten lokalisiert, jedoch konnten keine verbindenden Duplikationen ähnlich der Deletionsbrücken zwischen Translokationen gefunden werden. Weiterhin können die stromaufwärts bzw. stromabwärts liegenden Regionen eines Bruchpunktes an verschiedenen Translokationen beteiligt sein und über eine Rearrangementkette miteinander verbunden sein. Zum Beispiel wird zwischen den Chromosomen chr5, chr8 und chr10 eine solche Kette aus balancierten Rearrangements über die Translokationen t11_5p13.2/t12_5q13.2, t11_8p23.1/t17_8p23.1 und t12_10q23.33/t17_10q23.33 gebildet. Das Auftreten von 1) Translokations-verbindenden Deletionen und 2) Rearrangementketten zwischen Chromosomen sind Merkmale von Chromoplexy. Dabei handelt es sich um ein Muster komplexer chromosomaler Veränderungen, bei dem Segmente verschiedener Chromosomen rearrangiert werden und geschlossene Rearrangementketten mit kolokalisierenden Translokationsbruchpunkten bzw. Deletionsbrücken auftreten [13]. Das chromosomale Aberrationsmuster der Zelllinie Seax zeigt solche Rearrangementketten zwischen den Chromosomen 1) chr5, chr8, chr10, 2) chr1, chr4, chr16 und 3) chr2, chr3, chr6, chr9, chr17. Beschrieben wurde Chromoplexy u. a. bei Prostatakrebs und Melanomen [13], konnte aber auch kürzlich in CTCL gezeigt werden [74]. Bei Prostatakrebs mit Fusionen des Transkriptionsfaktorgens *ETS* sind die Rearrangementbruchpunkte geschlossener Rearrangementketten des Chromoplexy-Genoms in der Nähe von offenem, transkriptionell aktivem Chromatin positioniert [13]. Vermutet wird hier, dass die Kolokalisation von verschiedenen chromosomalen Loci in *Transcription Factories* zum Zeitpunkt des Auftretens von DNA-schädigenden Ereignissen die komplexen Rearrangements der Chromoplexy zur Folge hat. Auch die Translokationsbruchpunkte der Seax-Zelllinie zeigen eine Anreicherung von mit aktivem Chromatin assoziierten H3K4me3-Histonmodifikationen, insbesondere in dem Bruchpunkt benachbarten nicht-transloziertem Bereich. Das deckt sich mit Beobachtungen anderer Studien wo bereits nachgewiesen werden konnte, dass Gene

mit einer Anreicherung an aktivierenden Histonmodifikationen an den Promotoren anfälliger für Doppelstrangbrüche und Translokationen sind. Dabei konnte auch gezeigt werden, dass die Destabilisierung nicht aufgrund eines erhöhten Transkriptionslevels erfolgte, sondern möglicherweise durch die erhöhte Zugänglichkeit der DNA [72] [201] [46] [286] [260] [2]. Auch die Endonukleasen der V(D)J-Rekombination und des Immunglobulin-Klassenwechsels lösen bevorzugt an H3K4-methylierten Intervallen Doppelstrangbrüche und damit Rekombinationsereignisse aus [85] [368] [386] [180].

Nachfolgend wurde exemplarisch untersucht, ob an den Bruchpunkten der chromosomalen Veränderungen bestimmte Genomeigenschaften gehäuft auftreten. Zunächst wurde die Verteilung von nicht-pathogenen strukturellen Variationen der DGV-Datenbank, *Genomic Disorders*, *Fragile Sites* und evolutionären Bruchpunkten überprüft. Auch wenn es fallweise Übereinstimmungen mit *Fragile Sites* und in *Genomic Disorders* betroffenen Intervallen gab, so war dies keinesfalls charakteristisch für die große Mehrheit der Bruchpunkte. Ein Vergleich des DNA-Kopienzahlstatus der Bruchpunktregionen in den 20 Sézary-Patienten zeigt, dass 29 der Bruchpunkte mindestens bei einem der Patienten eine Deletion oder Duplikation aufweisen. Die höchste Übereinstimmung hat hierbei der Bruchpunkt des Rearrangements t21 auf Chromosom 19p13.3 mit 10 Patienten die hier eine Deletion aufweisen. Dies spricht für eine in Sézary-Genomen eher instabile Region. Ähnlich geringe Übereinstimmung fand sich auch zwischen den Bruchpunkten und den Grenzen von Chromatindomänen. Obwohl diese Grenzregionen potentiell anfälliger für Chromosomenbrüche sind, lagen die meisten Bruchpunkte in den inter-LAD Regionen bzw. nicht in der Nähe von TAD-Grenzen.

Schließlich wurde auch eine mögliche Assoziation der Bruchpunktregionen mit Sequenzen überprüft, welche die Ausbildung von non B-DNA Konformationen fördern. Keines der untersuchten Sequenzmerkmale, G4-Motive, *Mirror Repeats* oder zDNA fördernde Basenabfolgen, war in Bruchpunktnähe angereichert. Um den Einfluss rekombinations-basierter Mechanismen auf die Entstehung der chromosomalen Veränderungen abzuschätzen, wurde die Anwesenheit von *Alu*-Repeats und SDs an den Bruchpunkten überprüft. In Übereinstimmung mit der Literatur zu somatischen Chromosomenveränderungen konnte auch bei den hier untersuchten Bruchpunkten keine Häufung von SDs oder *Alu*-Repeats gefunden werden. Innerhalb der Bruchpunktregion konnten zwar paraloge SDs und gepaarte *self chains* ermittelt werden, allerdings traten diese zwischen Bruchpunktregionen nicht fusionierender Regionen auf, was gegen eine Beteiligung dieser Sequenzen an den Rearrangements spricht. Ein weiterer wichtiger Faktor für die Entstehung rekurrenter Chromosomenveränderungen ist die räumliche Nähe der beteiligten Chromosomenintervalle im

4.2. Somatische Chromosomenveränderungen und chromosomale Interaktionen

Zellkern, wie z. B. für die Entstehung von Translokationen in anderen Studien bereits gezeigt werden konnte. Ein Abgleich der Bruchpunktregionen mit der 4DGenome-Datenbank und Hi-C Interaktionsdaten zur lymphoblastoiden Zelllinie GM12878 ergab allerdings keine Hinweise auf hohe Interaktionswahrscheinlichkeiten der Bruchpunktregionen vor dem chromosomalen Rearrangement. Möglicherweise spielt hier aber auch die fehlende spezifische 3D-Organisation der Sézary-Vorläuferzellen eine Rolle, die natürlich in den hier genutzten Zelllinien nicht gegeben ist.

Das Fehlen einheitlicher Genomeigenschaften kann durchaus ein Hinweis einer falschen Definition der Rearrangementbruchpunkte sein oder erfolgte aufgrund der exemplarischen Auswahl der Genom- und Chromatineigenschaften. Möglicherweise weisen die Bruchpunktregionen aber auch auf eine allgemeine Heterogenität bezüglich der untersuchten Eigenschaften hin, die z. B. aufgrund eines komplexen Zusammenspiels verschiedener Genomeigenschaften destabilisiert wurden. Eine wichtige Rolle könnte dabei die transkriptionsbedingte Instabilität spielen, da nur 8 der 34 untersuchten Bruchpunkte (die über die Readinformation identifiziert wurden) keine Überschneidungen mit Genloci aufweisen.

Verlust des Tumorsuppressorgens *E2A* in Sézary-Tumorzellen

Sézary-Zellen sind bekannt für einen hohen Grad an genomischer Instabilität [423] [50]. Zu den rekurrenten Veränderungen in diesen Tumoren zählen Verluste ganzer Chromosomenarme bzw. großer Chromosomensegmente von 10q, 17p und Gewinne von 8q [459] [250] [43]. Die häufigste chromosomale Veränderung betrifft aber die Region 19p13.3, deren Verlust in 14 von 20 Patientenproben detektiert werden konnte [389]. Die 19p13.3-Deletion findet sich auch in der Seax-Zelllinie, wo sie in Kombination mit einer 19p13.3-Fusion an die 17q25.3-Region vorkommt. Die Region der kleinsten gemeinsamen Deletion von ca. 1,4 Mb (chr19: 1.368.087-2.824.434; hg18) umspannt dabei das Gen *E2A* (auch *transcription factor 3* - **TCF3** genannt), dessen Genprodukt einer Familie von Transkriptionsaktivatoren und -repressoren angehört (zusammengefasst in [107]). Der heterozygote Verlust von *E2A* konnte mittels FISH-Experimenten bestätigt werden und auch eine verringerte RNA- und Proteinexpression in *E2A*-defizienten Tumorzellproben nachgewiesen werden [389]. Weiterführende zellbiologische und molekularbiologische Untersuchungen einschließlich der ektopischen Expression des Gens in *E2A*-defizienten Zellen belegten eindeutig die Relevanz von *E2A* für die Entstehung des Sézary Syndroms. Vor dem Hintergrund der Häufigkeit von *E2A*-Verlusten und der Bedeutung von *E2A* für die Tumorentwicklung

[20] wurde nun untersucht, welche Gene durch diesen Gendosisverlust betroffen sind und ob diese Gene im Sinne einer gemeinsamen *Transcription Factory* räumliche Nähe im Kern zeigen. In anderen Zelllinien wurde bereits ein Zusammenhang zwischen Transkriptionsfaktoren und der zelltypspezifischen Dynamik der 3D-Kernarchitektur beschrieben [232], wobei eine Assoziation von intra- und interchromosomalen Interaktionen mit zelltypspezifischen Transkriptionsfaktorbindestellen nachgewiesen werden konnte. In dieser Untersuchung von Lin et al. zeigte auch E2A eine Beteiligung an der 3D-Organisation der Chromosomen der B-Zellen [232]. Die Untersuchung der durch E2A deregulierten Gene in Seax ergab allerdings keine erhöhte Interaktionswahrscheinlichkeit der E2A-Zielgene und damit keinen Anhaltspunkt, dass diese Zielgene in einer gemeinsamen *Transcription Factory* koreguliert werden.

Eine Analyse inwieweit sich der Verlust von *E2A* in T-Zellen auf die 3D-Organisation der Chromosomen auswirken könnte, ist gerade im Zusammenhang mit tumorspezifischen Rearrangementmustern interessant. Chromoplexy als möglicher transkriptionsbasierter Mechanismus der komplexen Aberrationsmuster in Sézary-Zellen könnte ein Hinweis auf die Rolle zelltypspezifischer Transkriptionsfaktoren sein. Weiterführende vergleichende Hi-C-Studien in Kombination mit einer gezielten Deletion des *E2A*-Gens in T-Zellen könnten hier einen Einblick geben.

5. Literatur

- [1] H. J. Abel und E. J. Duncavage. »Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches.« In: *Cancer genetics* 206.12 (2013), S. 432–40.
- [2] A. Abyzov u. a. »Analysis of deletion breakpoints from 1,092 humans reveals details of mutation mechanisms.« In: *Nature communications* 6 (2015), S. 7256.
- [3] A. Aguilera und T. García-Muse. »Causes of genome instability.« In: *Annual review of genetics* 47.1 (2013), S. 1–32.
- [4] F Albano u. a. »Genomic segmental duplications on the basis of the t(9;22) rearrangement in chronic myeloid leukemia.« In: *Oncogene* 29.17 (2010), S. 2509–16.
- [5] O. An u. a. »NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings.« In: *Nucleic acids research* 44.D1 (2016), S. D992–9.
- [6] A. Antonell u. a. »Evolutionary mechanisms shaping the genomic structure of the Williams-Beuren syndrome chromosomal region at human 7q11.23.« In: *Genome research* 15.9 (2005), S. 1179–88.
- [7] M. S. Apte und V. H. Meller. »Homologue pairing in flies and mammals: gene regulation when two are involved.« In: *Genetics research international* 2012 (2012), S. 430587.
- [8] J. L. Argueso u. a. »Double-strand breaks associated with repetitive DNA can reshape the genome.« In: *Proceedings of the National Academy of Sciences of the United States of America* 105.33 (2008), S. 11845–50.
- [9] L. Armengol u. a. »Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements.« In: *Human molecular genetics* 12.17 (2003), S. 2201–8.
- [10] L. Armengol u. a. »Murine segmental duplications are hot spots for chromosome and gene evolution.« In: *Genomics* 86.6 (2005), S. 692–700.
- [11] A Attarbaschi u. a. »Incidence and relevance of secondary chromosome abnormalities in childhood TEL/AML1+ acute lymphoblastic leukemia: an interphase FISH analysis.« In: *Leukemia* 18.10 (2004), S. 1611–6.
- [12] P. I. Axelrod, B Lorber und E. C. Vonderheid. »Infections complicating mycosis fungoides and Sézary syndrome.« In: *JAMA* 267.10 (1992), S. 1354–8.

- [13] S. C. Baca u. a. »Punctuated evolution of prostate cancer genomes.« In: *Cell* 153.3 (2013), S. 666–77.
- [14] U. Bacher u. a. »The incidence of submicroscopic deletions in reciprocal translocations is similar in acute myeloid leukemia, BCR-ABL positive acute lymphoblastic leukemia, and chronic myeloid leukemia.« In: *Haematologica* 90.4 (2005), S. 558–9.
- [15] J. A. Bailey u. a. »Segmental duplications: organization and impact within the current human genome project assembly.« In: *Genome research* 11.6 (2001), S. 1005–17.
- [16] J. a. Bailey und E. E. Eichler. »Primate segmental duplications: crucibles of evolution, diversity and disease.« In: *Nature reviews. Genetics* 7.7 (2006), S. 552–64.
- [17] J. a. Bailey, G. Liu und E. E. Eichler. »An Alu transposition model for the origin and expansion of human segmental duplications.« In: *American journal of human genetics* 73.4 (2003), S. 823–34.
- [18] J. a. Bailey u. a. »Hotspots of mammalian chromosomal evolution.« In: *Genome biology* 5.4 (2004), R23.
- [19] J. a. Bailey u. a. »Recent segmental duplications in the human genome.« In: *Science (New York, N.Y.)* 297.5583 (2002), S. 1003–7.
- [20] G. Bain u. a. »E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas.« In: *Molecular and cellular biology* 17.8 (1997), S. 4782–91.
- [21] G. S. Baldwin u. a. »DNA double helices recognize mutual sequence homology in a protein free environment.« In: *The journal of physical chemistry. B* 112.4 (2008), S. 1060–4.
- [22] E. Ballabio u. a. »MicroRNA expression in Sezary syndrome: identification, function, and diagnostic potential.« In: *Blood* 116.7 (2010), S. 1105–13.
- [23] T. D. Barber u. a. »Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers.« In: *Proceedings of the National Academy of Sciences of the United States of America* 105.9 (2008), S. 3443–8.
- [24] M. Barbieri u. a. »A model of the large-scale organization of chromatin.« In: *Biochemical Society transactions* 41.2 (2013), S. 508–12.
- [25] A. R. Barutcu u. a. »Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells«. In: *Genome Biology* 16.1 (2015), S. 214.
- [26] A. Barzel und M. Kupiec. »Finding a match: how do homologous sequences get together for recombination?« In: *Nature reviews. Genetics* 9.1 (2008), S. 27–37.
- [27] D. A. S. Batista u. a. »Multicolor fluorescence in situ hybridization (SKY) in mycosis fungoides and Sézary syndrome: search for recurrent chromosome abnormalities.« In: *Genes, chromosomes & cancer* 45.4 (2006), S. 383–91.
- [28] M. A. Batzer und P. L. Deininger. »Alu Repeats and Human Genomic Diversity«. In: *Nature Reviews Genetics* 3.5 (2002), S. 370–379.

- [29] M. Bayés u. a. »Mutational mechanisms of Williams-Beuren syndrome deletions.« In: *American journal of human genetics* 73.1 (2003), S. 131–51.
- [30] J.-M. Belton u. a. »Hi-C: a comprehensive technique to capture the conformation of genomes.« In: *Methods (San Diego, Calif.)* 58.3 (2012), S. 268–76.
- [31] N. L. van Berkum u. a. »Hi-C: a method to study the three-dimensional architecture of genomes.« In: *Journal of visualized experiments : JoVE* 39 (2010), S. 1–7.
- [32] C. Berthelot u. a. »The 3D organization of chromatin explains evolutionary fragile genomic regions.« In: *Cell reports* 10.11 (2015), S. 1913–24.
- [33] W. a. Bickmore und B. van Steensel. »Genome architecture: domain organization of interphase chromosomes.« In: *Cell* 152.6 (2013), S. 1270–84.
- [34] J. L. Biedler, L. Helson und B. A. Spengler. »Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture.« In: *Cancer research* 33.11 (1973), S. 2643–52.
- [35] M. L. Bochman, K. Paeschke und V. a. Zakian. »DNA secondary structures: stability and function of G-quadruplex structures.« In: *Nature reviews. Genetics* 13.11 (2012), S. 770–80. arXiv: [NIHMS150003](https://arxiv.org/abs/150003).
- [36] M. Boi u. a. »PRDM1/BLIMP1: a tumor suppressor gene in B and T cell lymphomas.« In: *Leukemia & lymphoma* 56.5 (2015), S. 1223–8.
- [37] S. Boissinot u. a. »Fitness cost of LINE-1 (L1) activity in humans.« In: *Proceedings of the National Academy of Sciences of the United States of America* 103.25 (2006), S. 9590–4.
- [38] A. Bolzer u. a. »Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes.« In: *PLoS biology* 3.5 (2005), e157.
- [39] B. A. M. Bouwman und W. de Laat. »Getting the genome in shape: the formation of loops, domains and compartments.« In: *Genome biology* 16.1 (2015), S. 154.
- [40] S. Boyle u. a. »The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells.« In: *Human molecular genetics* 10.3 (2001), S. 211–9.
- [41] M. R. Branco und A. Pombo. »Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations.« In: *PLoS biology* 4.5 (2006), e138.
- [42] J. M. Bridger u. a. »Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts.« In: *Current biology : CB* 10.3 (2000), S. 149–52.
- [43] V. Brito-Babapulle u. a. »p53 allele deletion and protein accumulation occurs in the absence of p53 gene mutation in T-prolymphocytic leukaemia and Sezary syndrome.« In: *British journal of haematology* 110.1 (2000), S. 180–7.
- [44] R. N. Buick, R. Pullano und J. M. Trent. »Comparative properties of five human ovarian adenocarcinoma cell lines.« In: *Cancer research* 45.8 (1985), S. 3668–76.
- [45] S. F. Bunting und A. Nussenzweig. »End-joining, translocations and cancer.« In: *Nature reviews. Cancer* 13.7 (2013), S. 443–54.

- [46] B. Burman u. a. »Histone modifications predispose genome regions to breakage and translocation.« In: *Genes & development* 29.13 (2015), S. 1393–402.
- [47] J. N. Burton u. a. »Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions.« In: *Nature biotechnology* 31.12 (2013), S. 1119–25.
- [48] J. N. Burton u. a. »Species-level deconvolution of metagenome assemblies with Hi-C-based contact probability maps.« In: *G3 (Bethesda, Md.)* 4.7 (2014), S. 1339–46.
- [49] O. Capozzi u. a. »A comprehensive molecular cytogenetic analysis of chromosome rearrangements in gibbons.« In: *Genome research* 22.12 (2012), S. 2520–8.
- [50] E. Caprini u. a. »Identification of key regions and genes important in the pathogenesis of Sezary syndrome by combining genomic and expression microarrays.« In: *Cancer research* 69.21 (2009), S. 8438–46.
- [51] L. Carbone u. a. »Evolutionary breakpoints in the gibbon suggest association between cytosine methylation and karyotype evolution.« In: *PLoS genetics* 5.6 (2009), e1000538.
- [52] H. Caron u. a. »The human transcriptome map: clustering of highly expressed genes in chromosomal domains.« In: *Science (New York, N.Y.)* 291.5507 (2001), S. 1289–92.
- [53] G. Cavalli und T. Misteli. »Functional implications of genome topology.« In: *Nature structural & molecular biology* 20.3 (2013), S. 290–9.
- [54] R. Z. Cer u. a. »Non-B DNA v2.0: a database of predicted non-B DNA-forming motifs and its associated tools.« In: *Nucleic acids research* 41.Database issue (2013), S. D94–D100.
- [55] E. V. Chambers, W. a. Bickmore und C. a. Semple. »Divergence of mammalian higher order chromatin structure is associated with developmental loci.« In: *PLoS computational biology* 9.4 (2013), e1003017.
- [56] S. Chambeyron und W. a. Bickmore. »Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription.« In: *Genes & development* 18.10 (2004), S. 1119–30.
- [57] T. Chandra u. a. »Global reorganization of the nuclear landscape in senescent cells.« In: *Cell reports* 10.4 (2015), S. 471–83.
- [58] C.-H. Chang, T. Gourley und D. Patel. »CIITA expression and Th1-Th2 balance.« In: *Journal of immunology (Baltimore, Md. : 1950)* 170.9 (2003), 4425; author reply 4425.
- [59] J. T. Chang und J. R. Nevins. »GATHER: a systems approach to interpreting genomic signatures.« In: *Bioinformatics (Oxford, England)* 22.23 (2006), S. 2926–33.
- [60] S.-L. Chang u. a. »Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations.« In: *PLoS genetics* 9.1 (2013), e1003232.
- [61] E. Y. Chen u. a. »Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.« In: *BMC bioinformatics* 14 (2013), S. 128.
- [62] H. Chen und P. C. Boutros. »VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R.« In: *BMC bioinformatics* 12.1 (2011), S. 35.

- [63] J.-M. Chen u. a. »Gene conversion: mechanisms, evolution and human disease.« In: *Nature reviews. Genetics* 8.10 (2007), S. 762–75.
- [64] J.-M. Chen u. a. »Genomic rearrangements in inherited disease and cancer.« In: *Seminars in cancer biology* 20.4 (2010), S. 222–33.
- [65] J.-M. Chen u. a. »Genomic rearrangements in inherited disease and cancer.« In: *Seminars in cancer biology* 20.4 (2010), S. 222–33.
- [66] W. Chen u. a. »Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing.« In: *European journal of human genetics : EJHG* 18.5 (2010), S. 539–43.
- [67] W. Chen u. a. »CGHPRO – a comprehensive data analysis tool for array CGH.« In: *BMC bioinformatics* 6 (2005), S. 85.
- [68] W. Chen u. a. »Mapping translocation breakpoints by next-generation sequencing.« In: *Genome research* 18.7 (2008), S. 1143–9.
- [69] B. Chenais. »Transposable elements in cancer and other human diseases.« In: *Current cancer drug targets* 15.3 (2015), S. 227–42.
- [70] Z. Cheng u. a. »A genome-wide comparison of recent chimpanzee and human segmental duplications.« In: *Nature* 437.7055 (2005), S. 88–93.
- [71] J. Cheung u. a. »Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence.« In: *Genome biology* 4.4 (2003), R25.
- [72] R. Chiarle u. a. »Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells.« In: *Cell* 147.1 (2011), S. 107–19.
- [73] F. Chiaromonte, V. B. Yap und W. Miller. »Scoring pairwise genomic sequence alignments.« In: *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* 126 (2002), S. 115–26.
- [74] J. Choi u. a. »Genomic landscape of cutaneous T cell lymphoma.« In: *Nature genetics* 47.9 (2015), S. 1011–9.
- [75] F. Ciabrelli und G. Cavalli. »Chromatin-driven behavior of topologically associating domains.« In: *Journal of molecular biology* 427.3 (2015), S. 608–25.
- [76] R. Ciccone u. a. »Reciprocal translocations: a trap for cytogenetists?« In: *Human genetics* 117.6 (2005), S. 571–82.
- [77] D. F. Conrad und M. E. Hurles. »The population genetics of structural variation.« In: *Nature genetics* 39.7 Suppl (2007), S30–6.
- [78] P. R. Cook. »A model for all genomes: the role of transcription factories.« In: *Journal of molecular biology* 395.1 (2010), S. 1–10.
- [79] D. N. Cooper u. a. »On the sequence-directed nature of human gene mutation: the role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease.« In: *Human mutation* 32.10 (2011), S. 1075–99.
- [80] A. Cournac, R. Koszul und J. Mozziconacci. »The 3D folding of metazoan genomes correlates with the association of similar repetitive elements.« In: *Nucleic acids research* 44.1 (2016), S. 245–55.

