

Genomweite Identifizierung von chromosomalen Bruchpunkten bei Tumoren unterschiedlicher Gewebe

Birgit Bortz, Stefanie Marczok, Heike Pospisil*

Zusammenfassung

Aufgrund der jährlich ansteigenden Anzahl an Krebsneuerkrankungen und der tödlichen Verläufe von malignen Tumoren gewinnt die vollständige Aufklärung der Tumorgenese und -progression immer mehr an Bedeutung. Für diese können Untersuchungen zu Bruchpunkten, die eine Kopienzahlvariation (CNV) in Krebsgenomen bewirken, genutzt werden. Es wurde eine Pipeline entwickelt, die in der Lage ist, CNVs, Bruchpunktregionen (BPRs) und Gene genomweit mit Hilfe von SNP-Array-Daten zu detektieren. Dazu wurden 2.820 Tumorproben aus 8 Tumorentitäten untersucht und mit 432 Proben aus gesundem Gewebe verglichen. In den Tumorproben wurden vierfach mehr BPRs detektiert, wobei unter 5 % der Gene in den Normalproben betroffen sind. Wir identifizierten 31 hochspezifische BPRs. Die am häufigsten vorkommende Variation umschließt das Gen *KIAA0513*, welches in Verbindung mit der Apoptose steht. Anhand der hier entwickelten Pipeline können erste Einblicke in CNV- und Bruchpunkt-Muster in Tumorgenomen gewonnen werden, die zu einem verbesserten Verständnis der Tumorgenese und somit zu einer verbesserten Diagnostik und Therapie von Krebserkrankungen führen können.

Abstract

The number of new cancer cases and the deadly courses of malignant tumors are still rising. An entire clarification of tumorigenesis and progression is required for better diagnostics and therapies. Here, we present an analysis of genomic breakpoint patterns, which might cause copy number variations (CNVs) in cancer genomes. Thus, a pipeline was developed which uses SNP array data for the genome-wide detection of CNVs, breakpoint regions (BPRs) and genes. 2,820 tumor samples from eight tumor entities were analyzed and compared with 432 samples from healthy tissue. Four times more BPRs were detected in the tumor samples and only less than 5 % of the affected genes were found to be involved in healthy tissues. We identified 31 highly specific BPRs. The most commonly occurring variation found in our study encloses the *KIAA0513* gene, which is associated with apoptosis. Based on these results, new insights in tumorigenesis are expected.

1. Einleitung

Die Zahl der Krebsneuerkrankungen steigt jedes Jahr weiter an. Der „Welt-Krebs-Bericht 2014“ zeigt auf, dass 2012 die Anzahl an Neuerkrankungen auf 12 Millionen gestiegen ist und 8,2 Millionen Menschen an Krebserkrankungen gestorben sind (Stewart & Wild 2014). Diese Zahlen weisen darauf hin, dass die Ursachen der Karzinogenese weitgehend ungeklärt und somit die Präventions-, Diagnose- und Therapiemöglichkeiten noch unzureichend sind.

Es ist schon lange bekannt, dass aufgrund der steigenden Lebenserwartung die Anzahl an Tumorerkrankungen zunimmt. Jedoch werden

zumeist schlechte Umwelteinflüsse und Verhaltensweisen sowie genetische Vorbelastungen als Hauptursachen für die Tumorentstehung genannt (Vogelstein et al. 2013). Nach einer neuen wissenschaftlichen Studie scheinen diese jedoch bei den meisten Tumorarten keine große Rolle zu spielen; sie wirken eher verstärkend auf die Entstehung ein, sodass sich diese Faktoren potenzieren. Dabei gehen die Autoren davon aus, dass während der Teilung zufällige Mutationen in regulatorischen Bereichen bzw. Genen auftreten, die nicht repariert werden (Tomasetti & Vogelstein 2015). Aufgrund dieser somatischen Mutati-

onen bildet sich die kanzerogene Zelle mit ihren typischen Eigenschaften heraus und weist somit (i) ein Wachstum trotz vorhandener Inhibitoren bzw. abwesender Wachstumsfaktoren, (ii) eine Resistenz gegen Apoptose, (iii) die Fähigkeit zur Angiogenese sowie (iv) Gewebeeinfiltration und (v) Metastasenbildung auf (Hanahan & Weinberg 2000). Diese Fähigkeiten können sich durch genetische Veränderungen, die nicht repariert werden, etablieren. Zu diesen genomischen Veränderungen werden die Kopienzahlvariationen (engl. copy number variations, CNVs) gezählt, bei denen eine vom Referenzgenom abweichende Kopienzahl auftritt. Diese entste-

* korrespondierende Autorin

hen durch fehlerhafte Reparaturmechanismen von DNA-Brüchen während der Replikation bzw. der Rekombination (Redon et al. 2006).

Zahlreiche Studien konnten ein vermehrtes Auftreten von CNVs (Yang et al. 2014) in humanen malignen Tumoren detektieren und einen funktionellen Zusammenhang zur Tumorentstehung nachweisen (Redon et al. 2006, McCarroll et al. 2008).

In einer unserer früheren Studien konnten wir feststellen, dass mit steigendem Tumorstadium vermehrt Regionen mit veränderten Kopienzahlen detektierbar sind. Aufgrund dessen lag die Vermutung nahe, dass die Entstehung von CNVs mit der Kanzerogenese korreliert. Es wurden bei Brustkrebstumoren von Mäusen bestimmte Bruchpunktbereiche identifiziert, in denen vermehrt CNVs auftreten. Diese Bereiche werden in der Literatur als *fragile sites* bezeichnet. In dieser Arbeit wurde untersucht, ob diese Bereiche für verschiedene Tumorarten spezifisch sind, um einen besseren Einblick in die Tumorprogression und auf die genomischen Auswirkungen zu erhalten (Standfuß et al. 2012). Allerdings basierte diese Arbeit auf wenigen Proben.

In anderen Publikationen konnten bereits sogenannte tumorartspezifische und allgemeine Bruchpunkt(BP)-Hotspots mit einer Auflösung von 1 Mbp beim Menschen detektiert werden (Li et al. 2012).

Es ist jedoch eine genauere Auflösung nötig, wenn solche Bruchpunkt-Muster neben der Aufklärung der Tumorgenese und -progression auch als Grundlage für die Entwicklung individueller und nebenwirkungsärmerer Medikamente und für frühzeitige und präzise Präventions- und Diagnoseverfahren genutzt werden sollen.