- [81] T Cremer und C Cremer. »Chromosome territories, nuclear architecture and gene regulation in mammalian cells.« In: *Nature reviews. Genetics* 2.4 (2001), S. 292–301.
- [82] C. Cristofolletti u. a. »Comprehensive analysis of PTEN status in Sezary syndrome.« In: *Blood* 122.20 (2013), S. 3511–20.
- [83] J. a. Croft u. a. »Differences in the localization and morphology of chromosomes in the human nucleus.« In: *The Journal of cell biology* 145.6 (1999), S. 1119–31.
- [84] Z. Dai, Y. Xiong und X. Dai. »Neighboring genes show interchromosomal colocalization after their separation.« In: *Molecular biology and evolution* 31.5 (2014), S. 1166–72.
- [85] J. A. Daniel und A. Nussenzweig. »Roles for histone H3K4 methyltransferase activities during immunoglobulin class-switch recombination.« In: *Biochimica et biophysica acta* 1819.7 (2012), S. 733–8. arXiv: [NIHMS150003](https://arxiv.org/abs/NIHMS150003).
- [86] C Danilowicz u. a. »Single molecule detection of direct, homologous, DNA/DNA pairing.« In: *Proceedings of the National Academy of Sciences of the United States of America* 106.47 (2009), S. 19824–9.
- [87] E. Darai-Ramqvist u. a. »Segmental duplications and evolutionary plasticity at tumor chromosome break-prone regions.« In: *Genome research* 18.3 (2008), S. 370–9.
- [88] J. Dekker und E. Heard. »Structural and functional diversity of Topologically Associating Domains.« In: *FEBS letters* 589.20 Pt A (2015), S. 2877–84.
- [89] J. Dekker, M. a. Marti-Renom und L. a. Mirny. »Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data.« In: *Nature reviews. Genetics* 14.6 (2013), S. 390–403.
- [90] J. Dekker u. a. »Capturing chromosome conformation.« In: *Science (New York, N.Y.)* 295.5558 (2002), S. 1306–11.
- [91] J. R. Dixon u. a. »Chromatin architecture reorganization during stem cell differentiation.« In: *Nature* 518.7539 (2015), S. 331–6.
- [92] J. R. Dixon u. a. »Topological domains in mammalian genomes identified by analysis of chromatin interactions.« In: *Nature* 485.7398 (2012), S. 376–80.
- [93] O. M. Dovey u. a. »Histone deacetylase 1 and 2 are essential for normal T-cell development and genomic stability in mice.« In: *Blood* 121.8 (2013), S. 1335–44.
- [94] H. G. Drexler, G Gaedicke und J Minowada. »Isoenzyme studies in human leukemia-lymphoma cell lines–1. Carboxylic esterase.« In: *Leukemia research* 9.2 (1985), S. 209–29.
- [95] H. G. Drexler und J Minowada. »Morphological, immunophenotypical and isoenzymatic profiles of human leukemia cells and derived T-cell lines.« In: *Hematological oncology* 7.2 (), S. 115–25.
- [96] Z. Du, Y. Zhao und N. Li. »Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription.« In: *Genome research* 18.2 (2008), S. 233–41.
- [97] B. L. Dumont und E. E. Eichler. »Signals of historical interlocus gene conversion in human segmental duplications.« In: *PloS one* 8.10 (2013), e75949.

- [98] F Dutly und A Schinzel. »Unequal interchromosomal rearrangements may result in elastin gene deletions causing the Williams-Beuren syndrome.« In: *Human molecular genetics* 5.12 (1996), S. 1893–8.
- [99] G. Ebert u. a. »Distribution of segmental duplications in the context of higher order chromatin organisation of human chromosome 7.« In: *BMC genomics* 15 (2014), S. 537.
- [100] E. Eden u. a. »Discovering motifs in ranked lists of DNA sequences.« In: *PLoS computational biology* 3.3 (2007), e39.
- [101] R. Edgar, M. Domrachev und A. E. Lash. »Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.« In: *Nucleic acids research* 30.1 (2002), S. 207–10.
- [102] E. E. Eichler, N. Archidiacono und M. Rocchi. »CAGGG repeats and the pericentromeric duplication of the hominoid genome.« In: *Genome research* 9.11 (1999), S. 1048–58.
- [103] J. E. Ellerman u. a. »Masquerader: high mobility group box-1 and cancer.« In: *Clinical cancer research : an official journal of the American Association for Cancer Research* 13.10 (2007), S. 2836–48.
- [104] W. Ellmeier. »Molecular control of CD4(+) T cell lineage plasticity and integrity.« In: *International immunopharmacology* 28.2 (2015), S. 813–7.
- [105] M. Eltsov u. a. »Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ.« In: *Proceedings of the National Academy of Sciences of the United States of America* 105.50 (2008), S. 19732–7.
- [106] E. Envervald u. a. »A regulatory role for the cohesin loader NIPBL in nonhomologous end joining during immunoglobulin class switch recombination.« In: *The Journal of experimental medicine* 210.12 (2013), S. 2503–13.
- [107] I Engel und C Murre. »The function of E- and Id proteins in lymphocyte development.« In: *Nature reviews. Immunology* 1.3 (2001), S. 193–9.
- [108] J. M. Engreitz, V. Agarwala und L. a. Mirny. »Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease.« In: *PloS one* 7.9 (2012), e44196.
- [109] F Erdogan u. a. »Impact of low copy repeats on the generation of balanced and unbalanced chromosomal aberrations in mental retardation.« In: *Cytogenetic and genome research* 115.3-4 (2006), S. 247–53.
- [110] J. Ernst und M. Kellis. »Discovery and characterization of chromatin states for systematic annotation of the human genome.« In: *Nature biotechnology* 28.8 (2010), S. 817–25.
- [111] J. Ernst u. a. »Mapping and analysis of chromatin state dynamics in nine human cell types.« In: *Nature* 473.7345 (2011), S. 43–9.
- [112] G. Escaramís, E. Docampo und R. Rabionet. »A decade of structural variants: description, history and methods to detect structural variation.« In: *Briefings in functional genomics* 14.5 (2015), S. 305–14.

- [113] a. K. Ewart u. a. »Hemizygoty at the elastin locus in a developmental disorder, Williams syndrome.« In: *Nature genetics* 5.1 (1993), S. 11–6.
- [114] M. Farré, T. J. Robinson und A. Ruiz-Herrera. »An Integrative Breakage Model of genome architecture, reshuffling and evolution: The Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity.« In: *BioEssays : news and reviews in molecular, cellular and developmental biology* 37.5 (2015), S. 479–88.
- [115] J. a. Fawcett und H. Innan. »The role of gene conversion in preserving rearrangement hotspots in the human genome.« In: *Trends in genetics : TIG* 29.10 (2013), S. 561–8.
- [116] L. Feuk, A. R. Carson und S. W. Scherer. »Structural variation in the human genome.« In: *Nature reviews. Genetics* 7.2 (2006), S. 85–97.
- [117] H. Fiegler u. a. »DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones.« In: *Genes, chromosomes & cancer* 36.4 (2003), S. 361–74.
- [118] J. T. Finch und A. Klug. »Solenoidal model for superstructure in chromatin.« In: *Proceedings of the National Academy of Sciences of the United States of America* 73.6 (1976), S. 1897–901.
- [119] H. V. Firth u. a. »DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources.« In: *American journal of human genetics* 84.4 (2009), S. 524–33.
- [120] T. C. Fischer u. a. »Genomic aberrations and survival in cutaneous T cell lymphomas.« In: *The Journal of investigative dermatology* 122.3 (2004), S. 579–86.
- [121] L. van der Fits u. a. »MicroRNA-21 expression in CD4+ T cells is regulated by STAT3 and is pathologically involved in Sézary syndrome.« In: *The Journal of investigative dermatology* 131.3 (2011), S. 762–8.
- [122] L. van der Fits u. a. »NOTCH1 signaling as a therapeutic target in Sézary syndrome.« In: *The Journal of investigative dermatology* 132.12 (2012), S. 2810–7.
- [123] P. Flicek u. a. »Ensembl 2012.« In: *Nucleic acids research* 40.Database issue (2012), S. D84–90.
- [124] A. Franceschini u. a. »STRING v9.1: protein-protein interaction networks, with increased coverage and integration.« In: *Nucleic acids research* 41.Database issue (2013), S. D808–15.
- [125] J. Fraser u. a. »An Overview of Genome Organization and How We Got There: from FISH to Hi-C.« In: *Microbiology and molecular biology reviews : MMBR* 79.3 (2015), S. 347–72.
- [126] S. Fredholm u. a. »STAT3 activation and infiltration of eosinophil granulocytes in mycosis fungoides.« In: *Anticancer research* 34.10 (2014), S. 5277–86.
- [127] A. J. Fritz u. a. »Wide-scale alterations in interchromosomal organization in breast cancer cells: defining a network of interacting chromosomes.« In: *Human molecular genetics* 23.19 (2014), S. 5133–46.
- [128] G. Fudenberg u. a. »High order chromatin architecture shapes the landscape of chromosomal alterations in cancer.« In: *Nature biotechnology* 29.12 (2011), S. 1109–13.
- [129] G. Fudenberg u. a. »High order chromatin architecture shapes the landscape of chromosomal alterations in cancer.« In: *Nature biotechnology* 29.12 (2011), S. 1109–13.

- [130] M. J. Fullwood u. a. »Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses.« In: *Genome research* 19.4 (2009), S. 521–32.
- [131] A. Fungtammasan u. a. »A genome-wide analysis of common fragile sites: what features determine chromosomal instability in the human genome?« In: *Genome research* 22.6 (2012), S. 993–1005.
- [132] T. Gaj, C. A. Gersbach und C. F. Barbas. »ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.« In: *Trends in biotechnology* 31.7 (2013), S. 397–405. arXiv: [NIHMS150003](#).
- [133] D. W. Garsed u. a. »The architecture and evolution of cancer neochromosomes.« In: *Cancer cell* 26.5 (2014), S. 653–67.
- [134] A. F. Gazdar u. a. »Mitogen requirements for the in vitro propagation of cutaneous T-cell lymphomas.« In: *Blood* 55.3 (1980), S. 409–17.
- [135] N. Gheldof u. a. »Structural variation-associated expression changes are paralleled by chromatin architecture modifications.« In: *PloS one* 8.11 (2013), e79973.
- [136] G. Giannuzzi u. a. »Hominoid fission of chromosome 14/15 and the role of segmental duplications.« In: *Genome research* 23.11 (2013), S. 1763–73.
- [137] S. Giavara u. a. »Yeast Nhp6A/B and mammalian Hmgbl facilitate the maintenance of genome stability.« In: *Current biology : CB* 15.1 (2005), S. 68–72.
- [138] S. Goetze u. a. »The three-dimensional structure of human interphase chromosomes is related to the transcriptome map.« In: *Molecular and cellular biology* 27.12 (2007), S. 4475–87.
- [139] A. Y. Grosberg, S. K. Nechaev und E. I. Shakhnovich. »The role of topological constraints in the kinetics of collapse of macromolecules.« In: *Journal de Physique* 49.12 (1988), S. 2095–2100.
- [140] A. Grosberg u. a. »Crumpled Globule Model of the Three-Dimensional Structure of DNA .« In: *Europhysics Letters* 23.5 (1993), S. 373–378.
- [141] C. Grunau u. a. »Mapping of the juxtacentromeric heterochromatin-euchromatin frontier of human chromosome 21.« In: *Genome research* 16.10 (2006), S. 1198–207.
- [142] W. Gu, F. Zhang und J. R. Lupski. »Mechanisms for human genomic rearrangements.« In: *PathoGenetics* 1.1 (2008), S. 4.
- [143] L. Guelen u. a. »Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions.« In: *Nature* 453.7197 (2008), S. 948–51.
- [144] A. A. Guerrero, C. Martínez-A und K. H. van Wely. »Merotelic attachments and non-homologous end joining are the basis of chromosomal instability.« In: *Cell division* 5 (2010), S. 13.
- [145] J. Guy u. a. »Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10q.« In: *Human molecular genetics* 9.13 (2000), S. 2029–42.
- [146] T. Haaf, K. Steinlein und M. Schmid. »Preferential somatic pairing between homologous heterochromatic regions of human chromosomes.« In: *American journal of human genetics* 38.3 (1986), S. 319–29.
- [147] M. Hackenberg u. a. »The biased distribution of Alus in human isochores might be driven by recombination.« In: *Journal of molecular evolution* 60.3 (2005), S. 365–77.

- [148] O. Hakim u. a. »Spatial congregation of STAT binding directs selective nuclear architecture during T-cell functional differentiation.« In: *Genome research* 23.3 (2013), S. 462–72.
- [149] P. Hallast u. a. »Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster.« In: *Genome research* 15.11 (2005), S. 1535–46.
- [150] C. Hallermann u. a. »Chromosomal aberration patterns differ in subtypes of primary cutaneous B cell lymphomas.« In: *The Journal of investigative dermatology* 122.6 (2004), S. 1495–502.
- [151] D. Hanahan und R. a. Weinberg. »Hallmarks of cancer: the next generation.« In: *Cell* 144.5 (2011), S. 646–74.
- [152] R. S. Hansen u. a. »Sequencing newly replicated DNA reveals widespread plasticity in human replication timing.« In: *Proceedings of the National Academy of Sciences of the United States of America* 107.1 (2010), S. 139–44.
- [153] L. Harewood und P. Fraser. »The impact of chromosomal rearrangements on regulation of gene expression.« In: *Human molecular genetics* 23.R1 (2014), R76–82.
- [154] M. A. Harris u. a. »The Gene Ontology (GO) database and informatics resource.« In: *Nucleic acids research* 32.Database issue (2004), S. D258–61.
- [155] D. A. Hartasánchez u. a. »Interplay of interlocus gene conversion and crossover in segmental duplications under a neutral scenario.« In: *G3 (Bethesda, Md.)* 4.8 (2014), S. 1479–89.
- [156] S. Heinz u. a. »Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.« In: *Molecular cell* 38.4 (2010), S. 576–89.
- [157] L. W. Hillier u. a. »The DNA sequence of human chromosome 7.« In: *Nature* 424.6945 (2003), S. 157–64.
- [158] E. L. Hockenhull u. a. »A complete physical contig and partial transcript map of the Williams syndrome critical region.« In: *Genomics* 58.2 (1999), S. 138–45.
- [159] J. E. Horvath u. a. »Lessons from the human genome: transitions between euchromatin and heterochromatin.« In: *Human molecular genetics* 10.20 (2001), S. 2215–23.
- [160] J. E. Horvath u. a. »Punctuated duplication seeding events during the evolution of human chromosome 2p11.« In: *Genome research* 15.7 (2005), S. 914–27.
- [161] C. Hou u. a. »Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains.« In: *Molecular cell* 48.3 (2012), S. 471–84.
- [162] K. D. Howarth u. a. »Large duplications at reciprocal translocation breakpoints that might be the counterpart of large deletions and could arise from stalled replication bubbles.« In: *Genome research* 21.4 (2011), S. 525–34.
- [163] M. R. Hübner, M. a. Eckersley-Maslin und D. L. Spector. »Chromatin organization and transcriptional regulation.« In: *Current opinion in genetics & development* 23.2 (2013), S. 89–95.
- [164] S. P. Hunger. »Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis.« In: *Blood* 87.4 (1996), S. 1211–24.

- [165] B. J. Huntly u. a. »Deletions of the derivative chromosome 9 occur at the time of the Philadelphia translocation and provide a powerful and independent prognostic indicator in chronic myeloid leukemia.« In: *Blood* 98.6 (2001), S. 1732–8.
- [166] M. Hurles und J. R. Lupski. »Recombination Hotspots in Nonallelic Homologous Recombination«. In: *Genomic Disorders*. Hrsg. von J. R. Lupski und P. Stankiewicz. Totowa, New Jersey: Humana Press, 2006, S. 341–355.
- [167] M. E. Hurles, E. T. Dermitzakis und C. Tyler-Smith. »The functional impact of structural variation in humans.« In: *Trends in genetics : TIG* 24.5 (2008), S. 238–45.
- [168] M. E. Hurles u. a. »Origins of chromosomal rearrangement hotspots in the human genome: evidence from the AZFa deletion hotspots.« In: *Genome biology* 5.8 (2004), R55.
- [169] S. T. Hwang u. a. »Mycosis fungoides and Sézary syndrome.« In: *Lancet (London, England)* 371.9616 (2008), S. 945–57.
- [170] a. J. Iafrate u. a. »Detection of large-scale variation in the human genome.« In: *Nature genetics* 36.9 (2004), S. 949–51.
- [171] J. Ibn-Salem u. a. »Deletions of chromosomal regulatory boundaries are associated with congenital disease.« In: *Genome biology* 15.9 (2014), S. 423.
- [172] M. Imakaev u. a. »Iterative correction of Hi-C data reveals hallmarks of chromosome organization.« In: *Nature methods* 9.10 (2012), S. 999–1003.
- [173] M. H. Imam u. a. »Incidence and survival patterns of cutaneous T-cell lymphomas in the United States.« In: *Leukemia & lymphoma* 54.4 (2013), S. 752–9.
- [174] S. Inoue u. a. »Self-assembly of double-stranded DNA molecules at nanomolar concentrations.« In: *Biochemistry* 46.1 (2007), S. 164–71.
- [175] International Human Genome Sequencing Consortium. »Finishing the euchromatic sequence of the human genome.« In: *Nature* 431.7011 (2004), S. 931–45.
- [176] D. Ioannou u. a. »Chromosome territory repositioning induced by PHA-activation of lymphocytes: A 2D and 3D appraisal.« In: *Molecular cytogenetics* 8.1 (2015), S. 47.
- [177] A. Itsara u. a. »Population analysis of large copy number variants and hotspots of human genetic disease.« In: *American journal of human genetics* 84.2 (2009), S. 148–61.
- [178] K. Izykowska u. a. »Identification of multiple complex rearrangements associated with deletions in the 6q23-27 region in Sézary syndrome.« In: *The Journal of investigative dermatology* 133.11 (2013), S. 2617–25.
- [179] R. Jäger u. a. »Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci.« In: *Nature communications* 6 (2015), S. 6178.
- [180] Y. Ji u. a. »The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci.« In: *Cell* 141.3 (2010), S. 419–31.
- [181] Z. Jiang u. a. »Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution.« In: *Nature genetics* 39.11 (2007), S. 1361–8.

- [182] J. Jurka. »Evolutionary impact of human Alu repetitive elements.« In: *Current opinion in genetics & development* 14.6 (2004), S. 603–8.
- [183] K Kaltoft u. a. »A continuous T-cell line from a patient with Sézary syndrome.« In: *Archives of dermatological research* 279.5 (1987), S. 293–8.
- [184] K Kaltoft u. a. »Establishment of two continuous T-cell strains from a single plaque of a patient with mycosis fungoides.« In: *In vitro cellular & developmental biology : journal of the Tissue Culture Association* 28A.3 Pt 1 (1992), S. 161–7.
- [185] T. V. Kalvik und T Arnesen. »Protein N-terminal acetyltransferases in cancer.« In: *Oncogene* 32.3 (2013), S. 269–76.
- [186] P. Kampstra. »Beanplot: A Boxplot Alternative for Visual Comparison of Distributions«. In: *Journal of Statistical Software, Code Snippets* 28.1 (2008), S. 1–9.
- [187] M Kanehisa und S Goto. »KEGG: kyoto encyclopedia of genes and genomes.« In: *Nucleic acids research* 28.1 (2000), S. 27–30.
- [188] M. Kanehisa u. a. »KEGG as a reference resource for gene and protein annotation.« In: *Nucleic acids research* 44.D1 (2016), S. D457–62.
- [189] S. Kang u. a. »SCTR regulates cell cycle-related genes toward anti-proliferation in normal breast cells while having pro-proliferation activity in breast cancer cells.« In: *International journal of oncology* 47.5 (2015), S. 1923–31.
- [190] N. Kaplan und J. Dekker. »High-throughput genome scaffolding from in vivo DNA interaction frequency.« In: *Nature biotechnology* 31.12 (2013), S. 1143–7. arXiv: [NIHMS150003](https://arxiv.org/abs/150003).
- [191] D. Karolchik u. a. »The UCSC Table Browser data retrieval tool.« In: *Nucleic acids research* 32.Database issue (2004), S. D493–6.
- [192] H. Kehrer-Sawatzki und D. N. Cooper. »Molecular mechanisms of chromosomal rearrangement during primate evolution.« In: *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 16.1 (2008), S. 41–56.
- [193] H. Kehrer-Sawatzki und D. N. Cooper. »Structural divergence between the human and chimpanzee genomes.« In: *Human genetics* 120.6 (2007), S. 759–78.
- [194] W. J. Kent u. a. »Evolution’s cauldron: duplication, deletion, and rearrangement in the mouse and human genomes.« In: *Proceedings of the National Academy of Sciences of the United States of America* 100.20 (2003), S. 11484–9.
- [195] W. J. Kent u. a. »The human genome browser at UCSC.« In: *Genome research* 12.6 (2002), S. 996–1006.
- [196] M. J. Kiel u. a. »Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sézary syndrome.« In: *Nature communications* 6 (2015), S. 8470.
- [197] P. M. Kim u. a. »Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history.« In: *Genome research* 18.12 (2008), S. 1865–74.

- [198] T.-M. Kim u. a. »Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes.« In: *Genome research* 23.2 (2013), S. 217–27.
- [199] J. Kind u. a. »Single-cell dynamics of genome-nuclear lamina interactions.« In: *Cell* 153.1 (2013), S. 178–92.
- [200] S. Kirsch u. a. »Evolutionary dynamics of segmental duplications from human Y-chromosomal euchromatin/heterochromatin transition regions.« In: *Genome research* 18.7 (2008), S. 1030–42.
- [201] I. a. Klein u. a. »Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes.« In: *Cell* 147.1 (2011), S. 95–106.
- [202] A. Klimopoulos, D. Sellis und Y. Almirantis. »Widespread occurrence of power-law distributions in inter-repeat distances shaped by genome dynamics.« In: *Gene* 499.1 (2012), S. 88–98.
- [203] W. P. Kloosterman u. a. »Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer.« In: *Genome biology* 12.10 (2011), R103.
- [204] E. Kolomietz u. a. »Primary chromosomal rearrangements of leukemia are frequently accompanied by extensive submicroscopic deletions and may lead to altered prognosis.« In: *Blood* 97.11 (2001), S. 3581–8.
- [205] E. Kolomietz u. a. »The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors.« In: *Genes, chromosomes & cancer* 35.2 (2002), S. 97–112.
- [206] R. D. Kornberg. »Chromatin structure: a repeating unit of histones and DNA.« In: *Science (New York, N.Y.)* 184.4139 (1974), S. 868–71.
- [207] A. A. Kornyshev und S. Leikin. »Sequence recognition in the pairing of DNA duplexes.« In: *Physical review letters* 86.16 (2001), S. 3666–9.
- [208] O. H. Krämer. »HDAC2: a critical factor in health and disease.« In: *Trends in pharmacological sciences* 30.12 (2009), S. 647–55.
- [209] M. Krzywinski u. a. »A set of BAC clones spanning the human genome.« In: *Nucleic acids research* 32.12 (2004), S. 3651–60.
- [210] M. Krzywinski u. a. »Circos: an information aesthetic for comparative genomics.« In: *Genome research* 19.9 (2009), S. 1639–45.
- [211] N. Kubben u. a. »Mapping of lamin A- and progerin-interacting genome regions.« In: *Chromosoma* 121.5 (2012), S. 447–64.
- [212] R. P. Kumar u. a. »Repeat performance: how do genome packaging and regulation depend on simple sequence repeats?« In: *BioEssays : news and reviews in molecular, cellular and developmental biology* 32.2 (2010), S. 165–74.
- [213] K. Küpper u. a. »Radial chromatin positioning is shaped by local gene density, not by gene expression.« In: *Chromosoma* 116.3 (2007), S. 285–306.
- [214] M. Kuroda u. a. »Alteration of chromosome positioning during adipocyte differentiation.« In: *Journal of cell science* 117.Pt 24 (2004), S. 5897–903.
- [215] B. R. Lajoie, J. Dekker und N. Kaplan. »The Hitchhiker’s guide to Hi-C analysis: practical guidelines.« In: *Methods (San Diego, Calif.)* 72 (2015), S. 65–75.

- [216] B Lamprecht u. a. »The tumour suppressor p53 is frequently nonfunctional in Sézary syndrome.« In: *The British journal of dermatology* 167.2 (2012), S. 240–6.
- [217] X. Lan u. a. »Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages.« In: *Nucleic acids research* 40.16 (2012), S. 7690–704.
- [218] S. S. Lange, D. L. Mitchell und K. M. Vasquez. »High mobility group protein B1 enhances DNA repair and chromatin modification after DNA damage.« In: *Proceedings of the National Academy of Sciences of the United States of America* 105.30 (2008), S. 10320–5.
- [219] S. S. Lange und K. M. Vasquez. »HMGB1: the jack-of-all-trades protein is a master DNA repair mechanic.« In: *Molecular carcinogenesis* 48.7 (2009), S. 571–80.
- [220] D. M. Larkin u. a. »Breakpoint regions and homologous synteny blocks in chromosomes have different evolutionary histories.« In: *Genome research* 19.5 (2009), S. 770–7.
- [221] S. M. Lauberth u. a. »H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation.« In: *Cell* 152.5 (2013), S. 1021–36.
- [222] H. Lee u. a. »Detection of structural variants involving repetitive regions in the reference genome.« In: *Journal of computational biology : a journal of computational molecular cell biology* 21.3 (2014), S. 219–33.
- [223] W. Lee u. a. »The mutation spectrum revealed by paired genome sequences from a lung cancer patient.« In: *Nature* 465.7297 (2010), S. 473–7.
- [224] R. Leinonen u. a. »The sequence read archive.« In: *Nucleic acids research* 39.Database issue (2011), S. D19–21.
- [225] C. Lemaitre u. a. »Analysis of fine-scale mammalian evolutionary breakpoints provides new insight into their relation to genome organisation.« In: *BMC genomics* 10 (2009), S. 335.
- [226] G. Li u. a. »ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing.« In: *Genome biology* 11.2 (2010), R22.
- [227] P. Lichter u. a. »Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries.« In: *Human genetics* 80.3 (1988), S. 224–34.
- [228] M. R. Lieber. »The mechanism of human nonhomologous DNA end joining.« In: *The Journal of biological chemistry* 283.1 (2008), S. 1–5.
- [229] E. Lieberman-Aiden u. a. »Comprehensive mapping of long-range interactions reveals folding principles of the human genome.« In: *Science (New York, N. Y.)* 326.5950 (2009), S. 289–93.
- [230] A. Lietz u. a. »Loss of bHLH transcription factor E2A activity in primary effusion lymphoma confers resistance to apoptosis.« In: *British journal of haematology* 137.4 (2007), S. 342–8.
- [231] W. M. Lin u. a. »Characterization of the DNA copy-number genome in the blood of cutaneous T-cell lymphoma patients.« In: *The Journal of investigative dermatology* 132.1 (2012), S. 188–97.
- [232] Y. C. Lin u. a. »A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate.« In: *Nature immunology* 11.7 (2010), S. 635–43.

- [233] E. V. Linardopoulou u. a. »Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication.« In: *Nature* 437.7055 (2005), S. 94–100.
- [234] G. Liu u. a. »Genome chaos: survival strategy during crisis.« In: *Cell cycle (Georgetown, Tex.)* 13.4 (2014), S. 528–37.
- [235] Y. Liu, R. Prasad und S. H. Wilson. »HMGB1: roles in base excision repair and related function.« In: *Biochimica et biophysica acta* 1799.1-2 (2014), S. 119–30.
- [236] D. P. Locke u. a. »Large-scale variation among human and great ape genomes determined by array comparative genomic hybridization.« In: *Genome research* 13.3 (2003), S. 347–57.
- [237] M. Long u. a. »The origin of new genes: glimpses from the young and old.« In: *Nature reviews. Genetics* 4.11 (2003), S. 865–75.
- [238] M. S. Longo u. a. »Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty.« In: *BMC genomics* 10.1 (2009), S. 334.
- [239] V. V. Lunyak und M. Atallah. »Genomic relationship between SINE retrotransposons, Pol III-Pol II transcription, and chromatin organization: the journey from junk to jewel.« In: *Biochemistry and cell biology = Biochimie et biologie cellulaire* 89.5 (2011), S. 495–504.
- [240] D. G. Lupiáñez u. a. »Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.« In: *Cell* 161.5 (2015), S. 1012–25.
- [241] J. R. Lupski. »Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits.« In: *Trends in genetics : TIG* 14.10 (1998), S. 417–22.
- [242] J. R. Lupski. »Genomic disorders ten years on.« In: *Genome medicine* 1.4 (2009), S. 42.
- [243] J. R. Lupski und P. Stankiewicz. »Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes.« In: *PLoS genetics* 1.6 (2005), e49.
- [244] M. Lutzner u. a. »Cutaneous T-cell lymphomas: the Sézary syndrome, mycosis fungoides, and related disorders.« In: *Annals of internal medicine* 83.4 (1975), S. 534–52.
- [245] M. Maechler u. a. *cluster: Cluster Analysis Basics and Extensions*. 2015.
- [246] K. Maeshima u. a. »Chromatin as dynamic 10-nm fibers.« In: *Chromosoma* 123.3 (2014), S. 225–37.
- [247] S. Magadum u. a. »Gene duplication as a major force in evolution.« In: *Journal of genetics* 92.1 (2013), S. 155–61.
- [248] N. Maizels und L. T. Gray. »The G4 genome.« In: *PLoS genetics* 9.4 (2013), e1003468.
- [249] T. Makino, A. McLysaght und M. Kawata. »Genome-wide deserts for copy number variation in vertebrates.« In: *Nature communications* 4 (2013), S. 2283.
- [250] X. Mao u. a. »Heterogeneous abnormalities of CCND1 and RB1 in primary cutaneous T-Cell lymphomas suggesting impaired cell cycle control in disease pathogenesis.« In: *The Journal of investigative dermatology* 126.6 (2006), S. 1388–95.
- [251] X.-j. Mao u. a. »Expression of HMGB1 and its clinical significance in T-cell lymphoma.« In: *Asian Pacific journal of cancer prevention : APJCP* 13.11 (2012), S. 5569–71.

- [252] N. V. Marella u. a. »Chromosomal rearrangements during human epidermal keratinocyte differentiation.« In: *Journal of cellular physiology* 221.1 (2009), S. 139–46.
- [253] M. Marotta u. a. »A common copy-number breakpoint of ERBB2 amplification in breast cancer colocalizes with a complex block of segmental duplications.« In: *Breast cancer research : BCR* 14.6 (2012), R150.
- [254] T. Marques-Bonet und E. E. Eichler. »The evolution of human segmental duplications and the core duplicon hypothesis.« In: *Cold Spring Harbor symposia on quantitative biology* 74 (2009), S. 355–62.
- [255] T. Marques-Bonet, S. Girirajan und E. E. Eichler. »The origins and impact of primate segmental duplications.« In: *Trends in genetics : TIG* 25.10 (2009), S. 443–54.
- [256] T. Marques-Bonet u. a. »A burst of segmental duplications in the genome of the African great ape ancestor.« In: *Nature* 457.7231 (2009), S. 877–81.
- [257] T. Marstrand u. a. »Validation of a diagnostic microRNA classifier in cutaneous T-cell lymphomas.« In: *Leukemia & lymphoma* 55.4 (2014), S. 957–8.
- [258] J. W. Martin u. a. »Analysis of segmental duplications, mouse genome synteny and recurrent cancer-associated amplicons in human chromosome 6p21-p12.« In: *Cytogenetic and genome research* 128.4 (2010), S. 199–213.
- [259] J. Mateos-Langerak u. a. »Spatially confined folding of chromatin in the interphase nucleus.« In: *Proceedings of the National Academy of Sciences of the United States of America* 106.10 (2009), S. 3812–7.
- [260] S. Mathas u. a. »Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma.« In: *Proceedings of the National Academy of Sciences of the United States of America* 106.14 (2009), S. 5831–6.
- [261] S. Mathas u. a. »Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma.« In: *Nature immunology* 7.2 (2006), S. 207–15.
- [262] V Matys u. a. »TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes.« In: *Nucleic acids research* 34.Database issue (2006), S. D108–10.
- [263] R. P. McCord u. a. »Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome.« In: *Genome research* 23.2 (2013), S. 260–9.
- [264] K. J. Meaburn, T. Misteli und E. Soutoglou. »Spatial genome organization in the formation of chromosomal translocations.« In: *Seminars in cancer biology* 17.1 (2007), S. 80–90.
- [265] K. J. Meaburn u. a. »Disease-specific gene repositioning in breast cancer.« In: *The Journal of cell biology* 187.6 (2009), S. 801–12.
- [266] P. Medstrand, L. N. van de Lagemaat und D. L. Mager. »Retroelement distributions in the human genome: variations associated with age and proximity to genes.« In: *Genome research* 12.10 (2002), S. 1483–95.

- [267] L Melchor u. a. »Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes.« In: *Oncogene* 27.22 (2008), S. 3165–75.
- [268] X Meng u. a. »Complete physical map of the common deletion region in Williams syndrome and identification and characterization of three novel genes.« In: *Human genetics* 103.5 (1998), S. 590–9.
- [269] M. Merkschlager und D. T. Odom. »CTCF and cohesin: linking gene regulatory elements with their targets.« In: *Cell* 152.6 (2013), S. 1285–97.
- [270] G. Merla u. a. »Copy number variants at Williams-Beuren syndrome 7q11.23 region.« In: *Human genetics* 128.1 (2010), S. 3–26.
- [271] G. Merla u. a. »Molecular Genetics of Williams-Beuren Syndrome«. In: *eLS*. Chichester, UK: John Wiley & Sons, Ltd, 2012.
- [272] W. Meuleman u. a. »Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence.« In: *Genome research* 23.2 (2013), S. 270–80.
- [273] L. R. Meyer u. a. »The UCSC Genome Browser database: extensions and updates 2013.« In: *Nucleic acids research* 41.Database issue (2013), S. D64–9.
- [274] H. Mi u. a. »PANTHER version 10: expanded protein families and functions, and analysis tools.« In: *Nucleic acids research* 44.D1 (2016), S. D336–42.
- [275] K. D. Mills, D. O. Ferguson und F. W. Alt. »The role of DNA breaks in genomic instability and tumorigenesis.« In: *Immunological reviews* 194 (2003), S. 77–95.
- [276] L. A. Mirny. »The fractal globule as a model of chromatin architecture in the cell.« In: *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 19.1 (2011), S. 37–51.
- [277] T. Misteli. »Beyond the sequence: cellular organization of genome function.« In: *Cell* 128.4 (2007), S. 787–800.
- [278] T. Misteli. »Higher-order genome organization in human disease.« In: *Cold Spring Harbor perspectives in biology* 2.8 (2010), a000794.
- [279] T. Misteli. »Self-organization in the genome.« In: *Proceedings of the National Academy of Sciences of the United States of America* 106.17 (2009), S. 6885–6.
- [280] F Mitelman. »Recurrent chromosome aberrations in cancer.« In: *Mutation research* 462.2-3 (2000), S. 247–53.
- [281] F. Mitelman, B. Johansson und F. Mertens. »The impact of translocations and gene fusions on cancer causation.« In: *Nature reviews. Cancer* 7.4 (2007), S. 233–45.
- [282] H. W. Moon u. a. »Incidence of submicroscopic deletions vary according to disease entities and chromosomal translocations in hematologic malignancies: investigation by fluorescence in situ hybridization.« In: *Cancer genetics and cytogenetics* 175.2 (2007), S. 166–8.
- [283] J.-S. Moon u. a. »mTORC1-Induced HK1-Dependent Glycolysis Regulates NLRP3 Inflammasome Activation.« In: *Cell reports* 12.1 (2015), S. 102–15.

- [284] L. Mora u. a. »Chromosome territory positioning of conserved homologous chromosomes in different primate species.« In: *Chromosoma* 115.5 (2006), S. 367–75.
- [285] M. a. Morrow u. a. »Overexpression of the Helix-Loop-Helix protein Id2 blocks T cell development at multiple stages.« In: *Molecular immunology* 36.8 (1999), S. 491–503.
- [286] R. Mostoslavsky, F. W. Alt und C. H. Bassing. »Chromatin dynamics and locus accessibility in the immune system.« In: *Nature immunology* 4.7 (2003), S. 603–6.
- [287] S. Müller u. a. »The evolutionary history of human chromosome 7.« In: *Genomics* 84.3 (2004), S. 458–67.
- [288] W. J. Murphy u. a. »Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps.« In: *Science (New York, N.Y.)* 309.5734 (2005), S. 613–7.
- [289] T. Nagano u. a. »Single-cell Hi-C reveals cell-to-cell variability in chromosome structure.« In: *Nature* 502.7469 (2013), S. 59–64.
- [290] C. Naughton u. a. »Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures.« In: *Nature structural & molecular biology* 20.3 (2013), S. 387–95.
- [291] N. Naumova u. a. »Organization of the mitotic chromosome.« In: *Science (New York, N.Y.)* 342.6161 (2013), S. 948–53.
- [292] M. Neusser u. a. »Evolutionarily conserved, cell type and species-specific higher order chromatin arrangements in interphase nuclei of primates.« In: *Chromosoma* 116.3 (2007), S. 307–20.
- [293] H. Neves u. a. »The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation.« In: *Blood* 93.4 (1999), S. 1197–207.
- [294] W. W. Nichols u. a. »Characterization of a new human diploid cell line–IMR-91.« In: *In vitro* 19.10 (1983), S. 797–804.
- [295] W. W. Nichols u. a. »Characterization of a new human diploid cell strain, IMR-90.« In: *Science (New York, N.Y.)* 196.4285 (1977), S. 60–3.
- [296] J. M. Nicholson und D. Cimini. »Cancer karyotypes: survival of the fittest.« In: *Frontiers in oncology* 3.June (2013), S. 148.
- [297] M. N. Nikiforova u. a. »Proximity of chromosomal loci that participate in radiation-induced rearrangements in human cells.« In: *Science (New York, N.Y.)* 290.5489 (2000), S. 138–41.
- [298] Y. Nishino u. a. »Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure.« In: *The EMBO journal* 31.7 (2012), S. 1644–53.
- [299] E. P. Nora u. a. »Spatial partitioning of the regulatory landscape of the X-inactivation centre.« In: *Nature* 485.7398 (2012), S. 381–5.
- [300] E. Nunez, X.-D. Fu und M. G. Rosenfeld. »Nuclear organization in the 3D space of the nucleus - cause or consequence?« In: *Current opinion in genetics & development* 19.5 (2009), S. 424–36.
- [301] N. A. O’Leary u. a. »Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.« In: *Nucleic acids research* 44.D1 (2016), S. D733–45.

- [302] A. B. Olshen u. a. »Circular binary segmentation for the analysis of array-based DNA copy number data.« In: *Biostatistics (Oxford, England)* 5.4 (2004), S. 557–72.
- [303] M. Onishi-Seebacher und J. O. Korb. »Challenges in studying genomic structural variant formation mechanisms: the short-read dilemma and beyond.« In: *BioEssays : news and reviews in molecular, cellular and developmental biology* 33.11 (2011), S. 840–50.
- [304] C. S. Osborne. »Molecular pathways: transcription factories and chromosomal translocations.« In: *Clinical cancer research : an official journal of the American Association for Cancer Research* 20.2 (2014), S. 296–300.
- [305] L. R. Osborne u. a. »PMS2-related genes flank the rearrangement breakpoints associated with Williams syndrome and other diseases on human chromosome 7.« In: *Genomics* 45.2 (1997), S. 402–6.
- [306] K. Osoegawa u. a. »A bacterial artificial chromosome library for sequencing the complete human genome.« In: *Genome research* 11.3 (2001), S. 483–96.
- [307] D. Ottaviani, M. LeCain und D. Sheer. »The role of microhomology in genomic structural variation.« In: *Trends in genetics : TIG* 30.3 (2014), S. 85–94.
- [308] L. a. Otten u. a. »Deregulated MHC class II transactivator expression leads to a strong Th2 bias in CD4+ T lymphocytes.« In: *Journal of immunology (Baltimore, Md. : 1950)* 170.3 (2003), S. 1150–7.
- [309] N. C. Ottoson u. a. »Cutting edge: T cell migration regulated by CXCR4 chemokine receptor signaling to ZAP-70 tyrosine kinase.« In: *Journal of immunology (Baltimore, Md. : 1950)* 167.4 (2001), S. 1857–61.
- [310] L. a. Parada, P. G. McQueen und T. Misteli. »Tissue-specific spatial organization of genomes.« In: *Genome biology* 5.7 (2004), R44.
- [311] S. T. Park, G. P. Nolan und X. H. Sun. »Growth inhibition and apoptosis due to restoration of E2A activity in T cell acute lymphoblastic leukemia cells.« In: *The Journal of experimental medicine* 189.3 (1999), S. 501–8.
- [312] A. L. Paterson u. a. »Mobile element insertions are frequent in oesophageal adenocarcinomas and can mislead paired-end sequencing analysis.« In: *BMC genomics* 16.1 (2015), S. 473.
- [313] A. Pavlicek u. a. »Traffic of genetic information between segmental duplications flanking the typical 22q11.2 deletion in velo-cardio-facial syndrome/DiGeorge syndrome.« In: *Genome research* 15.11 (2005), S. 1487–95.
- [314] N. Paweletz. »Walther Flemming: pioneer of mitosis research.« In: *Nature reviews. Molecular cell biology* 2.1 (2001), S. 72–5.
- [315] R. Peoples u. a. »A physical map, including a BAC/PAC clone contig, of the Williams-Beuren syndrome-deletion region at 7q11.23.« In: *American journal of human genetics* 66.1 (2000), S. 47–68.
- [316] L. a. Pérez Jurado u. a. »Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth.« In: *American journal of human genetics* 59.4 (1996), S. 781–92.

- [317] D. Peric-Hupkes u. a. »Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation.« In: *Molecular cell* 38.4 (2010), S. 603–13.
- [318] B. Phipson und G. K. Smyth. »Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn.« In: *Statistical applications in genetics and molecular biology* 9.1 (2010), Article39.
- [319] D. Pinkel u. a. »Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4.« In: *Proceedings of the National Academy of Sciences of the United States of America* 85.23 (1988), S. 9138–42.
- [320] D. Pinkel u. a. »High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.« In: *Nature genetics* 20.2 (1998), S. 207–11.
- [321] B. R. Pober. »Williams-Beuren syndrome.« In: *The New England journal of medicine* 362.3 (2010), S. 239–52.
- [322] A. Pombo und N. Dillon. »Three-dimensional genome architecture: players and mechanisms.« In: *Nature reviews. Molecular cell biology* 16.4 (2015), S. 245–57.
- [323] M. Popovic u. a. »Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS.« In: *Science (New York, N.Y.)* 224.4648 (1984), S. 497–500.
- [324] X. Qiao u. a. »Different effects of ZO-1, ZO-2 and ZO-3 silencing on kidney collecting duct principal cell proliferation and adhesion.« In: *Cell cycle (Georgetown, Tex.)* 13.19 (2014), S. 3059–75.
- [325] J. Z. Qin u. a. »Interleukin-7 and interleukin-15 regulate the expression of the bcl-2 and c-myc genes in cutaneous T-cell lymphoma cells.« In: *Blood* 98.9 (2001), S. 2778–83.
- [326] A. R. Quinlan und I. M. Hall. »BEDTools: a flexible suite of utilities for comparing genomic features.« In: *Bioinformatics (Oxford, England)* 26.6 (2010), S. 841–2.
- [327] A. R. Quinlan und I. M. Hall. »Characterizing complex structural variation in germline and somatic genomes.« In: *Trends in genetics : TIG* 28.1 (2012), S. 43–53.
- [328] R R Development Core Team. *Computational Many-Particle Physics*. Hrsg. von H. Fehske, R. Schneider und A. Weiße. Bd. 739. Lecture Notes in Physics 2.11.1. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, S. 409.
- [329] RStudio. *RStudio: Integrated development environment for R*. Boston, MA, 2012.
- [330] C. Rabl. »Über Zellteilung«. In: *Morph Jb* 10 (1885), S. 214–330.
- [331] U. Ralfkiaer u. a. »Diagnostic microRNA profiling in cutaneous T-cell lymphoma (CTCL).« In: *Blood* 118.22 (2011), S. 5891–900.
- [332] M. B. Ramocki u. a. »Recurrent distal 7q11.23 deletion including HIP1 and YWHAG identified in patients with intellectual disabilities, epilepsy, and neurobehavioral problems.« In: *American journal of human genetics* 87.6 (2010), S. 857–65.
- [333] S. S. P. Rao u. a. »A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.« In: *Cell* 159.7 (2014), S. 1665–80.

- [334] K. L. Reddy u. a. »Transcriptional repression mediated by repositioning of genes to the nuclear lamina.« In: *Nature* 452.7184 (2008), S. 243–7.
- [335] M. C. Reddy, J. Christensen und K. M. Vasquez. »Interplay between human high mobility group protein 1 and replication protein A on psoralen-cross-linked DNA.« In: *Biochemistry* 44.11 (2005), S. 4188–95.
- [336] L. T. Reiter u. a. »A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element.« In: *Nature genetics* 12.3 (1996), S. 288–97.
- [337] K. Rennstam u. a. »Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization.« In: *Cancer research* 63.24 (2003), S. 8861–8.
- [338] K. V. Revanna u. a. »GSV: a web-based genome synteny viewer for customized data.« In: *BMC bioinformatics* 12.1 (2011), S. 316.
- [339] D. Rhodes und H. J. Lipps. »G-quadruplexes and their regulatory roles in biology.« In: *Nucleic acids research* 43.18 (2015), S. 8627–37.
- [340] T. J. Richmond und C. a. Davey. »The structure of DNA in the nucleosome core.« In: *Nature* 423.6936 (2003), S. 145–50.
- [341] D. Rieder, Z. Trajanoski und J. G. McNally. »Transcription factories.« In: *Frontiers in genetics* 3.October (2012), S. 221.
- [342] R. J. Roberts u. a. »REBASE—a database for DNA restriction and modification: enzymes, genes and genomes.« In: *Nucleic acids research* 38.Database issue (2010), S. D234–6.
- [343] P. J. J. Robinson u. a. »EM measurements define the dimensions of the "30-nm"chromatin fiber: evidence for a compact, interdigitated structure.« In: *Proceedings of the National Academy of Sciences of the United States of America* 103.17 (2006), S. 6506–11.
- [344] W. P. Robinson u. a. »Delineation of 7q11.2 deletions associated with Williams-Beuren syndrome and mapping of a repetitive sequence to within and to either side of the common deletion.« In: *Genomics* 34.1 (1996), S. 17–23.
- [345] A. L. Rodd, K. Ververis und T. C. Karagiannis. »Current and Emerging Therapeutics for Cutaneous T-Cell Lymphoma: Histone Deacetylase Inhibitors.« In: *Lymphoma* 2012 (2012), S. 1–10.
- [346] J. J. Roix u. a. »Spatial proximity of translocation-prone gene loci in human lymphomas.« In: *Nature genetics* 34.3 (2003), S. 287–91.
- [347] A. V. Roschke u. a. »Karyotypic complexity of the NCI-60 drug-screening panel.« In: *Cancer research* 63.24 (2003), S. 8634–47.
- [348] V. Roukos, B. Burman und T. Misteli. »The cellular etiology of chromosome translocations.« In: *Current opinion in cell biology* 25.3 (2013), S. 357–64. arXiv: [NIHMS150003](https://arxiv.org/abs/150003).
- [349] A. Ruiz-Herrera und T. J. Robinson. »Evolutionary plasticity and cancer breakpoints in human chromosome 3.« In: *BioEssays : news and reviews in molecular, cellular and developmental biology* 30.11-12 (2008), S. 1126–37.