Um die Veränderung der Kopienzahl und die dazugehörigen Bruchpunkte in Tumorgenomen detektieren zu können, werden verschiedene Technologien genutzt. Die beste Auflösung der Ergebnisse liefern derzeit Hochdurchsatz-Sequenzierungstechnologien. Diese sind jedoch meistens sehr kosten- und zeitintensiv. Um ein schnelleres Screening zu ermöglichen, werden die klassischen Array-Technologien vermehrt verwendet. Bei diesem haben sich in letzter Zeit vor allem die SNP-Arrays als Alternative bewährt, da

sie ein gutes Verhältnis von der physikalischen Abdeckung und Schnelligkeit der Analyse bieten (Zhao et al. 2004). Zudem ist der Vorteil der SNP-Array-Analysen, dass ein großer Datenbestand frei zugänglicher Datensätze vorhanden ist.

Um zu untersuchen, ob spezifische Bruchpunkt-Muster in verschiedenen Tumorgenomen bzw. in spezifischen Tumorarten detektiert werden können, wurde eine Pipeline entwickelt, mit der es ermöglicht wird, eine genomweite Identifizierung und Analyse von chromosomalen Bruchpunktregionen und Segmenten veränderter Kopienzahl anhand von SNP-Array-Daten bei verschiedenen humanen malignen Tumoren zu realisieren (Marczok 2014).

2. Material und Methoden

2.1 SNP-Array-Daten

Maligne Tumordaten: Die verwendeten Tumordaten für die genomweite Identifizierung von Bruchpunktregionen (BPRs) und Segmenten veränderter Kopienzahl wurden der öffentlich zugänglichen GEO-(*Gene Expression Omnibus*)Datenbank vom NCBI (*National Center for Biotechnology Information*) entnommen (National Center for Biotechnology Information 2014). Aus dieser Datenbank wurden nur Daten, die mittels *Genome-Wide Human SNP Array 6.0* analysiert wurden, verwendet. Insgesamt erfolgte die Analyse von 2.820 malignen primären Tumortumoren aus 8 verschiedenen Tumorentitäten (Gehirntumore – 1.187 Proben, Brusttumore – 377 Proben, Kolorektaltumore – 189 Proben, Lungentumore – 291 Proben, Magentumore – 340 Proben, Nierentumore – 109 Proben, Ovarientumore – 207 Proben, Prostata Tumore – 120 Proben).

Referenzdaten: Als Referenzdaten wurden Daten des internationalen HapMap-Projekts (Phase 3, Release #3) (International HapMap Project 2009) zu einer Referenz zusammengeführt. Hierbei wurden nur solche Proben verwendet, die mit dem *Genome-Wide Human SNP Array 6.0* analysiert wurden (990 Proben).

Normalproben: Zusätzlich zu den malignen Primärtumortumoren wur-

den noch 432 Normalgewebeproben, ebenfalls von der GEO-Datenbank stammend (Ovariengewebe – 57 Proben, Gehirngewebe – 29 Proben, Lungengewebe – 62 Proben, Magengewebe – 148 Proben, Nierengewebe – 69), auf BPRs und Segmente veränderter Kopienzahl untersucht. Diese Analysen dienten zur Verifizierung der Krebspezifität potenzieller genomisch veränderter Regionen der malignen Primärtumortumoren.

2.2 Algorithmen

Die Bestimmung von Segmenten veränderter Kopienzahl und tumorentitätsspezifischen, krebspezifischen und allgemeinen BPRs anhand der SNP-Signalintensitäten beinhaltet drei wesentliche Schritte. Als Erstes erfolgte eine Präprozessierung der SNP-Array-Rohdaten, dann die Umrechnung der Signalintensitäten für jeden SNP in einen \log_2 -Ratio in Bezug zur Referenz und abschließend eine Segmentierung (Li & Olivier 2013, Karimpour-Fard et al. 2010). Die Präprozessierung erfolgte mit der Affymetrix Power Tools (APT) Software (Affymetrix 2014). Die weitere Verarbeitung der Daten wurde mit der freien Software R realisiert (R version 3.0.2) (R Core Team 2013). Die Segmentierung und anschließende CNV- und Bruchpunktregions-Bestimmung erfolgte für jede Tumortumore einzeln. Dabei wurde für jedes Segment das entsprechende Chromosom, die Start- und Endposition des Segments, die Anzahl der SNPs, die das jeweilige Segment umfasst, und der Mittelwert des jeweiligen Segments angegeben. Die Start- und Endposition eines Segments sind gekennzeichnet durch die genomischen Positionen des ersten und letzten SNPs eines Segments. Für die Bestimmung der genomischen Positionen der SNPs wurde das humane Referenzgenom hg19/GRCh37 genutzt (Abb. 1).

Erstellung der Referenz: Zur Identifizierung potenzieller genomischer Veränderungen wurde eine Referenz aus 990 HapMap-Daten generiert. Hierfür wurden die 990 Rohdaten zunächst mittels APT-Software (1.16.0 Linux-version) präprozessiert. Die weitere Datenanalyse erfolgte mit der R-Software. Die Signalintensitäten für beide

Allele wurden zusammengefasst und für jeden SNP eine durchschnittliche Signalintensität über alle Proben berechnet, welche als Referenz für die Analysen der Tumorproben verwendet wurde.

Präprozessierung und Bestimmung der Signalintensitäten: Zur Identifizierung von CNVs und BPRs bei allen 2.820 malignen Tumorproben verschiedener Tumorentitäten erfolgte zunächst die Präprozessierung der Rohdaten mittels APT. Hierbei wurde für die Quantilnormalisierung und Hintergrundkorrektur der Birdsee-v2-Algorithmus, unter der Verwendung der Standardparameter, angewendet. Bei diesem handelt es sich um einen Standardalgorithmus von Affymetrix für das Genome-Wide Human SNP Array 6.0 (Yang et al. 2014, Andrade et al. 2011). Die Präprozessierung der malignen Tumorproben wurde separat von den 432 Normalproben durchgeführt. Bei den malignen Tumorproben erfolgte die Präprozessierung jeweils für alle Proben einer Tumorentität. Im Anschluss an den Präprozessierungsschritt erfolgte die Allel-Summutation, um für jeden SNP eine Gesamtsignalintensität zu erhalten (Neuviel et al. 2011).