- [350] A. J. Saldanha. »Java Treeview–extensible visualization of microarray data.« In: *Bioinformatics (Oxford, England)* 20.17 (2004), S. 3246–8.
- [351] R. V. Samonte und E. E. Eichler. »Segmental duplications and the evolution of the primate genome.« In: *Nature reviews. Genetics* 3.1 (2002), S. 65–72.
- [352] J. Sandoval u. a. »MicroRNA expression profiling and DNA methylation signature for deregulated microRNA in cutaneous T-cell lymphoma.« In: *The Journal of investigative dermatology* 135.4 (2015), S. 1128–37.
- [353] D. Sankoff. »The where and wherefore of evolutionary breakpoints.« In: *Journal of biology* 8.7 (2009), S. 66.
- [354] M. Scales u. a. »visPIG—a web tool for producing multi-region, multi-track, multi-scale plots of genetic data.« In: *PloS one* 9.9 (2014), e107497.
- [355] A. Scally u. a. »Insights into hominid evolution from the gorilla genome sequence.« In: *Nature* 483.7388 (2012), S. 169–75.
- [356] S. W. Scherer u. a. »Human chromosome 7: DNA sequence and biology.« In: *Science (New York, N.Y.)* 300.5620 (2003), S. 767–72.
- [357] R. Schneider und R. Grosschedl. »Dynamics and interplay of nuclear architecture, genome organization, and gene expression.« In: *Genes & development* 21.23 (2007), S. 3027–43.
- [358] U Schneider, H. U. Schwenk und G Bornkamm. »Characterization of EBV-genome negative "nulländ "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma.« In: *International journal of cancer* 19.5 (1977), S. 621–6.
- [359] E Schröck und H Padilla-Nash. »Spectral karyotyping and multicolor fluorescence in situ hybridization reveal new tumor-specific chromosomal aberrations.« In: *Seminars in hematology* 37.4 (2000), S. 334–47.
- [360] C. Schubert. »The genomic basis of the Williams-Beuren syndrome.« In: *Cellular and molecular life sciences : CMLS* 66.7 (2009), S. 1178–97.
- [361] M. Schwartz und O. Hakim. »3D view of chromosomes, DNA damage, and translocations.« In: *Current opinion in genetics & development* 25 (2014), S. 118–25.
- [362] S. Schwartz u. a. »Human-mouse alignments with BLASTZ.« In: *Genome research* 13.1 (2003), S. 103–7.
- [363] S. Selvaraj u. a. »Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing.« In: *Nature biotechnology* 31.12 (2013), S. 1111–8.
- [364] D. Sen und W. Gilbert. »Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis.« In: *Nature* 334.6180 (1988), S. 364–6.
- [365] F. Serra u. a. *Structural features of the fly chromatin colors revealed by automatic three-dimensional modeling*. Techn. Ber. 2016, S. 1–29.
- [366] J. a. Shapiro und R. von Sternberg. »Why repetitive DNA is essential to genome function.« In: *Biological reviews of the Cambridge Philosophical Society* 80.2 (2005), S. 227–50.

- [367] A. J. Sharp u. a. »Segmental duplications and copy-number variation in the human genome.« In: *American journal of human genetics* 77.1 (2005), S. 78–88.
- [368] N. Shimazaki, A. G. Tsai und M. R. Lieber. »H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in trans in addition to tethering in cis: implications for translocations.« In: *Molecular cell* 34.5 (2009), S. 535–44. arXiv: [NIHMS150003](https://arxiv.org/abs/150003).
- [369] N. A. Sibbesen u. a. »Jak3, STAT3, and STAT5 inhibit expression of miR-22, a novel tumor suppressor microRNA, in cutaneous T-Cell lymphoma.« In: *Oncotarget* 6.24 (2015), S. 20555–69.
- [370] M Sigvardsson, M O’Riordan und R Grosschedl. »EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes.« In: *Immunity* 7.1 (1997), S. 25–36.
- [371] M Sigvardsson, M O’Riordan und R Grosschedl. »EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes.« In: *Immunity* 7.1 (1997), S. 25–36.
- [372] J. Sima und D. M. Gilbert. »Complex correlations: replication timing and mutational landscapes during cancer and genome evolution.« In: *Current opinion in genetics & development* 25 (2014), S. 93–100.
- [373] M. Simonis u. a. »High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology.« In: *Nature methods* 6.11 (2009), S. 837–42.
- [374] P. B. Sinclair u. a. »Large deletions at the t(9;22) breakpoint are common and may identify a poor-prognosis subgroup of patients with chronic myeloid leukemia.« In: *Blood* 95.3 (2000), S. 738–43.
- [375] J. L. Slightom, a. E. Blechl und O Smithies. »Human fetal G gamma- and A gamma-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes.« In: *Cell* 21.3 (1980), S. 627–38.
- [376] C. A. Sloan u. a. »ENCODE data at the ENCODE portal.« In: *Nucleic acids research* 44.D1 (2016), S. D726–32.
- [377] A. Smit, R Hubley und P Green. *RepeatMasker Open-3.0*.
- [378] A. M. Snijders u. a. »Shaping of tumor and drug-resistant genomes by instability and selection.« In: *Oncogene* 22.28 (2003), S. 4370–9.
- [379] S Solinas-Toldo u. a. »Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances.« In: *Genes, chromosomes & cancer* 20.4 (1997), S. 399–407.
- [380] I. Solovei u. a. »Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution.« In: *Cell* 137.2 (2009), S. 356–68.
- [381] M. J. Somerville u. a. »Severe expressive-language delay related to duplication of the Williams-Beuren locus.« In: *The New England journal of medicine* 353.16 (2005), S. 1694–701.
- [382] A. Sors u. a. »Down-regulating constitutive activation of the NF-kappaB canonical pathway overcomes the resistance of cutaneous T-cell lymphoma to apoptosis.« In: *Blood* 107.6 (2006), S. 2354–63.

- [383] P. T. Spellman und G. M. Rubin. »Evidence for large domains of similarly expressed genes in the *Drosophila* genome.« In: *Journal of biology* 1.1 (2002), S. 5.
- [384] P. Stankiewicz und J. R. Lupski. »Genome architecture, rearrangements and genomic disorders.« In: *Trends in genetics : TIG* 18.2 (2002), S. 74–82.
- [385] P. Stankiewicz und J. R. Lupski. »Structural variation in the human genome and its role in disease.« In: *Annual review of medicine* 61 (2010), S. 437–55.
- [386] A. Stanlie u. a. »Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination.« In: *Proceedings of the National Academy of Sciences of the United States of America* 107.51 (2010), S. 22190–5.
- [387] G. Starkebaum u. a. »Establishment of an IL-2 independent, human T-cell line possessing only the p70 IL-2 receptor.« In: *International journal of cancer* 49.2 (1991), S. 246–53.
- [388] C. Steidl u. a. »MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers.« In: *Nature* 471.7338 (2011), S. 377–81.
- [389] A. Steininger u. a. »Genomic loss of the putative tumor suppressor gene E2A in human lymphoma.« In: *The Journal of experimental medicine* 208.8 (2011), S. 1585–93.
- [390] P. J. Stephens u. a. »Complex landscapes of somatic rearrangement in human breast cancer genomes.« In: *Nature* 462.7276 (2009), S. 1005–10.
- [391] P. J. Stephens u. a. »Massive genomic rearrangement acquired in a single catastrophic event during cancer development.« In: *Cell* 144.1 (2011), S. 27–40.
- [392] C. T. Storlazzi u. a. »Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure.« In: *Genome research* 20.9 (2010), S. 1198–206.
- [393] K. Stout u. a. »Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease?« In: *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 7.5 (1999), S. 323–9.
- [394] M. R. Stratton, P. J. Campbell und P. A. Futreal. »The cancer genome.« In: *Nature* 458.7239 (2009), S. 719–24.
- [395] P. Strømme, P. G. Bjørnstad und K. Ramstad. »Prevalence estimation of Williams syndrome.« In: *Journal of child neurology* 17.4 (2002), S. 269–71.
- [396] A. Sturn, J. Quackenbush und Z. Trajanoski. »Genesis: cluster analysis of microarray data.« In: *Bioinformatics (Oxford, England)* 18.1 (2002), S. 207–8.
- [397] P. H. Sudmant u. a. »Diversity of human copy number variation and multicopy genes.« In: *Science (New York, N.Y.)* 330.6004 (2010), S. 641–6.
- [398] N. Sugawara, G. Ira und J. E. Haber. »DNA length dependence of the single-strand annealing pathway and the role of *Saccharomyces cerevisiae* RAD59 in double-strand break repair.« In: *Molecular and cellular biology* 20.14 (2000), S. 5300–9.
- [399] D. Szklarczyk u. a. »STRING v10: protein-protein interaction networks, integrated over the tree of life.« In: *Nucleic acids research* 43.Database issue (2015), S. D447–52.

- [400] M. Takata u. a. »Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells.« In: *The EMBO journal* 17.18 (1998), S. 5497–508.
- [401] H. Tanabe u. a. »Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates.« In: *Proceedings of the National Academy of Sciences of the United States of America* 99.7 (2002), S. 4424–9.
- [402] A. Tanaka u. a. »Epigenetic regulation of condensin-mediated genome organization during the cell cycle and upon DNA damage through histone H3 lysine 56 acetylation.« In: *Molecular cell* 48.4 (2012), S. 532–46.
- [403] S.-J. Tang. »Chromatin Organization by Repetitive Elements (CORE): A Genomic Principle for the Higher-Order Structure of Chromosomes.« In: *Genes* 2.3 (2011), S. 502–15.
- [404] N. A. Tchurikov u. a. »Hot spots of DNA double-strand breaks and genomic contacts of human rDNA units are involved in epigenetic regulation.« In: *Journal of molecular cell biology* 7.4 (2015), S. 366–82.
- [405] R. Teachenor u. a. »Biochemical and phosphoproteomic analysis of the helix-loop-helix protein E47.« In: *Molecular and cellular biology* 32.9 (2012), S. 1671–82.
- [406] L. Teng u. a. »4DGenome: a comprehensive database of chromatin interactions.« In: *Bioinformatics (Oxford, England)* 31.15 (2015), S. 2560–4.
- [407] M Thangavelu u. a. »Recurring structural chromosome abnormalities in peripheral blood lymphocytes of patients with mycosis fungoides/Sézary syndrome.« In: *Blood* 89.9 (1997), S. 3371–7.
- [408] L. I. Toledo u. a. »ATR prohibits replication catastrophe by preventing global exhaustion of RPA.« In: *Cell* 155.5 (2013), S. 1088–103.
- [409] A. Travers und G. Muskhelishvili. »DNA structure and function.« In: *The FEBS journal* 282.12 (2015), S. 2279–95.
- [410] T. J. Treangen und S. L. Salzberg. »Repetitive DNA and next-generation sequencing: computational challenges and solutions.« In: *Nature reviews. Genetics* 13.1 (2012), S. 36–46.
- [411] J. M. C. Tubio. »Somatic structural variation and cancer.« In: *Briefings in functional genomics* 14.5 (2015), S. 339–51.
- [412] D. J. Turner u. a. »Germline rates of de novo meiotic deletions and duplications causing several genomic disorders.« In: *Nature Genetics* 40.1 (2008), S. 90–95.
- [413] D. D. W. Twa u. a. »Recurrent genomic rearrangements in primary testicular lymphoma.« In: *The Journal of pathology* 236.2 (2015), S. 136–41.
- [414] M. Uddin u. a. »Genome-wide signatures of 'rearrangement hotspots' within segmental duplications in humans.« In: *PloS one* 6.12 (2011), e28853.
- [415] S. V. Ulianov u. a. »Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains.« In: *Genome research* 7.499 (2015).

- [416] A. Ungewickell u. a. »Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2.« In: *Nature genetics* 47.9 (2015), S. 1056–60.
- [417] M. C. Valero u. a. »Fine-scale comparative mapping of the human 7q11.23 region and the orthologous region on mouse chromosome 5G: the low-copy repeats that flank the Williams-Beuren syndrome deletion arose at breakpoint sites of an evolutionary inversion(s).« In: *Genomics* 69.1 (2000), S. 1–13.
- [418] N. Van der Aa u. a. »Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome.« In: *European journal of medical genetics* 52.2-3 (2009), S. 94–100.
- [419] J. P. Vaqué u. a. »PLCG1 mutations in cutaneous T-cell lymphomas.« In: *Blood* 123.13 (2014), S. 2034–43.
- [420] N. Vázquez-Salat u. a. »Gene conversion between mammalian CCR2 and CCR5 chemokine receptor genes: a potential mechanism for receptor dimerization.« In: *Genomics* 90.2 (2007), S. 213–24.
- [421] W. N. Venables und B. D. Ripley. *Modern Applied Statistics with S*. Statistics and Computing. New York, NY: Springer New York, 2002.
- [422] U. Verbovšek, C. J. F. Van Noorden und T. T. Lah. »Complexity of cancer protease biology: Cathepsin K expression and function in cancer progression.« In: *Seminars in cancer biology* 35 (2015), S. 71–84.
- [423] M. H. Vermeer u. a. »Novel and highly recurrent chromosomal alterations in Sézary syndrome.« In: *Cancer research* 68.8 (2008), S. 2689–98.
- [424] A. S. Véron u. a. »Close 3D proximity of evolutionary breakpoints argues for the notion of spatial synteny.« In: *BMC genomics* 12.1 (2011), S. 303.
- [425] R. Versteeg u. a. »The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes.« In: *Genome research* 13.9 (2003), S. 1998–2004.
- [426] M. Vietri Rudan u. a. »Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture.« In: *Cell reports* 10.8 (2015), S. 1297–309.
- [427] A. Vologodskii. »Brownian dynamics simulation of knot diffusion along a stretched DNA molecule.« In: *Biophysical journal* 90.5 (2006), S. 1594–7.
- [428] Y. Y. Wan. »GATA3: a master of many trades in immune regulation.« In: *Trends in immunology* 35.6 (2014), S. 233–42.
- [429] C. Wang und S. Huang. »Nuclear function of Alus.« In: *Nucleus (Austin, Tex.)* 5.2 (2014), S. 131–7.
- [430] C. Wang u. a. »Genome-wide analysis of local chromatin packing in *Arabidopsis thaliana*.« In: *Genome research* 25.2 (2015), S. 246–56.
- [431] M. Watanabe u. a. »Structural analysis of the interaction between Shiga toxin B subunits and linear polymers bearing clustered globotriose residues.« In: *Infection and immunity*. Statistics for Biology and Health 74.3 (2006). Hrsg. von R. Gentleman u. a., S. 1984–8.

- [432] B. Weckselblatt und M. K. Rudd. »Human Structural Variation: Mechanisms of Chromosome Rearrangements.« In: *Trends in genetics : TIG* 31.10 (2015), S. 587–99.
- [433] D. Weese u. a. »RazerS–fast read mapping with sensitivity control.« In: *Genome research* 19.9 (2009), S. 1646–54.
- [434] C. Weierich u. a. »Three-dimensional arrangements of centromeres and telomeres in nuclei of human and murine lymphocytes.« In: *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 11.5 (2003), S. 485–502.
- [435] J. Weischenfeldt u. a. »Phenotypic impact of genomic structural variation: insights from and for human disease.« In: *Nature reviews. Genetics* 14.2 (2013), S. 125–38.
- [436] B. Wen u. a. »Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.« In: *Nature genetics* 41.2 (2009), S. 246–50.
- [437] H. Wickham. »Reshaping Data with the reshape Package«. In: *Journal of Statistical Software* 21.12 (2007).
- [438] H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.
- [439] P. J. Wijchers und W. de Laat. »Genome organization influences partner selection for chromosomal rearrangements.« In: *Trends in genetics : TIG* 27.2 (2011), S. 63–71.
- [440] R. Willemze. »XV. Primary cutaneous lymphomas«. In: *Annals of Oncology* 22.Supplement 4 (2011), S. iv72–iv75.
- [441] R. Willemze u. a. »WHO-EORTC classification for cutaneous lymphomas.« In: *Blood* 105.10 (2005), S. 3768–85.
- [442] E. de Wit und W. de Laat. »A decade of 3C technologies: insights into nuclear organization.« In: *Genes & development* 26.1 (2012), S. 11–24.
- [443] C. L. Woodcock. »Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length.« In: *The Journal of cell biology* 125.1 (1994), S. 11–9.
- [444] C. L. Woodcock, J. P. Safer und J. E. Stanchfield. »Structural repeating units in chromatin. I. Evidence for their general occurrence.« In: *Experimental cell research* 97.1 (1976), S. 101–10.
- [445] A. Worcel, S. Strogatz und D. Riley. »Structure of chromatin and the linking number of DNA.« In: *Proceedings of the National Academy of Sciences of the United States of America* 78.3 (1981), S. 1461–5.
- [446] J. Wu u. a. »Structural alterations of the FAS gene in cutaneous T-cell lymphoma (CTCL).« In: *Archives of biochemistry and biophysics* 508.2 (2011), S. 185–91.
- [447] H. Xu u. a. »ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells.« In: *Database : the journal of biological databases and curation* 2013 (2013), bat045.
- [448] E. Yaffe und A. Tanay. »Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture.« In: *Nature genetics* 43.11 (2011), S. 1059–65.

- [449] W Yan u. a. »High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice.« In: *Molecular and cellular biology* 17.12 (1997), S. 7317–27.
- [450] D. H. Yasui u. a. »15q11.2-13.3 chromatin analysis reveals epigenetic regulation of CHRNA7 with deficiencies in Rett and autism brain.« In: *Human molecular genetics* 20.22 (2011), S. 4311–23.
- [451] A. Yates u. a. »Ensembl 2016.« In: *Nucleic acids research* 44.D1 (2016), S. D710–6.
- [452] N. Yawalkar u. a. »Profound loss of T-cell receptor repertoire complexity in cutaneous T-cell lymphoma.« In: *Blood* 102.12 (2003), S. 4059–66.
- [453] J. J. Yunis und O Prakash. »The origin of man: a chromosomal pictorial legacy.« In: *Science (New York, N.Y.)* 215.4539 (1982), S. 1525–30.
- [454] C.-z. Zhang, M. L. Leibowitz und D. Pellman. »Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements.« In: *Genes & development* 27.23 (2013), S. 2513–30.
- [455] Y. Zhang u. a. »Spatial organization of the mouse genome and its role in recurrent chromosomal translocations.« In: *Cell* 148.5 (2012), S. 908–21.
- [456] H. Zhao und G. Bourque. »Recovering genome rearrangements in the mammalian phylogeny.« In: *Genome research* 19.5 (2009), S. 934–42.
- [457] W. Zhou u. a. »Increased genome instability in human DNA segments with self-chains: homology-induced structural variations via replicative mechanisms.« In: *Human molecular genetics* 22.13 (2013), S. 2642–51.
- [458] X. Zhou u. a. »The Human Epigenome Browser at Washington University.« In: *Nature methods* 8.12 (2011), S. 989–90.
- [459] A. C. da Silva Almeida u. a. »The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome.« In: *Nature genetics* 47.12 (2015), S. 1465–70.

6. Verzeichnis der aus dieser Arbeit erfolgten Publikationen

A. Steininger*, M. Möbs*, R. Ullmann, K. Köchert, S. Kreher, B. Lamprecht, I. Anagnostopoulos, M. Hummel, J. Richter, M. Beyer, M. Janz, C.-D. Klemke, H. Stein, B. Dörken, W. Sterry, E. Schröck, S. Mathas, C. Assaf »Genomic loss of the putative tumor suppressor gene E2A in human lymphoma.« In: *The Journal of experimental medicine* 208.8 (2011), S. 1585–93.

G. Ebert*, A. Steininger*, R. Weißmann, V. Boldt, A. Lind-Thomsen, J. Grune, S. Badelt, M. Heßler, M. Peiser, M. Hitzler, L. R. Jensen, I. Müller, H. Hu, P. F. Arndt, A. W. Kuss, K. Tebel, R. Ullmann »Distribution of segmental duplications in the context of higher order chromatin organisation of human chromosome 7.« In: *BMC genomics* 15 (2014), S. 537.

* geteilte Autorenschaft

7. Anhang

Tabelle S1.: Änderung der Binpaaranzahl nach Anwendung der Filterkriterien bei einem Schwellenwert von > 15 IAs/Bin in Kombination mit den verschiedenen Schwellenwerten der Filterkategorien Interaktionsdistanz und SD-Überlapp. Verglichen wurde hierbei mit dem ungefilterten Datensatz (= 100 %). Tab. aus [99].

chr	Verwerfen von Bins mit SD	IAs/Bin	IA-Distanz	Binanzahl	verbleibende Bins (in %)
chr7	-	> 1	-	1301686	100
chr7	paraloge SDs	> 1	-	1301527	99,98
chr7	alle SDs	> 1	-	1174803	90,25
chr7	-	> 15	> 25 Mb	1474	0,11
chr7	paraloge SDs	> 15	> 25 Mb	1472	0,11
chr7	alle SDs	> 15	> 25 Mb	593	0,05

Tabelle S2.: Änderung der Anzahl an interagierenden Bins und resultierenden Interaktionsbündel nach Anwendung der Filterkategorien und allen untersuchten Schwellenwerten.

Datensatz 1-12: Entfernung der SD-Sequenzen aus dem Datensatz auf Binebene, Datensatz 13: Entfernung von nicht eindeutig im Referenzgenom verortbaren SD-Sequenzen aus dem Datensatz auf Readebene (siehe Abb. 3.1 J). Tab. aus [99].

Datensatz	chr	Filterkategorien			Anzahl interagierender Bins nach dem Filtern	Anzahl der Bins zur Generierung der IA-Bündel	Anzahl der resultierenden IA-Bündel	Abdeckung des chr7 durch die IA-Bündel in %	Anzahl an IA-Bündel die paraloge SDs verbinden
		Entfernung von Bins mit SDs	IAs/Bin	IA-Distanz					
1	chr7	-	> 15	> 25 Mb	1474	271	33	23	5
2	chr7	-	> 10	> 25 Mb	3774	966	40	42	10
3	chr7	-	10-50	> 25 Mb	3721	923	38	42	10
4	chr7	paraloge SDs	> 15	> 25 Mb	1472	266	32	23	4
5	chr7	paraloge SDs	> 10	> 25 Mb	3772	965	40	42	10
6	chr7	paraloge SDs	10-50	> 25 Mb	3719	922	38	42	10
7	chr7	alle SDs	> 15	> 25 Mb	593	215	72	20	0
8	chr7	alle SDs	> 10	> 25 Mb	1899	376	33	25	2
9	chr7	alle SDs	10-50	> 25 Mb	1885	366	32	25	2
10	chr7	-	> 15	> 10 Mb	2254	283	27	26	5
11	chr7	paraloge SDs	> 15	> 10 Mb	2252	276	26	26	4
12	chr7	alle SDs	> 15	> 10 Mb	904	181	41	18	1
13	chr7	-	> 15	> 25 Mb	711	77	11	9	1

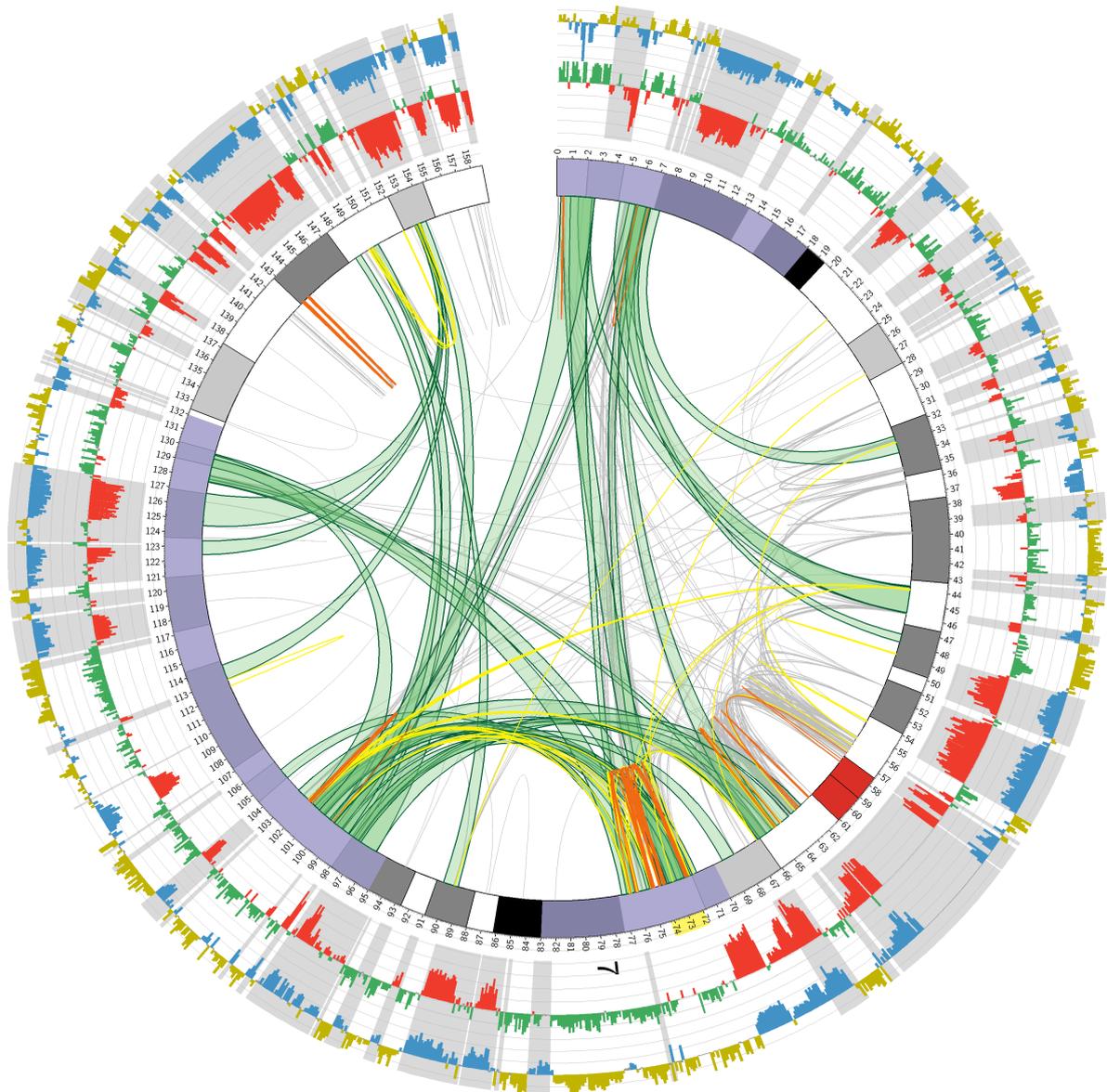


Abbildung S1.: Intrachromosomales Interaktionsmuster des humanen Chromosom 7 überlappt mit Regionen früh-apoptotisch degradiertes DNA.

Regionen intrachromosomaler Interaktion verbinden Domänen mit früh-apoptotischer DNA-Degradation (genomische Daten entsprechen der Zelllinie IMR90, soweit nicht anders vermerkt). **A** Acetylierungsprofil für das Histon H4K8; dunkelgelb: Hyperacetylierung, blau: Hypoacetylierung **B** DNA-Fragmentierung früh-apoptotisch (grün) und nicht degradiert (rot) **C** LAD-Grenzen als grau-unterlegte Intervalle (Tig3-Zellen; Lungenfibroblasten) **D** Idiogramm des Chromosom 7 mit in gelb markierter Williams-Beuren-Deletionsregion (72-74 Mb; hg18) **E** blau-transparente Schattierung markiert die von der Inversion betroffenen Segmente aus Abb. 3.4 **F** Interaktionsbündel als grün-transparente Bänder und **G** segmentale Duplikationen mit < 98 % (grau), 98-99 % (gelb) und > 99 % Sequenzsimilarität (orange).

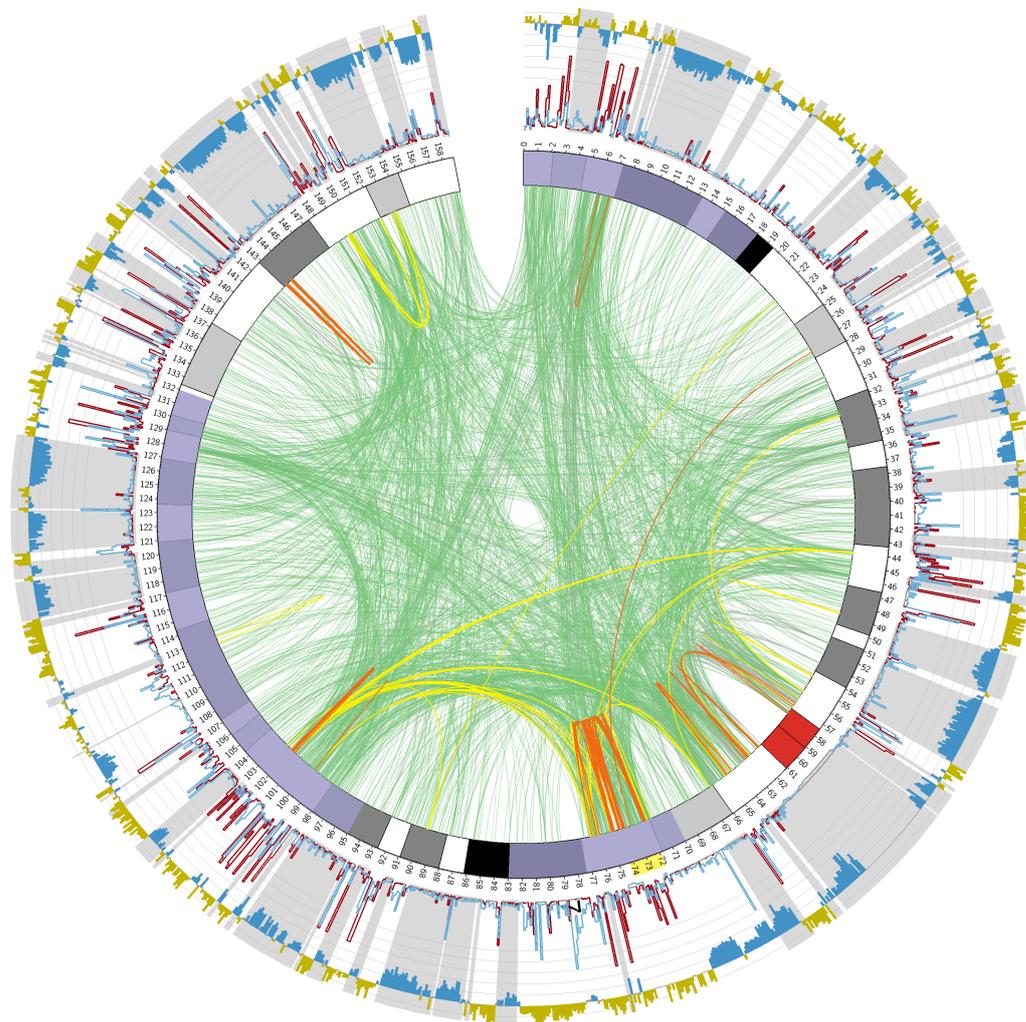


Abbildung S2.: Intrachromosomales Interaktionsmuster des humanen Chromosom 7 ohne Bündelung der Bindaten.

Darstellung entspricht Abb. 3.2 unter Austausch der Interaktionsbündel durch die direkten Interaktionen der gefilterten Binpaare. **A** Acetylierungsprofil für das Histon H4K8; dunkel-gelb: Hyperacetylierung, blau: Hypoacetylierung **B** RNA-Seq Readanzahl pro 100 kb Bin für kodierende (in rot) und nichtkodierende RNA (in blau) der Zelllinie IMR91L **C** LAD-Grenzen als grau-unterlegte Intervalle (Tig3-Zellen; Lungenfibroblasten) **D** Idiogramm des Chromosom 7 mit in gelb markierter Williams-Beuren-Deletionsregion (72-74 Mb; hg18) **E** blau-transparente Schattierung markiert die von der Inversion betroffenen Segmente aus Abb. 3.4 **F** Interagierende 20-kb-Binpaare und **G** segmentale Duplikationen mit < 98 % (grau), 98-99 % (gelb) und > 99 % Sequenzsimilarität (orange). Abb. aus [99].

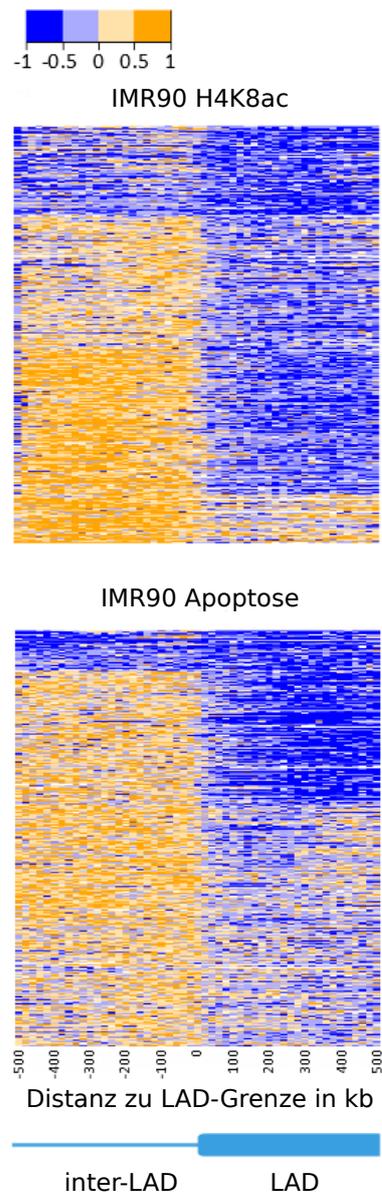


Abbildung S3.: LAD-Grenzen kolokalisieren mit Änderungen des Profils der apoptotischen DNA-Degradation und der Histon H4K8-Acetylierung.

Heatmaps zeigen die Intensitäten der H4K8-Acetylierung und der apoptotischen DNA-Degradation (in 10-kb-Bins) 500 kb stromaufwärts und stromabwärts der LAD-Grenzen. Nur LADs mit einer Minimalgröße von 500 kb und einer Minimaldistanz zur benachbarten LAD von 500 kb werden dargestellt. Bins mit hohen Intensitäten bezüglich apoptotischer DNA-Degradation bzw. H4K8-Hyperacetylierung in orange, Bins mit nicht degradierten DNA-Sequenzen bzw. hypoacetyliertem H4K8 in blau. Abb. aus [99].

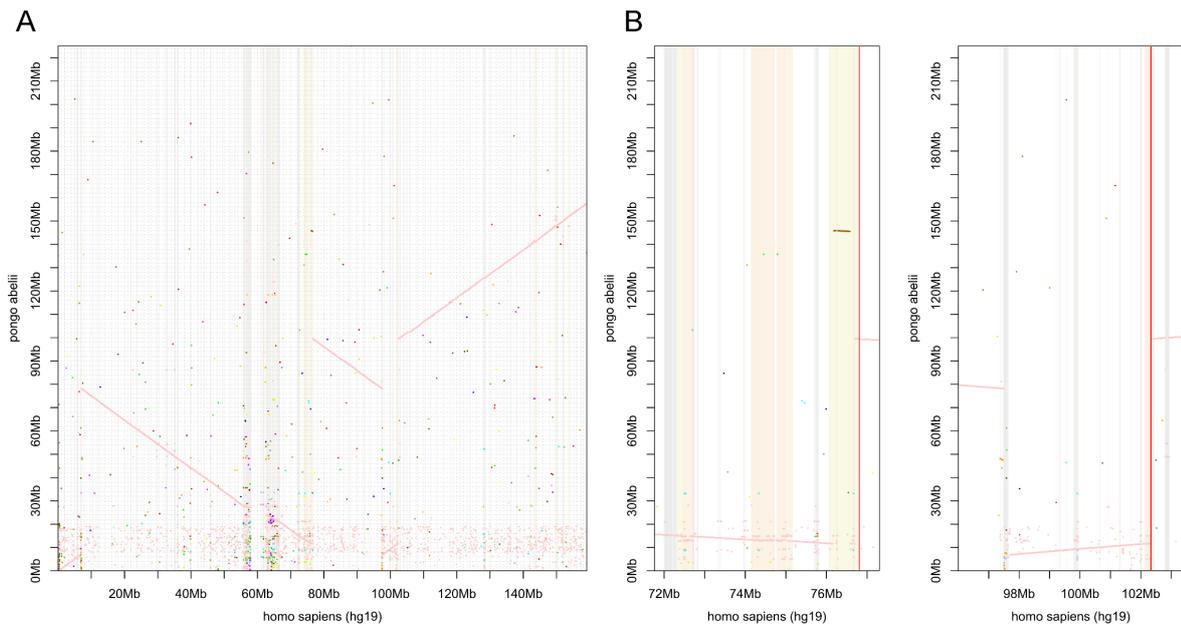


Abbildung S4.: Bruchpunkte der perizentrischen und parazentrischen Inversion des humanen Chromosom 7 während der Hominoidenevolution.

A Dot Plot des Alignments des Chromosom 7 des humanen (hg19) und des Orang-Utan Genoms (ponAbe2). Von links nach rechts aufsteigende Linien zeigen chromosomale Segmente mit direkter Homologie an, während absteigende Linien invertierte Homologien darstellen. Durch die Beteiligung eines Segmentes an beiden Inversionen zeigt das Segment bei ca. 100 Mb des Genoms von *Homo sapiens* eine direkte homologe Orientierung an. Bruchpunktregionen der Rearrangements erscheinen als Unterbrechungen der Linien (siehe B). **B** Ausschnitt der proximalen und distalen Bruchpunktregionen der parazentrischen Inversion (als vertikale rote Linien). Regionen mit segmentalen Duplikationen sind mittels farblicher Schattierungen entsprechend der Farbkodierung in Abb. 3.2 hervorgehoben. Die Chromosomen sind gemäß des Farbschemas des UCSC Genome Browsers repräsentiert (Abb. erstellt von Dr. Peter Arndt und aus [99]).

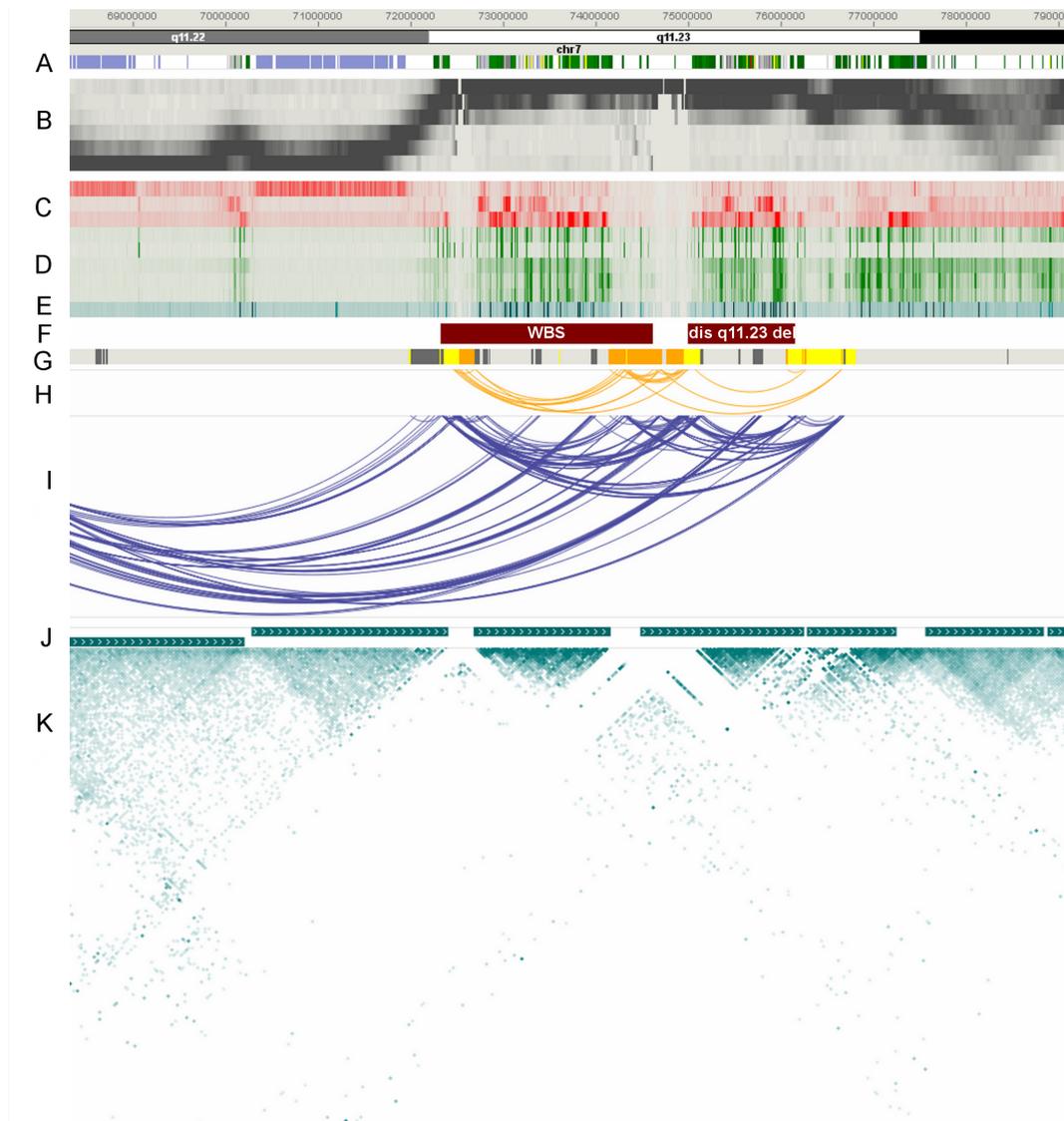


Abbildung S5.: Zusätzliche Daten zur Chromatinorganisation der WBS-Region

Alle Angaben sind in hg19 angegeben und stammen aus der Zelllinie IMR90 (soweit nicht anders beschrieben). Siehe auch Abb. 3.6. **A** Chromatinstatus entsprechend [110] [111]; Heterochromatin (hellblau), aktiv transkribierend (dunkelgrün), Enhancer (gelb), schwache Repression durch Polycomb (grau), starke Repression durch Polycomb (dunkelgrau), bivalenter Transkriptionsstartpunkt (rot) **B** Phasen der Replikationszeit von G_1 (oben) bis G_2 (unten) nach [152] [405] **C** Epigenetische Marker eines reprimierten Chromatinstatus (H3K9me1, H3K27me3, H4K20me1) in rot **D** Epigenetische Marker eines aktiven Chromatinstatus (H3K4me2, H3K4me3, H3K14ac, H3K18ac, H3K4me1, H3K27ac, H3K36me3) in grün **E** CTCF Bindestellen **F** In WBS bzw. dem distalen 7q11.23-Deletionssyndrom häufig deletierte Intervalle **G** Verteilung der SD-Cluster (Farbkodierung nach Abb. 3.2) **H** SD-Paraloge mit einer Sequenzsimilarität von über 99 % **I** Paraloge SDs aller Sequenzsimilaritäten **J** Topologische Domänen nach [92] **K** Zweidimensionale Heatmap der Interaktionen pro Binpaar nach [92] (Abbildung erstellt von G. Ebert und entnommen aus [99]).

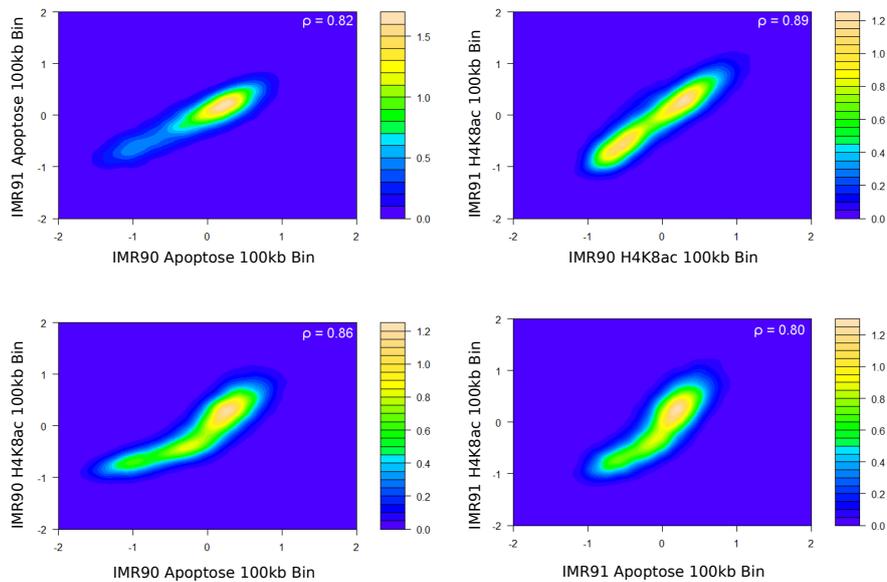


Abbildung S6.: DNA-Degradationsprofile korrelieren mit H4K8-Acetylierungsprofilen innerhalb und zwischen den Zelllinien IMR90 und IMR91L.

Reproduzierbarkeit des Degradationsprofils der apoptotischen DNA und des H4K8-Acetylierungsprofils zwischen den Zelllinien IMR90 und IMR91L (erste Reihe). DNA-Degradationsprofil korreliert mit dem H4K8-Acetylierungsprofil innerhalb der Zelllinien IMR90 bzw. IMR91L (zweite Reihe). Die Korrelationsprofile wurden über die zweidimensionale Kerndichteschätzung erstellt (Abb. entnommen aus [99]).

Tabelle S3.: Anteil der DNA-Kopienzahlveränderungen am Gesamtgenom von Tumorzellproben der 20 Sézary-Patienten.

Die Abdeckung des Genoms durch DNA-Kopienzahlunterschiede wie Deletionen oder Duplikationen (Del: Deletionen, Dup: Duplikationen). Bei der Berechnung wurden das Chromosom Y und nicht-annotierte Bereiche ausgeschlossen.