Segmentierung: Zur Detektion der genomischen Veränderungen bei malignen Tumoren wurde eine Segmentierung jeder Tumorprobe durchgeführt. Hierfür wurde zunächst pro Tumorprobe für jeden SNP der \log_2 -Ratio berechnet. Dabei wurde die jeweilige Gesamtsignalintensität der SNPs ins Verhältnis zu der entsprechenden Referenzintensität des SNPs gesetzt (Li & Olivier 2013, Le Scouarnec & Gribble 2012, Le Scouarnec & Gribble 2011). Basierend auf diesen \log_2 -Ratios erfolgte daraufhin die Segmentierung der 22 Autosomen, um das Genom in zusammenhängende Bereiche (Segmente) konstanter Kopienzahl zu unterteilen. Die Segmentierung erfolgte mittels des von Bioconductor zur Verfügung gestellten R-Pakets DNAcopy (Version 1.32.0), welches den Circular-Binary-Segmentation(CBS)-Algorithmus für die Segmentierung anwendet (Seshan & Olshen 2012). Hierbei wurde das Signifikanzlevel α auf 0,001, die Standardabweichung

(„undo.sd“) auf 0,5 und die Mindestanzahl der Marker („min.width“), die ein Segment enthalten muss, auf 4 gesetzt. Vor der jeweiligen Segmentierung erfolgt eine Erkennung der Ausreißer mit einer anschließenden Glättung der Daten („smooth.CNA“).

Bestimmung von genomischen Bereichen veränderter Kopienzahl und Genen innerhalb dieser Regionen: Für die Identifizierung genomischer Bereiche veränderter Kopienzahl wurden für jede Tumorprobe alle ermittelten Segmente analysiert. Jeder SNP-Position wurde der entsprechende Segmentmittelwert zugeordnet und

über alle Proben einer Tumorentität für jeden SNP ein durchschnittlicher Segmentmittelwert bestimmt. Für diese Werte wurde für jedes Chromosom einer Tumorentität separat der Mittelwert über alle SNPs gebildet, um das Grundsignal zu ermitteln. Anschließend wurde dann die Differenz aus jedem einzelnen SNP-Wert und dem entsprechenden Grundsignal bestimmt. Anhand dieser Werte wurden zusammenhängende Bereiche gesucht, die eine Abweichung von der Kopienzahl zu der Referenz aufweisen. Hierbei liegen amplifizierte Bereiche bei einem Wert $\geq 0,1$ vor und deletierte Bereiche bei einem Wert von $\leq -0,1$.

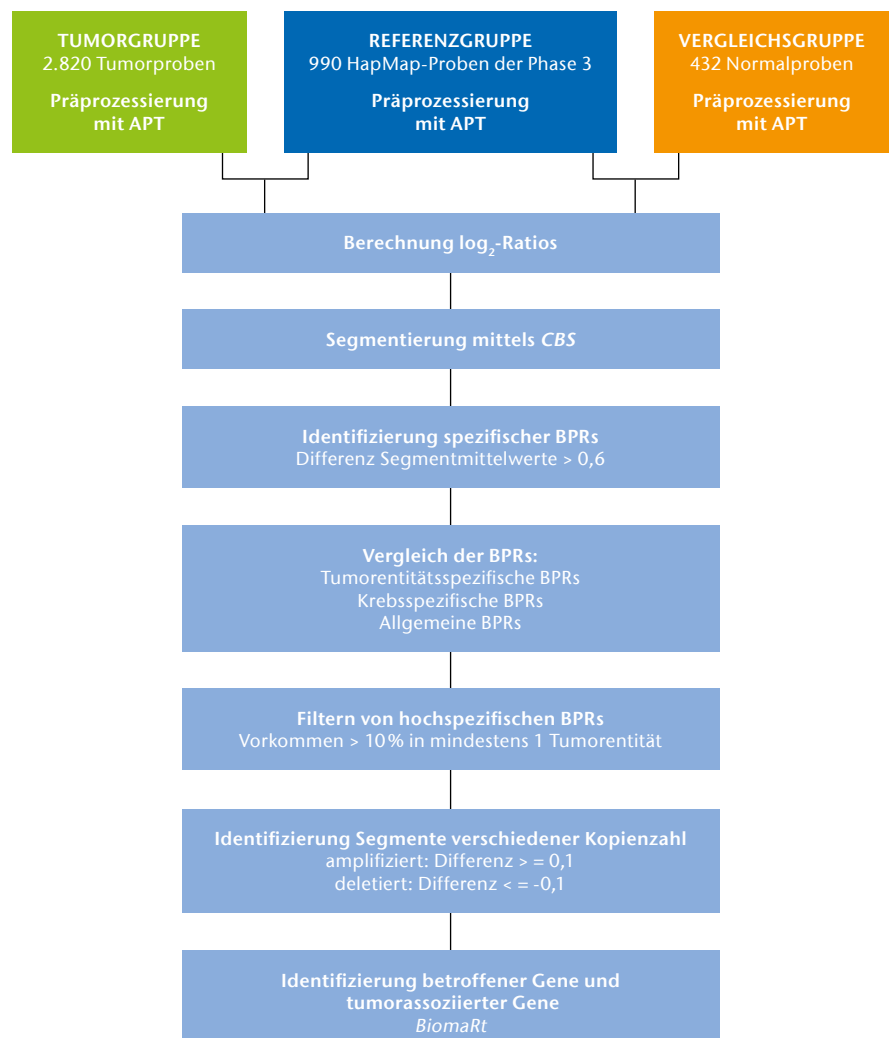


Abb. 1) Schematische Darstellung des Aufbaus der entwickelten Pipeline und der untersuchten Proben. Zunächst wurden 2.820 Tumorproben, 990 HapMap-Proben der Phase 3, die als Referenz dienen, und 432 Normalproben, die als Vergleichsgruppe dienen, mit dem frei zugänglichen Tool APT präprozessiert. Anschließend wurden sowohl die \log_2 -Ratios aus Tumorproben und Referenzdaten als auch aus Normalproben und Referenzdaten gebildet und anschließend mittels CBS das Genom jeder Probe segmentiert. Anhand dieser Segmentierung wurden die spezifischen BPRs identifiziert (Differenz Segmentmittelwerte $> 0,6$). Es wurden die hochspezifischen BPRs (Vorkommen $> 10\%$ in mindestens einer Tumorentität) und die zusammenhängenden Segmente verschiedener Kopienzahl (amplifiziert: Differenz $\geq 0,1$) und die darin kodierten Gene mittels BiomaRt identifiziert.

Innerhalb der ermittelten Regionen veränderter Kopienzahl wurde nach Genen gesucht, die durch einen deletierten oder amplifizierten genomischen Bereich beeinflusst werden können. Diese Untersuchung erfolgte mittels BiomaRt (Durinck et al. 2005). BiomaRt ist ein R-Paket, welches den Zugriff auf verschiedene Datenbanken ermöglicht. Hierbei wurde die Gendektion mittels Ensembl realisiert.

Bestimmung der Bruchpunktregionen: Zur Bestimmung der chromosomalen BPRs erfolgte für jede Tumorsegmente die Betrachtung aller ermittelten Segmente. Hierbei wurden nacheinander alle benachbarten Segmente des jeweiligen Chromosoms miteinander verglichen und eine potenzielle BPR dann angenommen, wenn die Differenz der Segmentmittelwerte zweier benachbarter Segmente $> 0,6$ betrug, das heißt, diese Segmente unterscheiden sich um eine Kopienzahl. Eine potenzielle BPR ergibt sich dann aus der genomischen Endposition des einen Segments und der Startposition des nachfolgenden Segments. Zwischen diesen beiden SNPs liegt ein potenzieller Bruchpunkt, der aufgrund des Microarray-Layouts nicht exakt bestimmt werden kann. Daher können wir nur Bruchpunktregionen und keine exakten Bruchpunkte bestimmen. Um abschließend spezifische BPRs bestimmen zu können, erfolgte die Bestimmung der Häufigkeit für das Auftreten jeder ermittelten BPR über alle Tumorsegmente.

3. Ergebnisse

3.1 Bruchpunkte allgemein

Mit den beschriebenen Methoden wurden insgesamt über alle 2.820 Tumorsegmente 67.076 verschiedene Bruchpunktregionen (BPRs) identifiziert. Davon wurden 13.374 BPRs gefunden, die mindestens in zwei verschiedenen Tumorsegmenten vorkommen. Die größte BPR erstreckt sich über 22.757.511 bp, die kürzeste über 10 bp; durchschnittlich umfasst eine BPR 6.804 bp. Es wurden 130 BPRs identifiziert, die in über 1 % aller Proben, 46 BPRs, die in über 2 % der Proben, und 10 BPRs, die in über 5 % der Proben vorkommen. In den

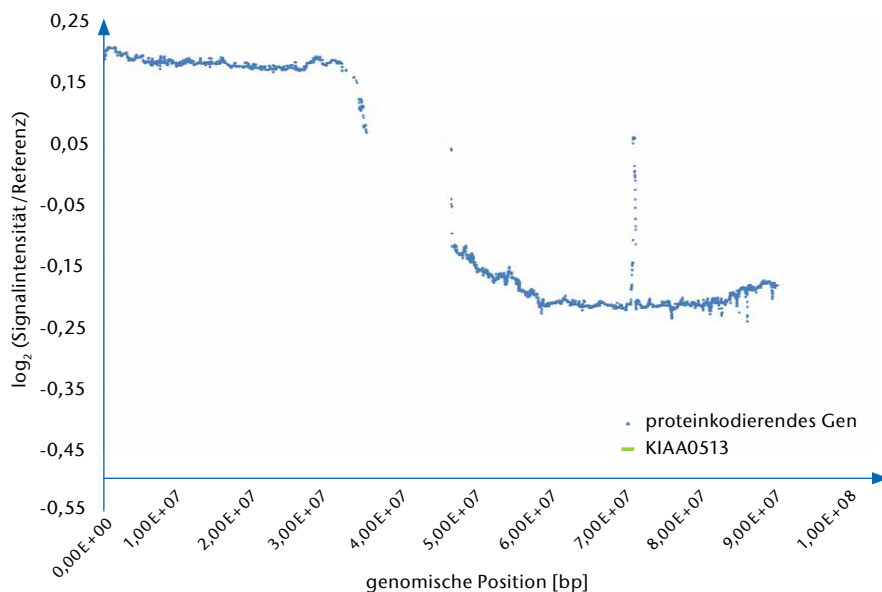


Abb. 2) Bestimmung der strukturellen Veränderungen des Chromosoms 16 der Brusttumore. Es sind die Signalintensitäten (\log_2 -Ratios) pro genomischer Position jedes SNPs in Basenpaaren gegeneinander aufgetragen. Das Grundsignal liegt bei $-0,05$. Es ist eine deutliche Amplifizierung des p-Arms und Deletion des q-Arms erkennbar. Zusätzlich ist innerhalb des q-Arms eine spezifische Amplifikation bei 70.854.381 bp bis 71.065.311 bp und eine spezifische Deletion bei 85.091.864 bp bis 85.092.748 bp, welches das Gen KIAA0513 (85.061.374 bp bis 85.127.836 bp) umschließt, erkennbar. Von 85.091.864 bp bis 85.092.892 bp liegen die zwei zugehörigen BPRs. Diese hatten über alle detektierten BPRs die höchste Vorkommenswahrscheinlichkeit aufgezeigt.

untersuchten Normalproben wurden weitere 16.549 BPRs identifiziert, die teilweise zusätzlich in den Tumorsegmenten vertreten sind. Davon kommen 8.449 BPRs nur in den Normalproben vor.

3.2 Bruchpunktvergleich zwischen den Tumorsegmenten zur Mustererkennung

Um spezifische Muster zu identifizieren, die auf eine Krebspezifität oder Tumorsegmentenspezifität hindeuten, wurden für die jeweiligen BPRs die relativen Häufigkeiten pro Tumorsegment bestimmt und miteinander verglichen. Es wurden 31 BPRs identifiziert, die in mindestens einer Tumorsegmenten über 10 % vorkommen. Diese BPRs wurden aufgrund ihres häufigen Auftretens als hochspezifische BPRs charakterisiert. Im Folgenden wird eine kleine Auswahl an Ergebnissen dieser Bereiche dargestellt (Marczok et al. 2015).

Es konnten zwei BPRs identifiziert werden, bei denen in 7 der 8 Tumorsegmenten (und in den Normalproben) bis zu 27 % (5 %) sowie bis zu 42 % (5 %) ein Bruch der DNA im Chromosom 16 bei ca. 85 Mbp aufgetreten ist.