Patient Nr.	Anteil an						Verhältnis Del/Dup
	Duplikationen		Deletionen		Aberrationen gesamt		
	in Mb	in %	in Mb	in %	in Mb	in %	
1	109,2	0,03	95,3	0,03	204,5	0,07	0,87
2	178,2	0,06	78,0	0,02	256,2	0,09	0,43
3	206,2	0,07	332,6	0,11	538,8	0,19	1,61
4	74,9	0,02	113,9	0,04	188,8	0,06	1,52
5	480,0	0,16	237,0	0,08	717,1	0,25	0,49
6	87,9	0,03	130,6	0,04	218,5	0,07	1,48
7	190,4	0,06	338,3	0,11	528,8	0,18	1,77
8	223,5	0,07	115,3	0,04	338,9	0,11	0,51
9	216,4	0,07	115,8	0,04	332,2	0,11	0,53
10	174,3	0,06	64,8	0,02	239,2	0,08	0,37
11	562,8	0,19	609,0	0,21	1171,8	0,41	1,08
12	175,4	0,06	57,1	0,02	232,5	0,08	0,32
13	108,4	0,03	186,2	0,06	294,6	0,10	1,71
14	114,3	0,04	166,9	0,05	281,3	0,09	1,45
15	653,8	0,23	57,1	0,02	711,0	0,25	0,08
16	270,3	0,09	349,6	0,12	620,0	0,21	1,29
17	766,1	0,27	546,3	0,19	1312,5	0,46	0,71
18	85,9	0,03	345,2	0,12	431,1	0,15	4,01
19	137,9	0,04	82,9	0,02	220,8	0,07	0,60
20	595,6	0,21	273,7	0,09	869,3	0,30	0,45
Ø		0,09		0,07		0,17	1,06

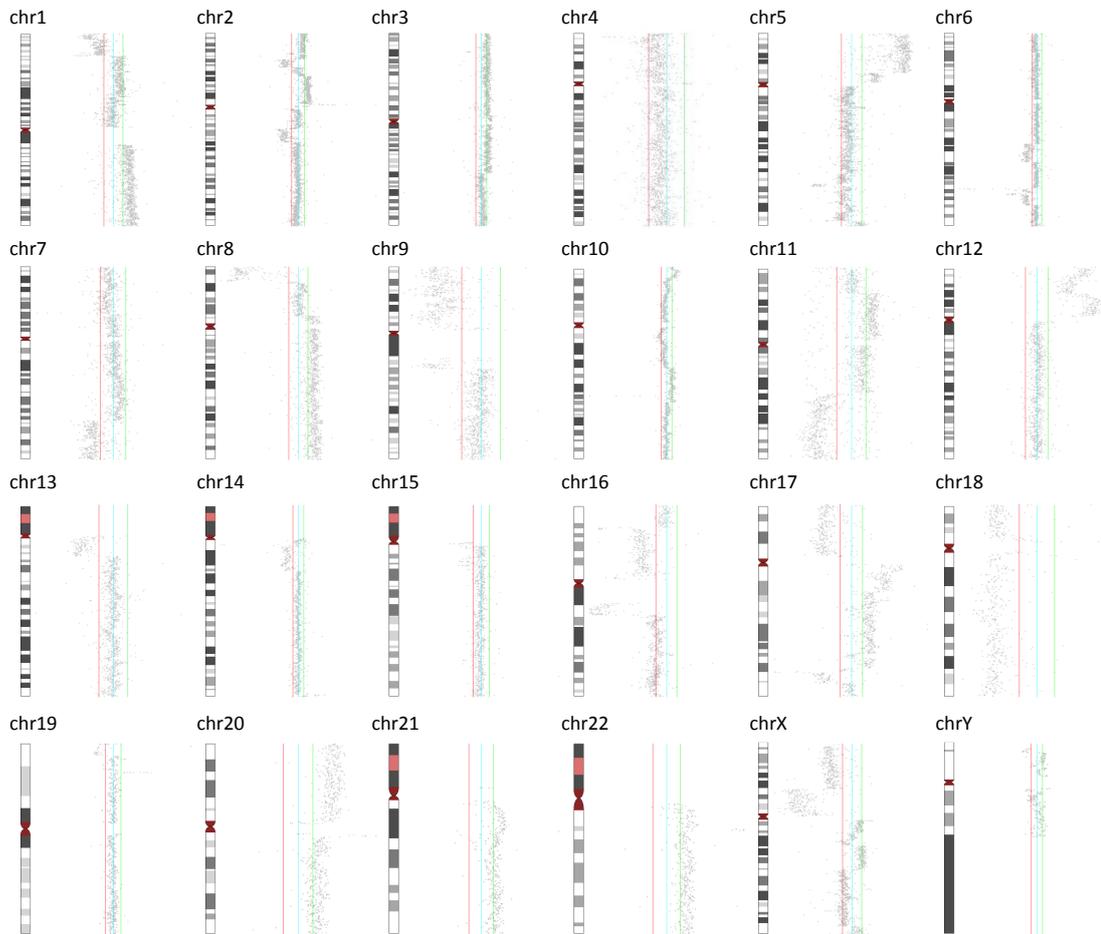


Abbildung S7.: DNA-Kopienzahlveränderungen der Sézary-Zelllinie Seax.

Log₂-Intensitäten der aCGH-Analyse sind neben den Idiogrammen entsprechend ihrer genomischen Position aufgetragen. Der Schwellenwert von -0,2 für die Bestätigung einer Deletion ist als rote vertikale Linie, der Schwellenwert für Duplikationen von 0,2 als grüne Linie markiert. Ein Überschreiten der Linien wird als DNA-Kopienzahlveränderung gewertet.

Tabelle S4.: Bestätigung der Deletion des *E2A*-Locus in Sézary-Tumorzellproben mittels FISH-Analysen.

Vergleich der aCGH- und FISH-Ergebnisse der Sézary-Tumorzellproben und der Zelllinie Seax. Für die FISH-Analyse werden die Ergebnisse als „Anzahl der Zellen mit Deletion/ Anzahl der untersuchten Zellen“ angegeben. Vier Tumorzellproben weisen mit FISH-Analysen Abweichungen von den aCGH-Ergebnissen auf. Die betreffenden Proben wurden für die beiden Analysemethoden zu unterschiedlichen Zeitpunkten entnommen. Tabelle aus [389].

Patient Nr.	<i>E2A</i> -Verlust detektiert mit		angereicherte Tumorzellen in %
	Array-CGH	FISH	
1	ja ^c	ja (149/200) ^e	96.7
2	ja ^c	n. d.	–
3	ja ^c	ja (198/200) ^e	99.1
4 ^a	ja ^c	ja (167/200) ^e	93.2
5	ja ^e	ja (90/200) ^e	97.9
6	ja ^d	n. d.	–
7 ^a	ja ^c	nein (14/200) ^e	99.2 ^f
8	ja ^e	n. d.	–
9	nein ^c	nein (3/200) ^e	97.4
10	nein ^c	nein (4/200) ^e	96.9
11 ^a	ja ^e	nein (12/200) ^d	n. d.
12	nein ^c	n. d.	–
13	ja ^c	ja (126/200) ^e	98.4 ^f
14	nein ^c	n. d.	–
15 ^a	nein ^{b,c}	ja (106/200) ^d	n. d.
16 ^a	nein ^e	ja (Del 28/200, Dup 46/200) ^e	97.2
17	ja ^e	ja (50/200) ^e	98.4
18	ja ^c	n. d.	–
19	ja ^e	n. d.	–
20	ja ^e	n. d.	–
Seax	ja	ja (190/200)	–

n. d.: nicht determiniert

Del: Deletion

Dup: Duplikation

^aTumorzellproben für aCGH- und FISH-Analysen wurden an unterschiedlichen Zeitpunkten entnommen (siehe Material und Methoden)

^bhypotriploide Zellen, Duplikation des Chromosom 19 mit Ausnahme der Region 19p13.2-13.3

^cAnalyse von PBMC Zellen

^dAnalyse von CD4-sortierten Zellen

^eAnalyse von V β -sortierten Zellen

^fder Reinheitsgrad wurde ausschließlich mit CD4-Antikörpern determiniert

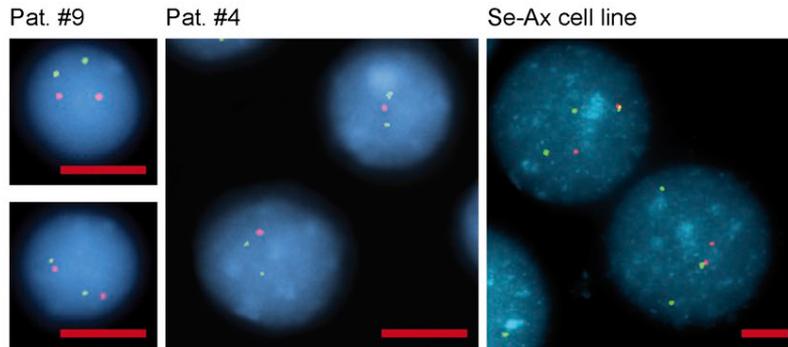


Abbildung S8.: Bestätigung der aCGH-Analyse der Sézary-Tumorzellproben mittels FISH-Analysen.

Beispiele der Zweifarben-FISH-Analyse von Tumorzellproben der Sézary-Patienten mit den Sonden für *E2A* (rot) und der Kontrolle 19qter (grün). Tumorzellen des Patienten 9 weisen jeweils zwei Signale der *E2A*-Sonde und der Kontrollsonden auf. Hier ist *E2A*, in Übereinstimmung mit den aCGH-Ergebnissen, nicht deletiert. Tumorzellen des Patienten 4 zeigen einen heterozygoten Verlust des *E2A*-Locus mit einem *E2A*-Signal und zwei Kontrollsignalen. Ebenfalls lässt sich bei der Sézary-Zelllinie Seax ein heterozygoter Verlust mit zwei *E2A*-Signalen und drei Kontrollsignalen detektieren (Roter Balken: 5 μm). Die Experimente wurden durchgeführt und die Abbildung erstellt von Prof. Dr. E. Schröck [389].

Tabelle S5.: Signifikant deregulierte Gene nach *E2A*-Rekonstitution durch die Konstrukte E47-fd bzw. E2A-myc in der Seax-Zelllinie.

Signifikant deregulierte Gene mit einem assoziierten $\log_2\text{FC}$ von > 1 bzw. < -1 und adjustiertem (adj.) p -Wert von $p < 0,01$. Die Genliste wurde über „Limma“ erstellt.

Genname	Agilent ID	E47-fd vs. Mock		E2A-myc vs. Mock	
		$\log_2\text{FC}$	adj. p -Wert	$\log_2\text{FC}$	adj. p -Wert
<i>SPRR1A</i>	A_23_P74012	10.29	2.937×10^{-06}	11.40	9.009×10^{-07}
<i>LTB</i>	A_23_P93348	7.48	1.401×10^{-06}	7.57	9.009×10^{-07}
<i>ELASB</i>	A_23_P200579	6.82	1.579×10^{-03}	5.73	1.408×10^{-03}
<i>TNNT2</i>	A_24_P927304	6.11	7.383×10^{-05}	5.12	5.225×10^{-05}
<i>ANKRD2</i>	A_23_P104555	6.10	2.354×10^{-05}	7.38	1.987×10^{-06}
<i>TNNT2</i>	A_23_P34700	5.90	1.360×10^{-04}	5.34	7.259×10^{-05}
<i>A_23_P111766</i>	A_23_P111766	5.67	2.640×10^{-03}	2.89	3.567×10^{-02}
<i>KRT5</i>	A_23_P218047	5.59	7.383×10^{-05}	6.40	1.028×10^{-05}
<i>SFTPC</i>	A_23_P95213	5.20	1.428×10^{-04}	6.27	1.648×10^{-05}
<i>CDKN1C</i>	A_23_P428129	5.17	2.354×10^{-05}	6.29	1.987×10^{-06}
<i>MYBPH</i>	A_23_P148737	5.15	1.567×10^{-02}	6.26	2.092×10^{-03}
<i>CDC42BPG</i>	A_23_P64051	5.04	1.069×10^{-04}	6.83	6.213×10^{-06}
<i>ZP1</i>	A_23_P1912	4.89	3.134×10^{-03}	3.50	7.259×10^{-03}
<i>RTN4R</i>	A_23_P132175	4.83	7.375×10^{-05}	7.54	2.623×10^{-06}
<i>JAG2</i>	A_23_P106024	4.79	3.058×10^{-05}	7.34	1.042×10^{-06}
<i>AQP3</i>	A_23_P112481	4.79	3.306×10^{-03}	5.05	8.409×10^{-04}
<i>CD27</i>	A_23_P48088	4.78	2.948×10^{-04}	4.79	8.902×10^{-05}
<i>KRT17</i>	A_23_P96158	4.78	1.845×10^{-02}	6.23	1.684×10^{-03}
<i>KRT15</i>	A_23_P27133	4.60	1.535×10^{-02}	7.22	4.756×10^{-04}
<i>LOC650517</i>	A_24_P887857	4.51	2.827×10^{-02}	5.87	2.774×10^{-03}
<i>ENST00000272035</i>	A_32_P16315	4.48	3.687×10^{-03}	3.63	4.494×10^{-03}
<i>AQP3</i>	A_23_P112482	4.44	7.588×10^{-03}	4.76	1.968×10^{-03}
<i>ENST00000311208</i>	A_24_P610945	4.30	1.956×10^{-02}	5.72	1.593×10^{-03}
<i>AGT</i>	A_23_P115261	4.26	3.638×10^{-03}	3.11	7.712×10^{-03}

Fortsetzung auf der nächsten Seite

Tabelle S5.: Signifikant deregulierte Gene nach *E2A*-Rekonstitution durch die Konstrukte *E47*-fd bzw. *E2A*-myc in der *Seax*-Zelllinie.

Signifikant deregulierte Gene mit einem assoziierten $\log_2\text{FC}$ von > 1 bzw. < -1 und adjustiertem (adj.) p -Wert von $p < 0,01$. Die Genliste wurde über „Limma“ erstellt.

Genname	Agilent ID	E47-fd vs. Mock		E2A-myc vs. Mock	
		$\log_2\text{FC}$	adj. p -Wert	$\log_2\text{FC}$	adj. p -Wert
<i>KLK12</i>	A_23_P500010	4.20	1.348×10^{-02}	5.00	1.962×10^{-03}
<i>BDKRB2</i>	A_23_P304897	4.17	2.401×10^{-03}	6.86	4.628×10^{-05}
<i>HIVEP3</i>	A_32_P59302	4.13	1.566×10^{-03}	4.28	3.946×10^{-04}
<i>KRT14</i>	A_24_P265346	4.11	2.922×10^{-02}	5.95	1.566×10^{-03}
<i>FLJ10324</i>	A_23_P111860	4.11	4.903×10^{-04}	6.02	1.785×10^{-05}
<i>CAND2</i>	A_23_P250102	4.11	1.257×10^{-04}	2.90	2.657×10^{-04}
<i>TNNT2</i>	A_24_P257022	4.08	1.001×10^{-03}	4.19	2.763×10^{-04}
<i>SYT12</i>	A_23_P421306	4.06	3.058×10^{-05}	4.26	6.135×10^{-06}
<i>ITGA6</i>	A_23_P210176	4.00	1.672×10^{-03}	2.34	1.249×10^{-02}
<i>BC014971</i>	A_23_P140884	4.00	8.985×10^{-03}	3.33	9.929×10^{-03}
<i>AQP1</i>	A_23_P372834	3.97	2.109×10^{-04}	3.37	1.710×10^{-04}
<i>MAL</i>	A_23_P17134	3.95	9.787×10^{-04}	3.69	4.744×10^{-04}
<i>CABP7</i>	A_24_P177236	3.83	4.792×10^{-03}	5.78	1.677×10^{-04}
<i>GOLGA8E</i>	A_32_P103669	3.72	1.227×10^{-02}	5.40	5.846×10^{-04}
<i>THC2647902</i>	A_32_P73507	3.71	2.354×10^{-05}	3.33	7.885×10^{-06}
<i>AK026750</i>	A_24_P589028	3.69	7.155×10^{-03}	5.37	3.106×10^{-04}
<i>KLK1</i>	A_23_P16252	3.68	1.840×10^{-03}	4.53	1.925×10^{-04}
<i>ADAM11</i>	A_23_P207345	3.57	1.718×10^{-04}	4.43	1.648×10^{-05}
<i>BAIAP3</i>	A_23_P163492	3.56	1.672×10^{-03}	6.38	2.039×10^{-05}
<i>TPSAB1</i>	A_23_P37702	3.56	1.672×10^{-03}	6.37	2.039×10^{-05}
<i>PDE6B</i>	A_23_P22143	3.51	3.602×10^{-03}	7.38	1.800×10^{-05}
<i>THC2632909</i>	A_32_P38093	3.46	8.782×10^{-04}	3.47	2.701×10^{-04}
<i>DYSF</i>	A_23_P39925	3.40	3.114×10^{-04}	4.68	1.716×10^{-05}
<i>ATOH8</i>	A_23_P142878	3.39	1.585×10^{-03}	3.02	1.026×10^{-03}
<i>CHDH</i>	A_23_P69293	3.37	4.772×10^{-05}	3.15	1.648×10^{-05}
<i>RGR</i>	A_23_P306941	3.37	2.992×10^{-02}	6.31	3.561×10^{-04}
<i>KRT86</i>	A_23_P363769	3.34	1.749×10^{-02}	4.14	2.130×10^{-03}
<i>DYSF</i>	A_23_P39931	3.34	1.257×10^{-04}	4.72	6.213×10^{-06}
<i>ENTPD8</i>	A_23_P44335	3.31	4.656×10^{-05}	4.03	4.040×10^{-06}
<i>KRT72</i>	A_32_P48825	3.31	9.434×10^{-03}	3.12	5.251×10^{-03}
<i>GPR20</i>	A_23_P159237	3.31	1.749×10^{-02}	4.77	8.960×10^{-04}
<i>PLCH2</i>	A_24_P62469	3.28	4.362×10^{-04}	5.70	7.254×10^{-06}
<i>TMEM80</i>	A_23_P332960	3.24	2.247×10^{-04}	4.44	1.390×10^{-05}
<i>WNT10A</i>	A_23_P102113	3.19	7.310×10^{-05}	2.70	4.360×10^{-05}
<i>TIMP2</i>	A_23_P107401	3.11	8.521×10^{-04}	1.84	5.294×10^{-03}
<i>WNT10A</i>	A_23_P102117	3.03	1.718×10^{-04}	2.45	1.848×10^{-04}
<i>FCGBP</i>	A_23_P21495	3.03	1.573×10^{-04}	5.74	2.465×10^{-06}
<i>KHK</i>	A_23_P5845	3.01	7.383×10^{-05}	3.55	9.096×10^{-06}
<i>TGM2</i>	A_32_P86763	2.98	3.209×10^{-03}	4.51	1.018×10^{-04}
<i>FXYD2</i>	A_23_P161769	2.97	8.712×10^{-04}	5.11	1.392×10^{-05}
<i>ACTG2</i>	A_23_P39955	2.97	2.741×10^{-03}	2.34	3.574×10^{-03}
<i>LOC124976</i>	A_24_P8371	2.93	9.571×10^{-03}	3.94	7.112×10^{-04}
<i>ARHGAP9</i>	A_23_P64661	2.92	1.967×10^{-03}	3.40	2.757×10^{-04}
<i>ADA</i>	A_23_P210482	2.92	4.656×10^{-05}	2.79	1.427×10^{-05}
<i>ALDH1L1</i>	A_23_P258887	2.88	2.044×10^{-03}	4.04	1.027×10^{-04}
<i>MYO7A</i>	A_23_P98786	2.82	3.021×10^{-02}	5.02	4.965×10^{-04}
<i>TMC8</i>	A_23_P346093	2.81	1.236×10^{-03}	3.65	8.902×10^{-05}
<i>CRIP2</i>	A_23_P112798	2.81	7.310×10^{-05}	5.85	9.009×10^{-07}
<i>KCNQ1</i>	A_23_P429977	2.79	7.280×10^{-05}	5.32	1.042×10^{-06}
<i>NROB2</i>	A_23_P160800	2.77	1.682×10^{-02}	1.89	4.931×10^{-02}
<i>VWF</i>	A_23_P105562	2.74	4.912×10^{-03}	4.77	7.352×10^{-05}
<i>EPS8L2</i>	A_23_P147822	2.70	6.756×10^{-03}	3.36	7.418×10^{-04}
<i>PTGDS</i>	A_23_P146554	2.70	3.736×10^{-03}	3.30	4.332×10^{-04}
<i>PREX1</i>	A_23_P413641	2.69	8.501×10^{-04}	3.34	7.352×10^{-05}
<i>TGM2</i>	A_24_P923251	2.69	3.138×10^{-02}	3.92	1.589×10^{-03}
<i>LAMB2</i>	A_23_P21382	2.62	9.258×10^{-04}	6.42	3.123×10^{-06}
<i>C20orf58</i>	A_23_P17316	2.62	1.001×10^{-03}	6.30	3.690×10^{-06}
<i>GFAP</i>	A_24_P31165	2.60	5.457×10^{-03}	2.36	3.473×10^{-03}
<i>TMEM130</i>	A_23_P349966	2.60	7.575×10^{-03}	3.23	8.464×10^{-04}
<i>ENST00000368804</i>	A_32_P98072	2.59	4.188×10^{-02}	2.63	1.729×10^{-02}
<i>SH3TC1</i>	A_23_P41390	2.59	9.606×10^{-04}	3.85	3.205×10^{-05}
<i>SPRR1A</i>	A_23_P348208	2.59	1.355×10^{-02}	3.59	8.370×10^{-04}

Fortsetzung auf der nächsten Seite

Tabelle S5.: Signifikant deregulierte Gene nach *E2A*-Rekonstitution durch die Konstrukte *E47*-fd bzw. *E2A*-myc in der *Seax*-Zelllinie.

Signifikant deregulierte Gene mit einem assoziierten \log_2FC von > 1 bzw. < -1 und adjustiertem (adj.) p -Wert von $p < 0,01$. Die Genliste wurde über „Limma“ erstellt.

Genname	Agilent ID	E47-fd vs. Mock		E2A-myc vs. Mock	
		\log_2FC	adj. p -Wert	\log_2FC	adj. p -Wert
<i>AL109708</i>	A_24_P823096	2.57	7.064×10^{-03}	3.50	4.663×10^{-04}
<i>COL5A2</i>	A_23_P33196	2.56	1.116×10^{-03}	2.11	1.084×10^{-03}
<i>WNT4</i>	A_23_P11787	2.55	7.958×10^{-04}	5.06	6.213×10^{-06}
<i>CHDH</i>	A_23_P357185	2.54	1.257×10^{-04}	2.23	7.352×10^{-05}
<i>NLRP5</i>	A_23_P4962	2.51	3.602×10^{-03}	3.73	1.350×10^{-04}
<i>MLPH</i>	A_23_P165778	2.51	2.391×10^{-02}	1.88	4.415×10^{-02}
<i>SARDH</i>	A_24_P35400	2.51	4.065×10^{-04}	5.37	3.123×10^{-06}
<i>PTP4A3</i>	A_23_P112774	2.48	2.762×10^{-03}	5.22	1.492×10^{-05}
<i>BC025775</i>	A_32_P491499	2.46	3.053×10^{-02}	5.92	8.772×10^{-05}
<i>PLAU</i>	A_23_P24104	2.45	9.723×10^{-03}	2.45	3.810×10^{-03}
<i>RHBDF1</i>	A_23_P163682	2.44	1.988×10^{-02}	3.78	7.076×10^{-04}
<i>AVIL</i>	A_23_P53257	2.43	3.563×10^{-02}	3.67	1.528×10^{-03}
<i>NTRK1</i>	A_23_P34804	2.42	3.542×10^{-04}	1.93	4.199×10^{-04}
<i>TMEM38A</i>	A_23_P101392	2.41	2.824×10^{-03}	3.47	1.253×10^{-04}
<i>CBFA2T3</i>	A_23_P500741	2.40	1.632×10^{-03}	6.00	3.935×10^{-06}
<i>OXT</i>	A_24_P382579	2.40	2.357×10^{-02}	3.09	2.344×10^{-03}
<i>GLIS1</i>	A_23_P316612	2.39	4.981×10^{-03}	5.39	1.785×10^{-05}
<i>KRT80</i>	A_24_P331704	2.37	5.921×10^{-03}	3.63	1.882×10^{-04}
<i>BC030813</i>	A_24_P357847	2.36	1.718×10^{-04}	1.82	2.435×10^{-04}
<i>RASSF4</i>	A_24_P295590	2.36	1.257×10^{-04}	2.35	3.878×10^{-05}
<i>STAB1</i>	A_23_P32500	2.33	4.445×10^{-03}	5.30	1.636×10^{-05}
<i>KCNIP3</i>	A_23_P356004	2.30	1.170×10^{-04}	3.62	3.725×10^{-06}
<i>BC030813</i>	A_32_P39440	2.29	1.170×10^{-04}	1.69	1.728×10^{-04}
<i>C19orf36</i>	A_23_P50674	2.29	7.028×10^{-03}	3.46	2.495×10^{-04}
<i>RIMS3</i>	A_23_P319583	2.29	4.020×10^{-03}	3.08	2.649×10^{-04}
<i>ENST00000295339</i>	A_32_P159192	2.28	1.267×10^{-04}	1.68	2.186×10^{-04}
<i>LOC651758</i>	A_32_P35947	2.26	1.362×10^{-03}	5.47	3.935×10^{-06}
<i>ENST00000331696</i>	A_32_P148118	2.24	2.948×10^{-04}	1.82	2.984×10^{-04}
<i>HBEGF</i>	A_24_P140608	2.24	6.380×10^{-03}	1.89	6.325×10^{-03}
<i>SYT13</i>	A_23_P161659	2.21	8.858×10^{-04}	2.39	1.848×10^{-04}
<i>IGSF3</i>	A_24_P182461	2.19	3.558×10^{-02}	2.17	1.635×10^{-02}
<i>BC032451</i>	A_23_P21800	2.18	6.694×10^{-04}	1.66	1.004×10^{-03}
<i>AX721087</i>	A_23_P33511	2.18	1.931×10^{-02}	4.29	1.683×10^{-04}
<i>NTSR1</i>	A_23_P371039	2.17	1.298×10^{-02}	5.02	4.287×10^{-05}
<i>FBLN2</i>	A_23_P143981	2.17	1.409×10^{-02}	4.84	5.447×10^{-05}
<i>AF131813</i>	A_24_P623768	2.16	1.535×10^{-02}	1.53	3.794×10^{-02}
<i>NTRK1</i>	A_24_P265506	2.15	2.108×10^{-04}	1.50	5.376×10^{-04}
<i>CAMK2B</i>	A_23_P42882	2.15	1.785×10^{-03}	5.46	3.980×10^{-06}
<i>DTX1</i>	A_24_P290751	2.15	7.064×10^{-03}	3.06	3.543×10^{-04}
<i>THBS1</i>	A_23_P206212	2.14	1.967×10^{-03}	1.36	9.294×10^{-03}
<i>RASSF4</i>	A_23_P75260	2.13	5.712×10^{-04}	2.29	1.251×10^{-04}
<i>FLJ22671</i>	A_23_P60990	2.13	9.301×10^{-03}	4.67	4.117×10^{-05}
<i>KCNH6</i>	A_23_P390984	2.12	1.162×10^{-03}	3.55	1.946×10^{-05}
<i>THC2500892</i>	A_24_P924631	2.12	2.031×10^{-03}	1.86	1.504×10^{-03}
<i>DMBT1</i>	A_23_P86599	2.11	1.298×10^{-02}	4.63	5.457×10^{-05}
<i>A_24_P490109</i>	A_24_P490109	2.11	3.500×10^{-04}	1.57	6.257×10^{-04}
<i>SLC9A3R2</i>	A_24_P364838	2.11	5.712×10^{-04}	3.54	1.053×10^{-05}
<i>PGA3</i>	A_23_P150547	2.11	3.568×10^{-03}	1.08	4.560×10^{-02}
<i>FUT7</i>	A_23_P170534	2.09	6.842×10^{-03}	2.69	6.301×10^{-04}
<i>CHIT1</i>	A_23_P126278	2.09	1.579×10^{-03}	1.86	1.004×10^{-03}
<i>SERINC2</i>	A_23_P147326	2.08	4.908×10^{-02}	2.10	2.047×10^{-02}
<i>AY062331</i>	A_24_P639701	2.07	3.114×10^{-04}	1.59	4.607×10^{-04}
<i>IGKC</i>	A_24_P272146	2.07	1.808×10^{-04}	1.54	3.080×10^{-04}
<i>MYO5B</i>	A_23_P101193	2.06	9.434×10^{-03}	1.83	7.356×10^{-03}
<i>FXYD2</i>	A_24_P196562	2.05	1.579×10^{-03}	3.91	1.427×10^{-05}
<i>COL5A2</i>	A_23_P10391	2.03	1.566×10^{-03}	1.29	6.439×10^{-03}
<i>MGLL</i>	A_24_P226008	2.01	4.981×10^{-03}	2.41	6.708×10^{-04}
<i>LOC388610</i>	A_24_P693986	2.00	2.357×10^{-02}	2.18	6.087×10^{-03}
<i>BCL2L11</i>	A_24_P122921	2.00	3.632×10^{-04}	2.35	4.610×10^{-05}
<i>THBS1</i>	A_24_P142118	1.99	2.475×10^{-04}	1.37	6.903×10^{-04}
<i>PTPRE</i>	A_24_P213494	1.99	1.068×10^{-02}	3.18	2.793×10^{-04}
<i>LOC339240</i>	A_32_P46238	1.98	3.687×10^{-03}	3.49	4.933×10^{-05}

Fortsetzung auf der nächsten Seite

Tabelle S5.: Signifikant deregulierte Gene nach *E2A*-Rekonstitution durch die Konstrukte *E47*-fd bzw. *E2A*-myc in der *Seax*-Zelllinie.