Zusätzlich wurden BPRs detektiert, die lediglich in einer Tumorsegmenten vermehrt aufgetreten sind. Im Chromosom

17 wurde in Gehirntumorsegmenten zu 14,32 % eine BPR detektiert. In den restlichen Tumorsegmenten trat diese jeweils in weniger als 0,6 % der Proben auf.

Acht der hochspezifischen BPRs konnten auch in den Normalproben mit einer Häufigkeit von mehr als 5 % detektiert werden. Die Häufigkeit des Vorkommens in den Normalproben ist aber immer geringer als mindestens eine der Häufigkeiten von den Tumorsegmenten.

Aufgrund der Ergebnisse der BPR-Bestimmung klassifizierten wir die BPRs in vier verschiedene Gruppen, abhängig von der Häufigkeit des Auftretens der BPR und ob diese auch in den Normalproben gefunden wurde. Die Einteilung erfolgte in

1. tumorsegmentenspezifische BPRs (Häufigkeit ≥ 1 % ausschließlich in einer Tumorsegmenten (TE) oder ≥ 1 % in einer TE und gesundem Gewebe) (2.278 detektierte BPRs),
2. krebspezifische BPRs (Häufigkeit ≥ 1 % in mehr als einer TE und in Normalgewebe $< 0,5$ %) (231 detektierte BPRs) und
3. die allgemeinen BPRs (Häufigkeit ≥ 1 % in mehr als einer TE und $\geq 0,5$ % Normalgewebe) (207 detektierte BPRs).

Die Gruppe der nicht krebsspezifischen BPRs beinhaltet jene, welche mit einer Häufigkeit von weniger als 1 % in allen Tumorentitäten (72.809 detektierte BPRs) detektiert wurden.

3.3 Vergleich der Bereiche verschiedener Kopienzahl über verschiedene Tumorentitäten

Um den eventuellen Einfluss der BPRs auf die Tumorentstehung und -progression zu analysieren, wurden die Genombereiche untersucht, die eine veränderte Kopienzahl aufweisen, um festzustellen, ob in diesen bestimmte Gene lokalisiert sind. Es wurden sowohl genomische Bereiche mit verringerter als auch erhöhter Kopienzahl identifiziert.

Es konnten insgesamt 563 (10 in Normalproben) zusammenhängende Segmente mit einer von der Referenz abweichenden Kopienzahl detektiert werden. Innerhalb dieser Segmente liegen 20.084 (7 in Normalproben) verschiedene Gene (protein kodierend und nicht protein kodierend) von denen bei 169 (keine in Normalproben) bekannt ist, dass sie mit der Tumorentstehung und -progression in Verbindung stehen.

In 7 von 8 Tumorentitäten konnte in Chromosom 16 bei ca. 85 Mbp im Bereich des Genes *KIAA0513* eine niedrigere Kopienzahl als im genomischen

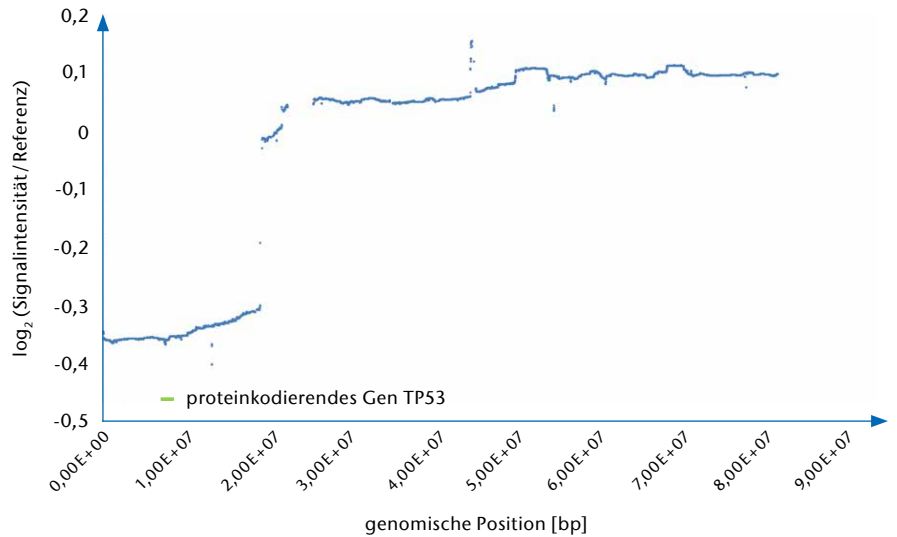


Abb. 3) Bestimmung der strukturellen Veränderungen des Chromosoms 17 der Gehirntumore. Es sind die Signalintensitäten (\log_2 -Ratios) pro genomischer Position jedes SNPs in Basenpaaren gegeneinander aufgetragen. Das Grundsignal liegt bei -0,04. Es ist eine deutliche Deletion des p-Armes und Amplifizierung des q-Armes erkennbar. Innerhalb der deutlichen Amplifizierung zwischen 6.689 bp und 18.917.915 bp des p-Armes liegt der kodierende Bereich des Gens TP53 (7.565.097 bp – 7.590.856 bp). Die zugehörige Bruchpunktregion ist bei 8.917.915 bp bis 19.168.912 bp lokalisiert.

Mittel detektiert werden. Zudem konnte in diesem Chromosom bei ca. 70 Mbp lediglich in Brustkrebs eine prägnante Amplifikation detektiert werden (Erhöhung um bis zu 2,6 vom Grundsignal) (Abb. 2). Zusätzlich ist bei Gehirntumorgeweben im Chromosom 17 der Bereich des p-Arms deletiert (um 0,6 geringer als das Grundsignal) (Abb. 3). Es konnte jedoch auch eine leichte Deletion

in Brust-, Kolorektal- und Ovarientumormproben (ca. 0,1 geringer als das Grundsignal) festgestellt werden. Dieser Bereich umschließt das Gen TP53. In Chromosom 8 konnte bei ca. 45 Mbp ein Bereich mit verringerter Kopienzahl in Gehirntumoren (unter 0,3 vom Grundsignal) und eine weniger ausgeprägte bei ca. 6 Mbp identifiziert werden. In diesem Bereich ist das Gen *RP11-115C21.2* kodiert (Abb. 4). Bei den Normalgewebeproben ist die Veränderung der Kopienzahl immer geringer als mindestens eine der Veränderungen einer Tumorentität.

4. Diskussion

4.1 Vergleich identifizierter BPRs und Segmente veränderter Kopienzahl zwischen Tumormproben und Normalproben

Die von uns durchgeführten genomweiten Analysen zeigen, dass die Anzahl der insgesamt ermittelten Bruchpunktregionen (BPRs) und Bereiche veränderter Kopienzahl in Tumorgenomen höher ist als in gesunden Genomen. Hierbei ist die Anzahl der verschiedenen BPRs über alle Tumormproben um ein Vierfaches höher als bei den Normalproben. Auch der Vergleich der Anzahl der Segmente veränderter Kopienzahl zwischen

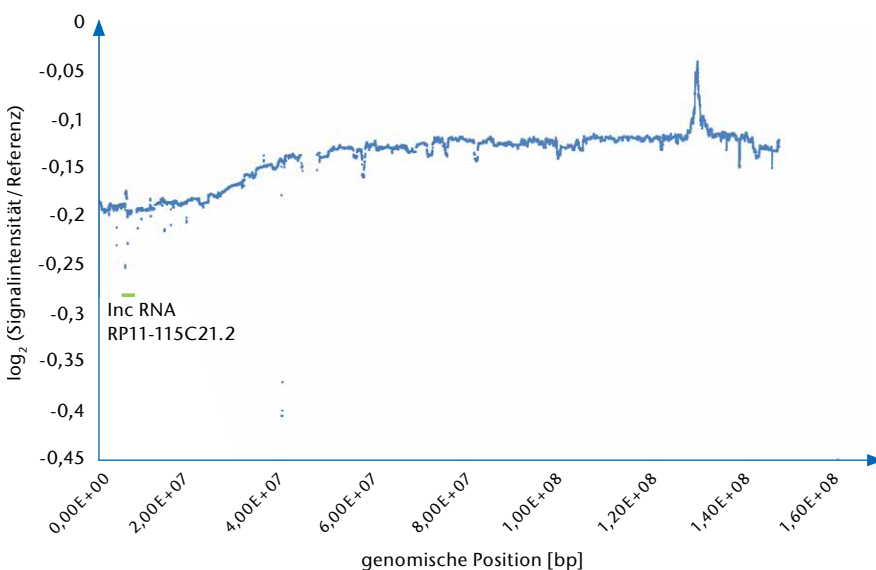


Abb. 4) Bestimmung der strukturellen Veränderungen des Chromosoms 8 der Gehirntumore. Es sind die Signalintensitäten (\log_2 -Ratios) pro genomischer Position jedes SNPs in Basenpaaren gegeneinander aufgetragen. Das Grundsignal liegt bei -0,14. Es ist eine deutliche Deletion bei 43.778.914 bp bis 46.924.211 bp und eine weitere weniger spezifische bei ca. 6 Mbp erkennbar. Letzterer Bereich umschließt den kodierenden Bereich des *RP11-115C21.2* (Chr. 8: 6261.072 bp – 6.264.663 bp).

den einzelnen Tumorentitäten und den Normalproben zeigt, dass diese bei den Tumorproben höher ist. Dies konnte auch durch andere Studien bestätigt werden (Yang et al. 2014, Beroukhi et al. 2010). Durchschnittlich wurden 70 Segmente pro Tumorentität ermittelt und nur 10 bei den Normalproben. Die geringere Anzahl an BPRs und Segmenten veränderter Kopienzahl in Normalproben im Vergleich zu Tumorproben unterstützt die Aussage, dass genomische Veränderungen und vor allem CNVs auch zu 4,8–9,5 % in gesunden Genomen vorkommen. Neben dem geringeren Auftreten der Segmente veränderter Kopienzahl in den Normalproben konnten wir innerhalb dieser eine geringere Anzahl betroffener Gene, die laut Zarrei et al. (2015) keinen Einfluss auf den Phänotyp haben können, detektieren. Dabei ist vom besonderen Interesse, dass keine Gene betroffen waren, die als tumorassoziiert bekannt sind.

4.2 Vergleich identifizierter BPRs und Segmente veränderter Kopienzahl zwischen verschiedenen Tumorentitäten

Anhand unserer Analysen konnten gleiche BPRs und Segmente veränderter Kopienzahl in mehreren Tumorentitäten gefunden werden. Zur Untersuchung, ob es tumorentitätsspezifische, krebspezifische oder allgemeine genomische Veränderungen sind, erfolgte der Vergleich dieser Veränderungen zwischen den 8 verschiedenen Tumorentitäten. Dabei konnten wir feststellen, dass scheinbar Muster bei den genomweit detektierten BPRs und Segmenten veränderter Kopienzahl existieren.

Hierbei kann die Zuordnung der ermittelten BPRs zu einer entsprechenden Gruppe einen möglichen Ansatz bieten, um weitere Erkenntnisse über die komplexen Prozesse der Tumorgenese und -progression zu erhalten.

Beim Vergleich der drei krebsrelevanten Gruppen ist erkennbar, dass die meisten BPRs tumorentitätsspezifisch sind. Jede Tumorentität weist zu einem gewissen Anteil ein individuelles BPR-Muster auf, was ein Indiz dafür sein könnte, dass die Tumorgenese und -progression zumindest teilweise auf individuelle genomische Veränderun-

gen zurückzuführen ist. Ein möglicher Grund hierfür könnten gewebsspezifische Unterschiede sein, denn es ist bekannt, dass jedes Gewebe unterschiedlich differenziert ist, sodass verschiedene genomische Bereiche aktiv sind. Das bewirkt unter anderem eine unterschiedliche Replikationshäufigkeit. Beispielsweise ist die Wahrscheinlichkeit des Auftretens eines malignen Herztumors sehr gering, da die Kardiomyozyten postmitotische Zellen sind (McAllister & Fenoglio 1978). Folglich kann die Differenzierung einer Zelle einen entscheidenden Einfluss auf das individuelle BPR-Muster haben und somit auch auf die Krebsentstehung und -progression.