Signifikant deregulierte Gene mit einem assoziierten $\log_2\text{FC}$ von > 1 bzw. < -1 und adjustiertem (adj.) p -Wert von $p < 0,01$. Die Genliste wurde über „Limma“ erstellt.

Genname	Agilent ID	E47-fd vs. Mock		E2A-myc vs. Mock	
		$\log_2\text{FC}$	adj. p -Wert	$\log_2\text{FC}$	adj. p -Wert
<i>C19orf36</i>	A_24_P32295	1.98	3.687×10^{-03}	3.51	4.825×10^{-05}
<i>SCGB1D1</i>	A_23_P127781	1.98	7.998×10^{-03}	2.72	5.193×10^{-04}
<i>RASA4</i>	A_24_P943263	1.97	1.068×10^{-02}	3.30	2.191×10^{-04}
<i>PRKCZ</i>	A_23_P51187	1.97	4.051×10^{-02}	3.62	5.846×10^{-04}
<i>MGC29506</i>	A_23_P84596	1.97	1.573×10^{-03}	1.88	6.707×10^{-04}
<i>PACSIN1</i>	A_23_P258088	1.96	2.254×10^{-03}	1.39	5.466×10^{-03}
<i>AF088004</i>	A_24_P880000	1.96	2.948×10^{-04}	1.51	3.985×10^{-04}
<i>SH3PXD2A</i>	A_23_P35456	1.95	2.934×10^{-02}	2.28	5.233×10^{-03}
<i>AATK</i>	A_23_P10559	1.93	2.745×10^{-03}	3.23	4.755×10^{-05}
<i>NANOS1</i>	A_24_P383609	1.91	1.535×10^{-02}	2.71	8.509×10^{-04}
<i>SOHLH1</i>	A_24_P307964	1.90	3.614×10^{-02}	7.36	8.496×10^{-06}
<i>PRODH</i>	A_23_P68786	1.90	6.380×10^{-03}	3.92	3.885×10^{-05}
<i>MYEOV</i>	A_23_P360240	1.90	7.367×10^{-03}	2.46	6.560×10^{-04}
<i>A_24_P571824</i>	A_24_P571824	1.90	2.651×10^{-02}	3.19	5.969×10^{-04}
<i>GZMH</i>	A_23_P128993	1.89	4.312×10^{-04}	1.69	2.572×10^{-04}
<i>DKFZp434K1815</i>	A_23_P360874	1.89	6.126×10^{-03}	2.71	2.828×10^{-04}
<i>UNC93A</i>	A_23_P214408	1.88	7.718×10^{-04}	4.06	3.935×10^{-06}
<i>ZBTB47</i>	A_24_P71700	1.87	2.670×10^{-02}	3.97	1.575×10^{-04}
<i>ZBP1</i>	A_24_P395966	1.86	5.372×10^{-04}	2.14	7.365×10^{-05}
<i>PTPRE</i>	A_23_P138495	1.86	5.335×10^{-04}	3.00	1.236×10^{-05}
<i>SLC12A7</i>	A_23_P61688	1.85	8.654×10^{-04}	2.91	2.039×10^{-05}
<i>GNAZ</i>	A_23_P416581	1.85	1.672×10^{-03}	4.79	3.725×10^{-06}
<i>C14orf78</i>	A_23_P311912	1.85	1.687×10^{-02}	4.67	3.489×10^{-05}
<i>ENST00000377221</i>	A_24_P66578	1.84	1.669×10^{-03}	1.45	2.230×10^{-03}
<i>BU587941</i>	A_32_P201292	1.84	5.457×10^{-03}	1.77	2.503×10^{-03}
<i>UNC93A</i>	A_24_P103434	1.84	7.028×10^{-03}	4.08	2.728×10^{-05}
<i>ZBP1</i>	A_23_P259141	1.83	8.579×10^{-04}	1.84	2.592×10^{-04}
<i>CUTL2</i>	A_23_P2543	1.83	3.670×10^{-02}	2.61	2.230×10^{-03}
<i>CCDC120</i>	A_23_P73540	1.83	5.921×10^{-03}	1.74	2.896×10^{-03}
<i>TMEM80</i>	A_24_P90349	1.82	3.785×10^{-02}	3.03	9.537×10^{-04}
<i>HIVEP3</i>	A_23_P35330	1.82	3.414×10^{-03}	1.90	9.064×10^{-04}
<i>TCF1</i>	A_24_P372913	1.82	1.669×10^{-03}	2.45	1.012×10^{-04}
<i>SNTB1</i>	A_23_P95029	1.80	3.602×10^{-03}	1.18	1.230×10^{-02}
<i>RGS16</i>	A_23_P217845	1.78	3.626×10^{-02}	2.03	7.690×10^{-03}
<i>ENST00000292728</i>	A_24_P83758	1.78	6.039×10^{-03}	3.48	4.610×10^{-05}
<i>DKFZp434K191</i>	A_24_P687305	1.76	4.907×10^{-02}	2.13	7.585×10^{-03}
<i>SERINC2</i>	A_23_P147331	1.76	1.227×10^{-02}	1.70	5.741×10^{-03}
<i>MYO7B</i>	A_23_P209799	1.74	1.729×10^{-02}	2.64	6.715×10^{-04}
<i>TTC9</i>	A_23_P14508	1.74	5.475×10^{-03}	1.75	1.923×10^{-03}
<i>C4BPB</i>	A_23_P319598	1.73	3.660×10^{-02}	3.12	5.846×10^{-04}
<i>A_23_P72330</i>	A_23_P72330	1.73	9.723×10^{-03}	1.72	3.877×10^{-03}
<i>ID1</i>	A_23_P252306	1.70	4.792×10^{-03}	1.04	2.637×10^{-02}
<i>RGS16</i>	A_23_P320578	1.70	6.722×10^{-03}	1.89	1.395×10^{-03}
<i>BC073815</i>	A_23_P341860	1.70	3.703×10^{-03}	3.70	1.648×10^{-05}
<i>NPYR</i>	A_23_P29282	1.69	2.826×10^{-03}	2.43	1.293×10^{-04}
<i>GIMAP8</i>	A_24_P132383	1.68	2.824×10^{-03}	1.76	7.607×10^{-04}
<i>BM461836</i>	A_32_P42236	1.64	5.457×10^{-03}	3.27	3.885×10^{-05}
<i>KCNN3</i>	A_23_P126167	1.63	9.787×10^{-04}	2.52	2.675×10^{-05}
<i>KRT83</i>	A_24_P410408	1.63	8.412×10^{-03}	2.08	8.409×10^{-04}
<i>RPIP8</i>	A_24_P91094	1.63	3.602×10^{-03}	2.96	4.117×10^{-05}
<i>GIMAP8</i>	A_23_P168388	1.60	1.535×10^{-02}	1.95	2.041×10^{-03}
<i>SLC12A7</i>	A_23_P251855	1.60	4.803×10^{-03}	2.75	7.511×10^{-05}
<i>FGFR3</i>	A_23_P500501	1.59	4.231×10^{-03}	2.76	6.441×10^{-05}
<i>KLF15</i>	A_24_P20327	1.59	8.209×10^{-03}	1.44	5.475×10^{-03}
<i>MOV10L1</i>	A_23_P143774	1.58	2.302×10^{-02}	1.35	2.310×10^{-02}
<i>MGC9712</i>	A_23_P157247	1.56	2.787×10^{-03}	2.92	2.628×10^{-05}
<i>BC032451</i>	A_24_P677559	1.56	1.573×10^{-03}	1.02	5.597×10^{-03}
<i>PLEKHG3</i>	A_23_P76901	1.55	1.881×10^{-02}	2.08	1.493×10^{-03}
<i>TRIM29</i>	A_23_P203267	1.52	4.187×10^{-02}	2.58	9.378×10^{-04}
<i>PCYT2</i>	A_24_P404245	1.52	3.785×10^{-02}	2.90	4.332×10^{-04}
<i>IQSEC1</i>	A_23_P6746	1.51	1.468×10^{-02}	2.53	2.982×10^{-04}
<i>ADAM15</i>	A_23_P159227	1.51	2.801×10^{-02}	1.82	4.223×10^{-03}

Fortsetzung auf der nächsten Seite

Tabelle S5.: Signifikant deregulierte Gene nach *E2A*-Rekonstitution durch die Konstrukte *E47*-fd bzw. *E2A*-myc in der *Seax*-Zelllinie.

Signifikant deregulierte Gene mit einem assoziierten \log_2FC von > 1 bzw. < -1 und adjustiertem (adj.) p -Wert von $p < 0,01$. Die Genliste wurde über „Limma“ erstellt.

Genname	Agilent ID	E47-fd vs. Mock		E2A-myc vs. Mock	
		\log_2FC	adj. p -Wert	\log_2FC	adj. p -Wert
<i>AKR1B10</i>	A_23_P93641	1.51	1.937×10^{-02}	1.62	5.430×10^{-03}
<i>BDKRB1</i>	A_23_P128744	1.51	1.298×10^{-02}	3.92	2.039×10^{-05}
<i>LOC388610</i>	A_24_P734953	1.49	3.558×10^{-02}	1.42	1.997×10^{-02}
<i>CABLES1</i>	A_23_P422851	1.49	1.196×10^{-02}	1.46	5.339×10^{-03}
<i>ADRA2C</i>	A_23_P256158	1.49	4.445×10^{-03}	3.01	2.837×10^{-05}
<i>PTPRE</i>	A_24_P213503	1.48	3.909×10^{-03}	2.69	4.482×10^{-05}
<i>ADAM11</i>	A_23_P502158	1.45	4.904×10^{-02}	2.22	1.955×10^{-03}
<i>AMBP</i>	A_23_P256504	1.44	3.563×10^{-02}	2.24	1.256×10^{-03}
<i>AQP1</i>	A_23_P19894	1.43	1.956×10^{-02}	1.64	3.865×10^{-03}
<i>CPAMD8</i>	A_23_P67198	1.42	2.824×10^{-03}	1.88	1.966×10^{-04}
<i>MYH7B</i>	A_24_P130621	1.42	3.198×10^{-02}	3.96	4.027×10^{-05}
<i>DDR1</i>	A_23_P93311	1.42	2.302×10^{-02}	2.34	5.662×10^{-04}
<i>F2RL3</i>	A_23_P208422	1.41	8.412×10^{-03}	2.44	1.470×10^{-04}
<i>HBEGF</i>	A_23_P213944	1.41	7.919×10^{-03}	1.24	6.382×10^{-03}
<i>BIN1</i>	A_23_P165333	1.38	1.061×10^{-02}	1.25	7.283×10^{-03}
<i>GRIN2C</i>	A_23_P49546	1.36	3.736×10^{-03}	2.96	1.692×10^{-05}
<i>DDR1</i>	A_24_P123601	1.36	3.456×10^{-02}	2.25	8.509×10^{-04}
<i>TBC1D1</i>	A_23_P73023	1.34	6.126×10^{-03}	1.29	2.898×10^{-03}
<i>GNAT1</i>	A_24_P320036	1.33	1.368×10^{-02}	2.41	1.820×10^{-04}
<i>THC2678045</i>	A_32_P174285	1.32	4.234×10^{-02}	3.81	4.369×10^{-05}
<i>ZNF238</i>	A_23_P200512	1.31	7.408×10^{-03}	1.25	3.915×10^{-03}
<i>KIF1A</i>	A_23_P395460	1.31	2.326×10^{-02}	3.99	1.785×10^{-05}
<i>PVRL1</i>	A_24_P50890	1.31	4.246×10^{-03}	3.48	7.238×10^{-06}
<i>DDR1</i>	A_24_P367289	1.27	2.801×10^{-02}	1.99	9.208×10^{-04}
<i>DAB2IP</i>	A_23_P123848	1.24	1.471×10^{-02}	2.17	2.402×10^{-04}
<i>ADAMTSL2</i>	A_23_P20697	1.22	4.263×10^{-02}	3.95	2.379×10^{-05}
<i>AKR1B10</i>	A_24_P129341	1.21	1.714×10^{-02}	1.69	1.039×10^{-03}
<i>BDKRB1</i>	A_24_P919850	1.21	1.834×10^{-02}	3.16	3.081×10^{-05}
<i>RP11-529110.4</i>	A_23_P52286	1.21	9.649×10^{-03}	1.26	2.999×10^{-03}
<i>HTRA3</i>	A_23_P395438	1.20	1.682×10^{-02}	2.52	9.808×10^{-05}
<i>SLC12A7</i>	A_24_P60634	1.18	1.877×10^{-02}	2.32	1.677×10^{-04}
<i>TESC</i>	A_23_P76538	1.17	2.656×10^{-02}	1.94	6.467×10^{-04}
<i>PSD2</i>	A_23_P30358	1.17	2.445×10^{-02}	1.62	1.589×10^{-03}
<i>STMN3</i>	A_23_P165927	1.17	2.517×10^{-02}	3.82	1.392×10^{-05}
<i>RP11-93B14.6</i>	A_23_P384551	1.15	7.028×10^{-03}	2.31	4.700×10^{-05}
<i>TMEM129</i>	A_23_P354175	1.15	3.771×10^{-02}	1.57	2.896×10^{-03}
<i>BSND</i>	A_23_P23029	1.13	6.380×10^{-03}	2.17	5.474×10^{-05}
<i>NINJ1</i>	A_23_P169137	1.12	1.343×10^{-02}	1.52	9.537×10^{-04}
<i>SEMA3F</i>	A_24_P300076	1.10	2.497×10^{-02}	2.37	1.340×10^{-04}
<i>LOC340888</i>	A_24_P152845	1.09	1.596×10^{-02}	1.67	5.974×10^{-04}
<i>OSTbeta</i>	A_23_P436284	1.09	2.967×10^{-02}	2.41	1.404×10^{-04}
<i>ARHGEF16</i>	A_23_P114670	1.09	1.384×10^{-02}	4.18	3.725×10^{-06}
<i>SMPD3</i>	A_23_P163567	1.06	1.400×10^{-02}	1.23	2.450×10^{-03}
<i>AK024399</i>	A_23_P76136	1.04	4.370×10^{-02}	1.10	1.402×10^{-02}
<i>A_32_P162322</i>	A_32_P162322	1.02	2.639×10^{-02}	1.20	4.520×10^{-03}
<i>CSRFP1</i>	A_24_P225448	1.02	1.705×10^{-02}	1.68	3.888×10^{-04}
<i>WNT4</i>	A_23_P382607	1.01	2.280×10^{-02}	3.82	6.213×10^{-06}
<i>PDGFRA</i>	A_23_P300033	-1.05	4.447×10^{-02}	-1.00	2.536×10^{-02}

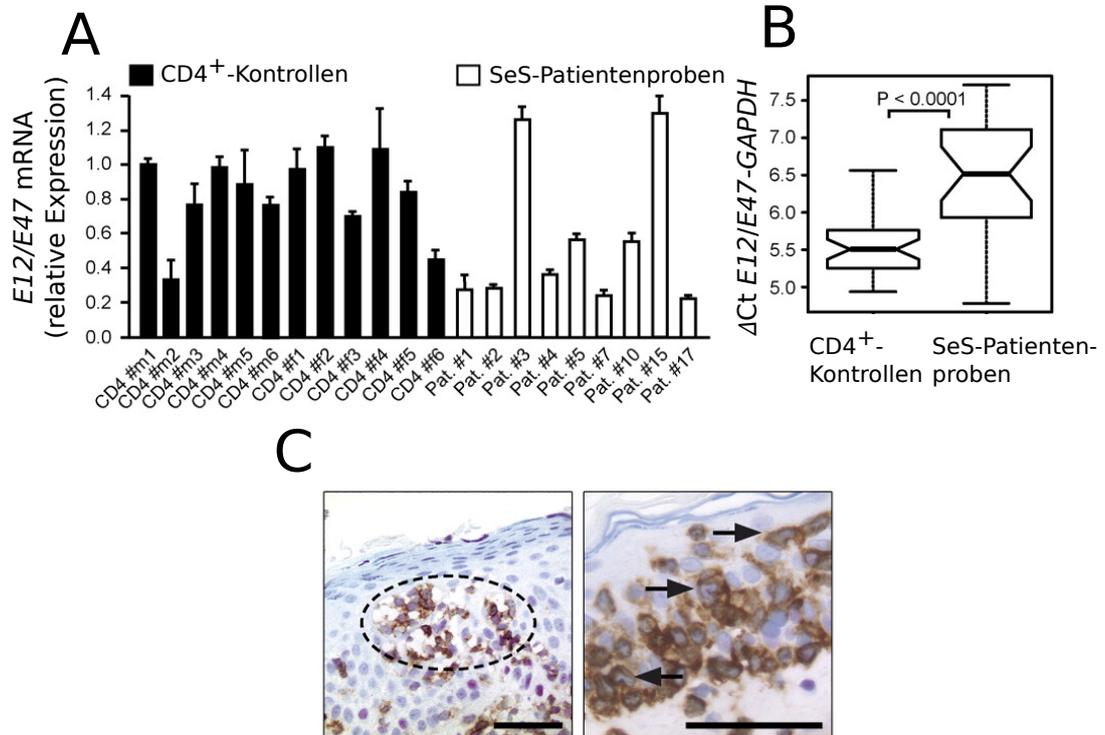


Abbildung S9.: Nachweis des E2A-Verlustes auf RNA- und Proteinebene.

A Relative Quantifizierung der mRNA der *E2A*-Transkriptvarianten *E47* und *E12* mittels RT-qPCR im Vergleich zur Kontrolle *GAPDH*. Tumorzellproben der Sézary-Patienten zeigen im Vergleich zu den CD4⁺-T-Zell-Kontrollen eine verminderte *E12/E47*-Expression (m-männlich, f-weiblich) **B** Boxplot der ΔCt -Werte für die beiden *E2A*-Transkriptvarianten in 12 CD4⁺-T-Zell-Kontrollen im Vergleich zu den Sézary-Patientenproben. Die ΔCt -Werte von *E12/E47* sind über den Vergleich mit *GAPDH* normalisiert worden. Unter Annahme eines stabilen *GAPDH*-mRNA-Niveaus zwischen den Tumor- und Kontrollzellen exprimieren die Patientenprobenzellen im Vergleich zu Kontrollzellen signifikant weniger *E12/E47* relativ zu *GAPDH*. Angegeben sind die Resultate aus einem von zwei unabhängig durchgeführten Experimenten. **C** Beispiele des immunhistochemischen Nachweises des E2A-Verlustes in Hautbiopsien von Sézary-Patienten. Die Zellen wurden einer Doppelfärbung mit dem T-Zellmarker CD3 (braun) und den E2A-Transkriptvarianten E12/E47 (violett) unterzogen. Die gestrichelte Umrandung entspricht einem Pautrier-Mikroabszess, einer Akkumulation von malignen T-Lymphozyten in der Epidermis. Pfeile markieren Tumorzellen mit cerebriformem Nukleus. Der schwarze Balken repräsentiert 50 μm (Experimente durchgeführt und Abbildungen erstellt von Dr. M. Möbs und Dr. S. Mathas, Abbildung aus [389]).

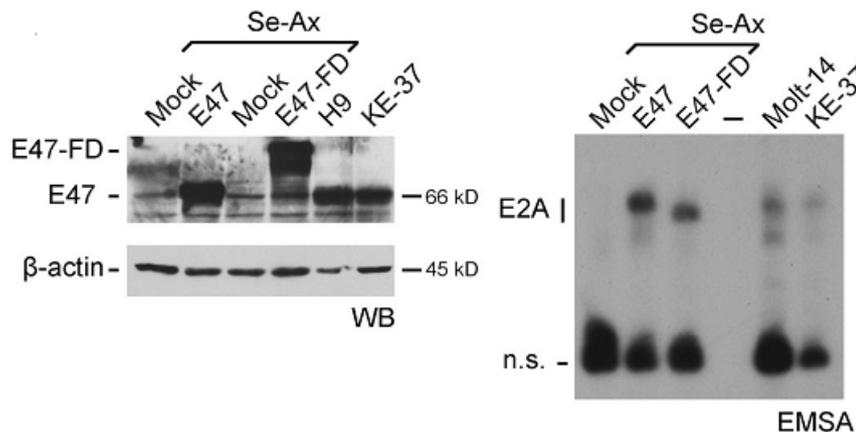


Abbildung S10.: Erhöhte Proteinexpression und E-Box-Bindeaktivität von E2A nach Rekonstitution in Seax-Zellen

Linke Seite: Western-Blot-Analyse zur Bestimmung der E2A-Proteinexpression nach der Transfektion mit den Konstrukten des E47-Homodimers (E47-fd), E2A-myc (E47) und dem Kontrollplasmid (Mock) in der Seax-Zelllinie. Die Ko-Transfektion mit dem Plasmid pEGFP und die anschließende Isolation der pEGFP⁺-Zellen ermöglichte die Anreicherung erfolgreich transfizierter Zellen. Die Immunmarkierung erfolgte mittels des E2A-Antikörpers und des Antikörpers der Ladekontrolle β -Aktin. Als weitere Kontrollen wurden ebenfalls Proteinlysate von untransfizierten, E2A-positiven Zelllinien aufgetragen. Rechte Seite: Analyse der E2A-Bindeaktivität über EMSA. Protein-Extrakte des Zellkerns des Kontrollexperimentes (Mock) und der *E2A*-transfizierten Zellen (E2A-myc und E2A-fd) wurden mit Sequenzen des μ E5-Elementes inkubiert und elektrophoretisch getrennt. Die anschließende Immunmarkierung erfolgte mittels E2A-Antikörper (n. s.: nicht-spezifischer Komplex). Als Vergleich wurden Protein-Kernextrakte der untransfizierten E2A-positiven Zelllinien H9 und KE-37 ebenfalls der Analyse unterzogen. Die Resultate sind nur für eines der vier unabhängigen Experimente gezeigt (Abbildung erstellt von Dr. S. Mathas und entnommen aus [389]).