Die am zweithäufigsten vorkommende Gruppe der ermittelten BPRs ist die krebspezifische. Die BPRs, die dieser Klasse zugeordnet werden, sind nur in Tumorgenomen und nicht in gesunden Genomen detektierbar, sodass diese Hinweise auf die Mechanismen der allgemeinen Tumorgenese und -progression geben können, die die gemeinsamen Eigenschaften von malignen Tumoren (Hallmarks of Cancer) hervorrufen (Hanahan & Weinberg 2000). Diese können bedeutend für die Aufklärung von allgemeinen Krebsrisikofaktoren, die die Krebsentstehung begünstigen, in Bezug auf bestimmte genomische Positionen sein.

Die allgemeinen BPRs stellen die dritte krebsrelevante Gruppe dar. Die identifizierten BPRs und Segmente veränderter Kopienzahl, die in den Tumorproben vorkommen, sind mitunter auch in den untersuchten Normalproben ermittelt worden. Dabei ist erkennbar, dass teilweise die entsprechenden BPRs in den Tumorproben mit einer sehr viel höheren Auftretenswahrscheinlichkeit vorkommen als in den Normalproben. Diese Gruppe unterstützt die Theorie der *fragile sites*. Hierbei wird davon ausgegangen, dass das gesunde Genom an bestimmten genomischen Stellen instabiler ist als an anderen. Die geringere Stabilität macht die DNA in diesen Bereichen anfälliger für Brüche (Gorgoulis et al. 2005). Aufgrund der hohen Teilungsrate der Krebszellen ist die Wahrscheinlichkeit für weitere Brüche in diesen Abschnitten sehr hoch. Dieses Phänomen konnte, wie auch schon bei beispielsweise Thys et al. (2015), durch das erhöhte Vorkommen

der von uns ermittelten allgemeinen BPRs in den einzelnen Tumorentitäten verifiziert werden.

4.3 Vorliegen von Hotspot-Bereichen

Beim Vergleich der strukturellen Veränderungen ist zum einen ein erhöhtes Auftreten mehrerer benachbarter BPRs und zum anderen das Überlappen von Segmenten gleich veränderter Kopienzahl zu erkennen. Dies impliziert das Vorkommen von sogenannten Hotspot-Bereichen, welche entweder auf technische, d. h. durch die Auflösung bzw. das Layout der Detektionsmethode (Affymetrix 2009) oder biologische Gegebenheiten zurückgeführt werden können. Letztere können auf das Vorliegen von *fragile sites* hindeuten, welche Genombereiche darstellen, die aufgrund geringerer Stabilität leichter brechen (Gorgoulis et al. 2005). Unsere Ergebnisse und die von Tomasetti & Vogelstein (2015) (zufälliger Bruch der DNA) lassen vermuten, dass die DNA in bestimmten Bereichen anfälliger ist, zufällig zu brechen.

Zudem ist erwähnenswert, dass es scheinbar zwei verschiedene Ausprägungsformen von CNVs gibt. Neben denen, bei denen lediglich Bereiche mit relativ kleiner Ausdehnung (bis einige Mbp) betroffen sind (Abb. 4), konnten vermehrt Strukturvariationen detektiert werden, bei denen komplette chromosomale Arme einer Kopienzahlveränderung unterlagen (Abb. 2 und 3). Die Bedeutung dieser Ereignisse bedarf jedoch noch weiterer Untersuchungen.

4.4 Identifizierung von 31 hochspezifischen BPRs

Um die Bedeutung der detektierten hochspezifischen BPRs einschätzen zu können, wurden sie mit den identifizierten Segmenten veränderter Kopienzahl und den potenziell betroffenen Genbereichen verglichen. Dabei konnten Zusammenhänge zu sowohl proteinkodierenden als auch nicht kodierenden Genen (z. B. lncRNA) hergestellt werden. Dabei sind 30 % der identifizierten Gene bereits in Zusammenhang mit der Tumorgenese und -progression gestellt worden. Die meisten der 31 hochspezifischen

BPRs können der allgemeinen Gruppe (58 %) zugeordnet werden. Die hierbei mit der höchsten Auftretswahrscheinlichkeit ermittelte BPR ist innerhalb des proteinkodierenden Gens *KIAA0513* lokalisiert (Chr. 16: 85.061.374 bp–85.127.836 bp). Es ist bekannt, dass es eine wichtige Rolle bei immunologischen, synaptischen und apoptotischen Signalwegen spielt. Somit könnte dieses Gen bei einer Deletion in diesem Bereich, wie wir sie in 7 der 8 Tumorentitäten und den Normalproben detektiert haben, zu einer Störung der Expression des Proteins führen, sodass es der Funktion als Signalmolekül in der Apoptose nicht nachkommen (Lauriat et al. 2006) und somit die Tumorgenese begünstigen kann.

25 % der hochspezifischen BPRs können den krebspezifischen BPRs zugeordnet werden. Das Gen *RP11-115C21.2* (Chr. 8: 6.261.072 bp–6.264.663 bp), welches für eine *long non-coding* RNA (lnc RNA) kodiert, liegt unmittelbar in der Nähe einer BPR in über 75 % der TE (> 1 % Vorkommen) und ist in 6 Tumorentitäten innerhalb von verschiedenen deletierten Segmenten lokalisiert. In den letzten Jahren wurde die Bedeutung der lnc RNAs für die Tumor- und Mutagenese aufgedeckt. Diese können entweder als Onkogene oder als Tumorsuppressorgene wirken, indem sie die Chromatinstruktur verändern (Gutschner et al. 2013, Gutschner & Diederichs 2012) und die Transkription von proteinkodierenden Genen beeinflussen können (Huarte et al. 2010). Aufgrund unserer Analysen lässt sich vermuten, dass aufgrund des in diesem Genbereich ermittelten deletierten Segments zellzyklus- oder apoptoseregulierende Eigenschaften (Xue et al. 2012) verloren gehen und somit das Gen *RP11-115C21.2* als Tumorsuppressor wirkt.

Lediglich 16 % der 31 hochspezifischen BPRs können den tumorentitätsspezifischen BPRs zugeordnet werden. Dies lässt vermuten, dass je signifikanter eine BPR ist, desto höher ist die Wahrscheinlichkeit, dass diese in mehreren Tumorentitäten vorkommt. Dies wurde schon von einer Studie zu somatischen Kopienalterationen vermutet (Beroukhi et al. 2010), sodass möglicherweise die genomischen Veränderungen, die eine

Krebserkrankung begünstigen, bei den verschiedenen Tumorentitäten sehr ähnlich und nur einige wenige individuelle genomische Veränderungen kennzeichnend für die einzelnen Tumorentitäten sind.