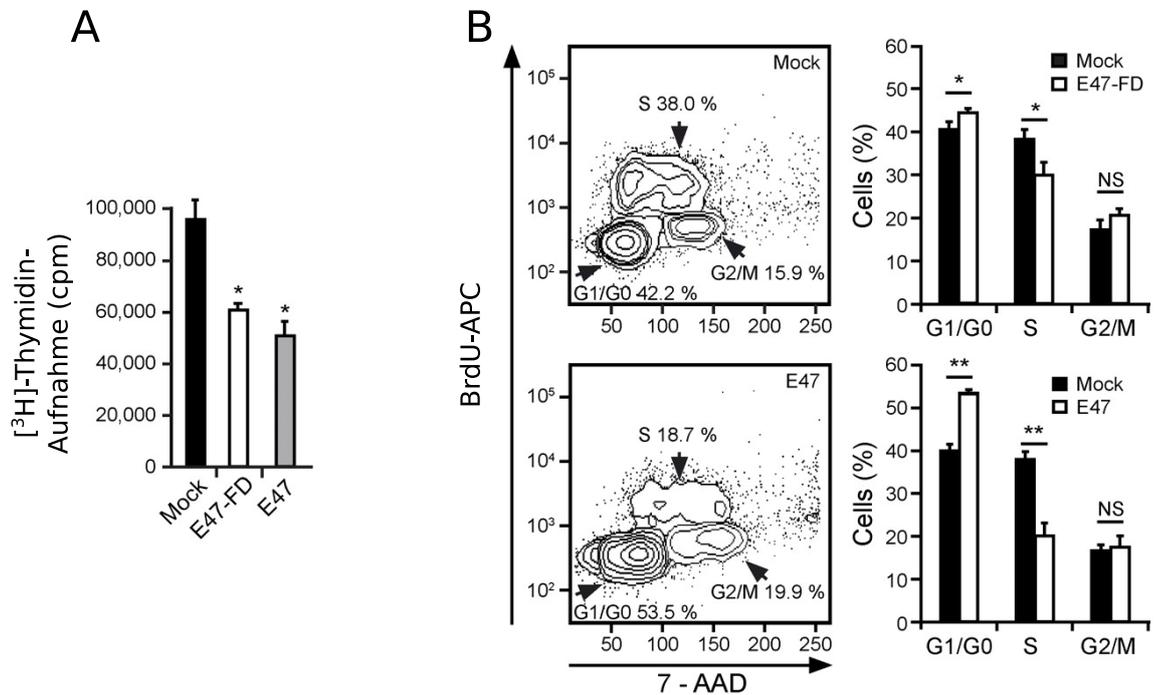


Abbildung S11.: Änderung der Zellproliferation bei Rekonstitution von *E2A* in Seax-Zellen.

A Zellproliferationsassay von Seax-Zellen. Die Transfektion der Seax-Zellen erfolgte für 48 Stunden mit den Konstrukten des E47-Homodimers (E47-fd), E2A-myc (E47) und dem Kontrollplasmid (Mock). Die Ko-Transfektion mit dem Plasmid pEGFP und die anschließende Isolation der pEGFP⁺-Zellen ermöglichte die Anreicherung erfolgreich transfizierter Zellen. Die aufgereinigten Zellen wurden mit 1 μ Ci des Tritium-markierten Nukleosids [³H]-Thymidin für weitere 20 Stunden inkubiert. Während der S-Phase werden diese Nukleoside mit in die neu synthetisierte DNA integriert. Die Zellproliferation kann dadurch über die Intensität des radioaktiven Signals bestimmt werden (counts per minute; **cpm**). Dargestellt sind Resultate aus einem von drei Experimenten (* $p < 0,001$). **B** Zellzyklusanalyse der Seax-Zellen über BrdU-Markierung. Die gemäß Abb. S11 A transfizierten Zellen wurden zusätzlich 30 Minuten mit dem Thymidin-Analogen BrdU inkubiert. Das in die DNA aufgenommene BrdU wurde über Durchflusszytometrie detektiert. Zur Bestimmung der Gesamt-DNA wurden die Zellen zusätzlich mit 7-AAD behandelt. Links: Exemplarische BrdU/7-AAD-Zellzyklusprofile der Zwei-Farben-Durchflusszytometrie (BrdU-APC: BrdU-Antikörper Allophycocyanin-konjugiert). Rechts: Vergleich der Anteile (in %) an DNA-synthetisierenden Zellen, die in den Zellzyklusphasen G₁/G₀, S bzw. G₂/M vorliegen. Dargestellt sind Resultate aus einem von drei Experimenten (**NS**: nicht signifikant, * $p < 0,05$, ** $p < 0,005$; Abbildung erstellt von Dr. S. Mathas und entnommen aus [389]).

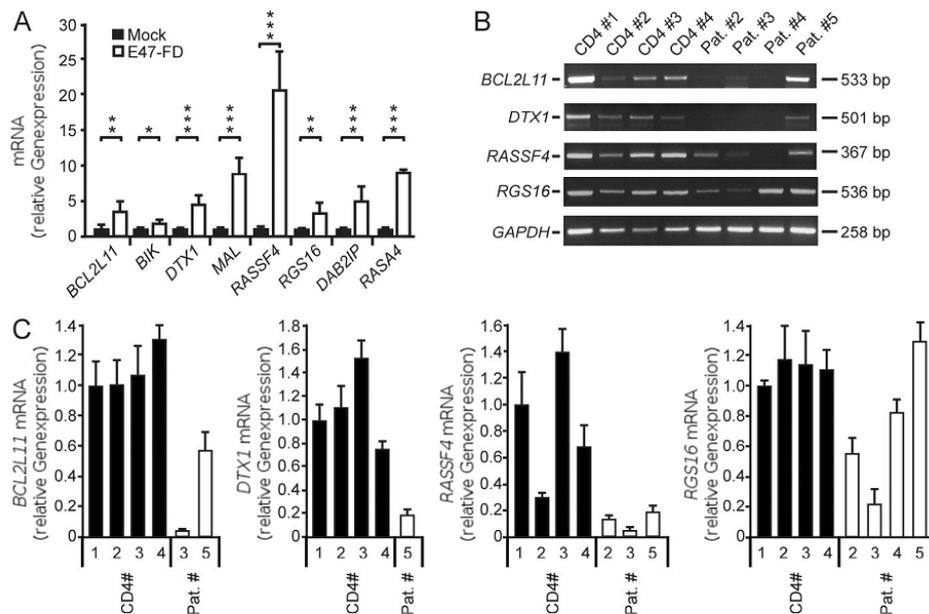


Abbildung S12.: Verifizierung der Genexpressionsarray-Analyse anhand von Beispielgenen.

A RT-qPCR-Analyse der Beispielgene nach transienter Transfektion der Seax-Zellen mit dem *E2A*-Konstrukt E47-FD im Vergleich zu dem Kontrollplasmid. Die Fehlerbalken entsprechen einem 95%-Konfidenzintervall. Die Resultate repräsentieren eines der vier durchgeführten Experimente (*, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$). **B** Semiquantitative PCR der mRNA-Expression *E2A*-deregulierter Gene bzw. der Kontrolle *GAPDH* in CD4-Kontrollzellen (CD4) und Patiententumorproben (Pat). **C** RT-qPCR der Gene mit mRNA-Expression unter B, im Vergleich zwischen CD4-Kontrollzellen und Patiententumorproben. Die Fehlerbalken entsprechen einem 95%-Konfidenzintervall und die Resultate repräsentieren eines der vier durchgeführten Experimente (durchgeführt von Dr. S. Mathas; Abbildung aus [389]).

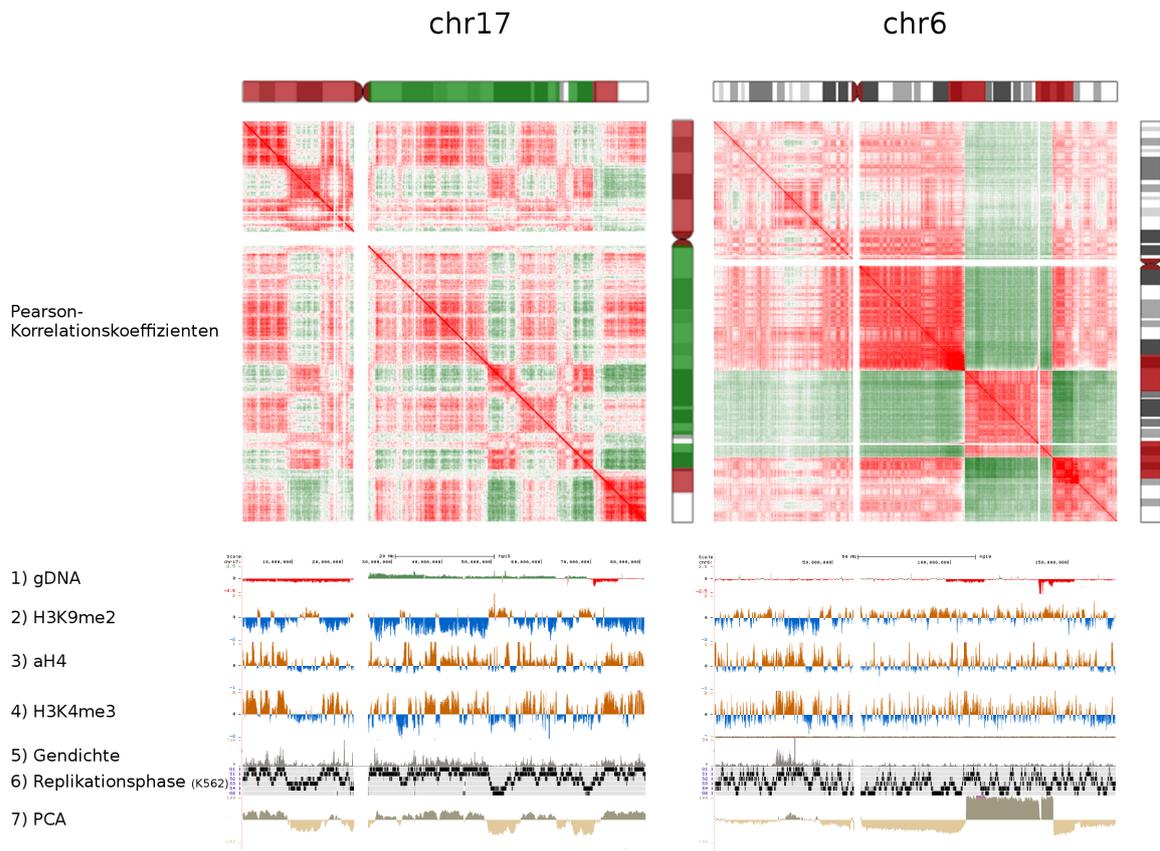


Abbildung S13.: Chromosomenkompartimentierung der Chromosomen 6 und 17.

Vergleich der Matrizen der Pearson-Korrelationskoeffizienten mit verschiedenen Chromatineigenschaften: 1) der DNA-Kopienzahl der aCGH-Analyse (rot: Deletionen, grün: Duplikationen), den Histonmodifikationen von Seax, 2) H3K9me2 (gelb: Hypermethylierung, blau: Hypomethylierung), 3) aH4 (gelb: Hyperacetylierung, blau: Hypoacetylierung), 4) H3K4me3 (gelb: Hypermethylierung, blau: Hypomethylierung), 5) der Gendichte, 6) der Replikationszeiten der Zelllinie K562 und 7) der Hauptkomponentenanalyse (PCA) der Hi-C-Daten der Seax-Zelllinie. PCA des Chromosom 17 repräsentiert aktive A- und inaktive B-Kompartimente (in grau bzw. beige), während die PCA-Analyse des Chromosom 6 chromosomale Rearrangements reflektiert.

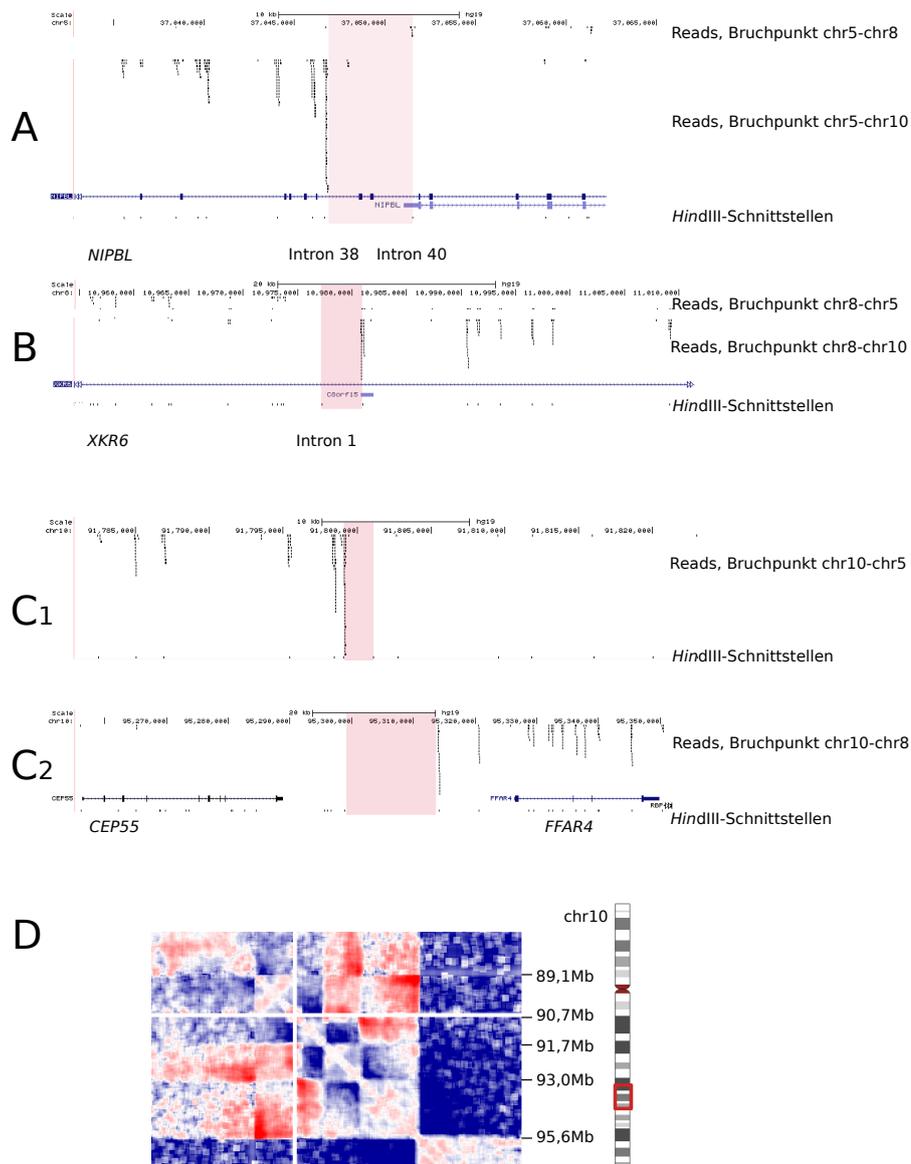


Abbildung S14.: Vergrößerte Ansicht der Bruchpunktregionen der Rearrangements zwischen den Chromosomen chr5, chr8 und chr10.

Eine Änderung der Readpaarhäufigkeiten innerhalb der Rearrangementbruchpunktregionen deutet auf das *HindIII*-Fragment, welches den Bruchpunkt umspannt (rot unterlegt). Gene, in denen dieses Fragment lokalisiert ist, sind möglicherweise direkt an der Translokation beteiligt. **A** Bruchpunktregion des Chromosom 8 innerhalb des Gens *NIPBL*. **B** Bruchpunktregion des Chromosom 8 innerhalb des Gens *XKR6*. **C** Bruchpunktregion des Chromosom 10 zwischen den Genen *CEP55* und *FFAR4*. Die Bruchpunkte zwischen chr10:chr5 und chr10:chr8 liegen hier ca. 3,5 Mb weit voneinander entfernt. **D** Ausschnitt der normalisierten Interaktionsmatrix intrachromosomaler Interaktionen des Chromosom 10 (Bingröße: 250 kb mit der Fenstergröße 50 kb). Zwischen den Bruchpunkten chr10:chr5 und chr10:chr8 treten im Chromosom 10 mehrere intrachromosomale Rearrangements auf.

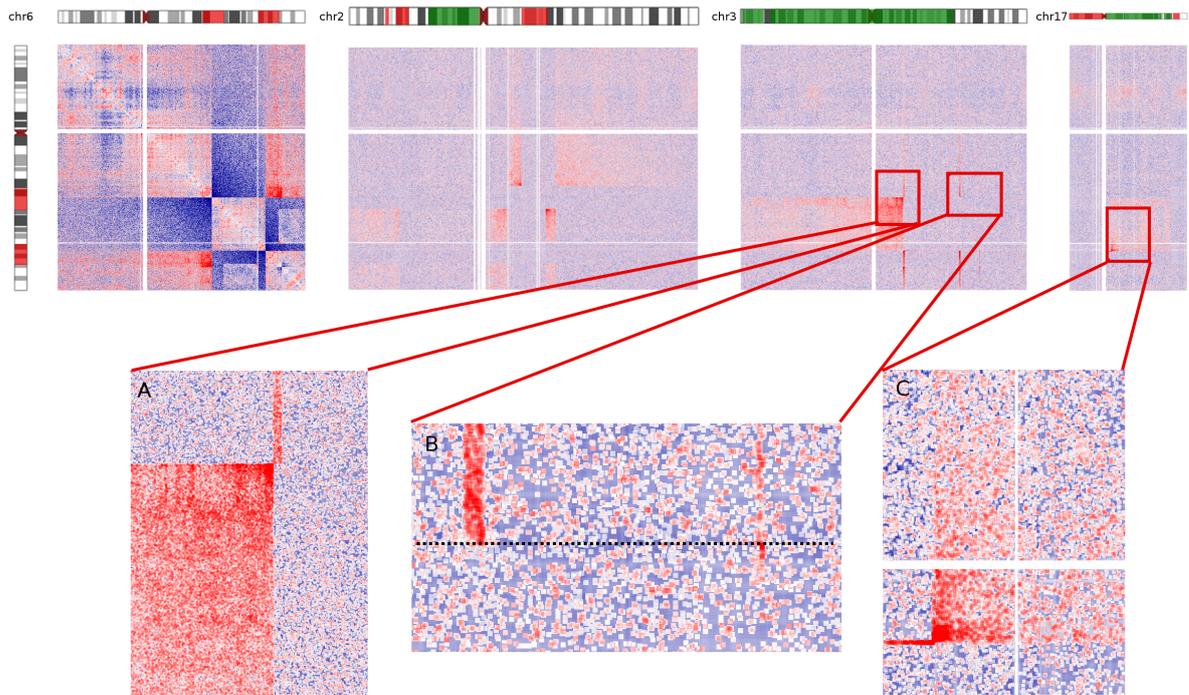


Abbildung S15.: Normalisierte Interaktionsmatrizen der chromosomalen Interaktionen des Chromosom 6 mit den Chromosomen chr2, chr3 und chr17. Die Ausschnitte zeigen Beispiele der Interaktionsmuster, die auf Rearrangements hinweisen. **A** 3q13.2 zeigt erhöhte Interaktionshäufigkeiten mit 6q21 (Rearrangement t7). Die Bruchpunkte sind möglicherweise durch das wenige kb-große 3q26.1-Segment getrennt, interagieren jedoch durch die geringe lineare Distanz zueinander. **B** Der 6q21-Bruchpunkt des translozierten 14 MB 6q-Segmentes (gestrichelte Linie) zeigt, dass die Region 3q25.1 mit der am chr6 verbleibenden Bruchpunktregion assoziiert ist (oberhalb der Linie), während die 3q26.2-Region mit dem translozierten Segment stärker interagiert, daher vermutlich mit diesem fusioniert (unterhalb der Linie). **C** Reziproke Translokation der Regionen 6q24.2 und 17q11.2 (Rearrangement t15). Bingröße der Matrizen: 250 kb mit einer Fenstergröße von 50 kb. Chromosomale Aberrationen sind entsprechend ihrer Lokalisation grün (Duplikationen) oder rot (Deletionen) am Idiogramm markiert.

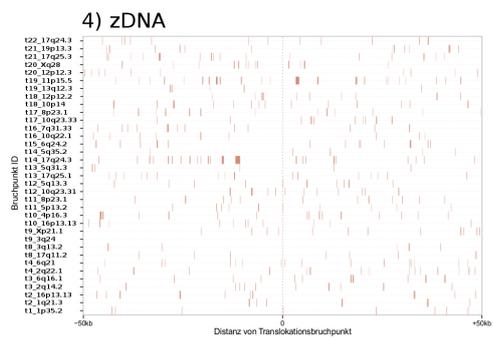
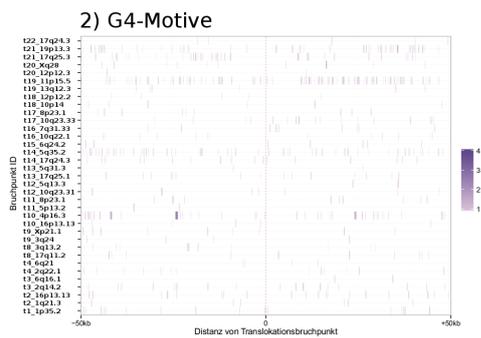
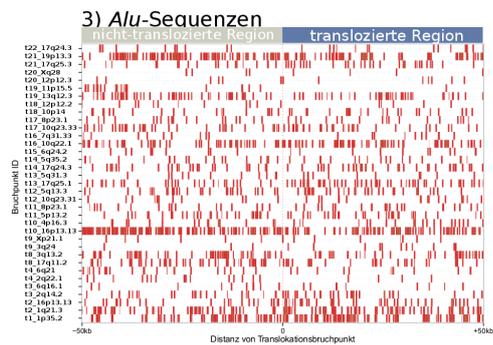
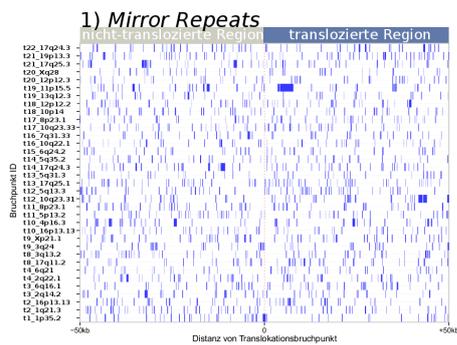


Abbildung S16.: Sequenzeigenschaften der Translokationsbruchpunkte. Sequenzeigenschaften der Regionen - 50 kb (nicht-translozierte Region) und + 50 kb (translozierte Region) der Seax-Rearrangementbruchpunkte. 1) *Mirror Repeats*, 2) G4-Motive (Intensität entspricht der Anzahl an G4-Motiven pro 100 bp), 3) *Alu*-Sequenzen, 4) zDNA.

A. Curriculum Vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

B. Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und unter Verwendung keiner anderen als der von mir angegebenen Quellen und Hilfsmittel verfasst habe. Ferner erkläre ich, dass ich bisher weder an der Freien Universität Berlin noch anderweitig versucht habe, eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

Berlin, den 31.03.2016

.....