Eine der hochspezifischen BPRs auf Chr. 17 ist nur in den Gehirntumoren mit ausreichender Häufigkeit detektiert worden. In drei weiteren Tumorentitäten ist der komplette Bereich deletiert. Es konnten mehrere tumorassoziierte Gene detektiert werden, die in diesem Bereich potenziell beeinflusst werden. Das bekannteste Gen ist das *TP53* (7.565.097 bp–7.590.856 bp), welches den Tumorsuppressor p53 kodiert. Dieser hat wichtige Funktionen in der Aufrechterhaltung des Zellzyklus, in der Regelung der Apoptose sowie des Alterungsprozesses, bei DNA-Reparaturen und Veränderungen des Metabolismus (Lane 1992, Peifer et al. 2012). Zudem ist an diesem Beispiel erkennbar, dass aufgrund der hohen strikten Bedingungen zur Detektion von spezifischen BPRs lediglich in einer Tumorentität eine BPR detektiert wurde, jedoch noch in weiteren Tumorentitäten Bereiche veränderter Kopienzahl identifiziert wurden. Diese Schwellenwerte wurden jedoch gewählt, um lediglich die hochspezifischen BPRs für eine Mustererkennung nutzen zu können.

Diese Auswahl an Beispielen zeigt, dass die Pipeline geeignet ist, um einen ersten Eindruck über potenzielle Bruchpunkt-Muster in malignen Tumorgenomen bekommen zu können und dass sich die Ergebnisse mit den in der Literatur beschriebenen Erkenntnissen bestätigen lassen. Somit kann die hier beschriebene Methode genutzt werden, um neue Einsichten in die Tumorgenese zu gewinnen.

5. Ausblick

Die hier identifizierten Bruchpunktregionen (BPRs) müssen noch weiter analysiert und präzisiert werden, da aufgrund des SNP-Array-Layouts nicht alle Bereiche des Genoms abgedeckt sind, sodass nicht alle genomischen Bereiche auf Bruchpunkt- und CNV-Regionen untersucht werden konnten. Des Weiteren nehmen die ermittelten Bruchpunkt- und CNV-Regionen teilweise sehr weite Bereiche des Genoms ein. Mit Hilfe der Analyse von Whole-

Genome-Sequenzierungs-Daten und Whole-Exome-Sequenzierungs-Daten (z. B. aus dem Projekt The Cancer Genome Atlas (TCGA)) können die Bruchpunkt- und CNV-Regionen basengenau bestimmt werden, womit eine genauere Aussage über den Einfluss solcher genomischer Muster auf die Tumorgenese und -progression möglich ist. Damit können potenziell kritische Gene identifiziert werden, die eine entscheidende Rolle für die Krebsprädisposition spielen. Die ermittelten genomischen Muster können sowohl Hinweise auf eine mögliche familiäre Krebsprädisposition liefern als auch als diagnostische Marker für die Tumorcharakterisierung dienen. Somit wäre eine personalisierte Vorsorge, Behandlung und Nachsorge von Patienten möglich und die Grundlagen für die Entwicklung wesentlich zeiteffizienterer, kostengünstigerer und risikoärmerer Präventions- und Diagnoseverfahren geschaffen, sodass lediglich eine Blutuntersuchung eine patientenspezifische Krebsprävention, -diagnose und -therapie ermöglichen kann. So könnten die Heilungschancen der Betroffenen verbessert werden, da frühzeitig eine Behandlung begonnen und die Überlebensrate gesteigert werden kann, bevor erste sichtbare Symptome erkennbar sind.

Die aus diesem Projekt gewonnenen Erkenntnisse können somit die Grundlage auf dem Weg zur Entwicklung effektiverer präventiver, diagnostischer und therapeutischer Maßnahmen bilden, um so die Zahl der Neuerkrankungen und Todesfälle bei malignen Tumorerkrankungen zu minimieren sowie die Überlebensrate der Betroffenen zu erhöhen.

LITERATUR

Affymetrix (2009) Genome-Wide Human SNP Array 6.0. Data Sheet, rev 2:1–4

Affymetrix (2014) Affymetrix Power Tools. http://www.affymetrix.com/estore/partners_programs/programs/developer/tools/power_tools.affx#1_1. Accessed 27 May 2014

Andrade M de, Atkinson EJ, Bamlet WR, Matsumoto ME, Maharjan S, Slager SL, Vachon CM, Cunningham JM, Kardia SLR (2011) Evaluating the Influence of Quality Control Decisions and Software Algorithms on SNP Calling for the Affymetrix 6.0 SNP Array Platform. *Hum Hered* 71(4):221–233. doi: 10.1159/000328843

Beroukhi R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, Mc Henry KT, Pinchback RM, Ligon AH, Cho Y, Haery L, Greulich H, Reich M, Winckler W,

- Lawrence MS, Weir BA, Tanaka KE, Chiang DY, Bass AJ, Loo A, Hoffman C, Prensner J, Liefeld T, Gao Q, Yecies D, Signoretti S, Maher E, Kaye FJ, Sasaki H, Tepper JE, Fletcher JA, Taberero J, Baselga J, Tsao M, Demichelis F, Rubin MA, Janne PA, Daly MJ, Nucera C, Levine RL, Ebert BL, Gabriel S, Rustgi AK, Antonescu CR, Ladanyi M, Letai A, Garraway LA, Loda M, Beer DG, True LD, Okamoto A, Pomeroy SL, Singer S, Golub TR, Lander ES, Getz G, Sellers WR, Meyerson M (2010) The landscape of somatic copy-number alteration across human cancers. *Nature* 463(7283):899–905. doi: 10.1038/nature08822
- Durinck S, Moreau Y, Kasprzyk A, Davis S, Moor B de, Brazma A, Huber W (2005) BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 21(16):3439–3440. doi: 10.1093/bioinformatics/bti525
- Gorgoulis VG, Vassiliou LF, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, Venere M, Dittulio RA, Kastrinakis NG, Levy B, Kletsas D, Yoneta A, Herlyn M, Kittas C, Halazonetis TD (2005) Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 434(7035):907–913. doi: 10.1038/nature03485
- Gutschner T, Diederichs S (2012) The hallmarks of cancer. A long non-coding RNA point of view. *RNA Biol* 9(6):703–719. doi: 10.4161/rna.20481
- Gutschner T, Hämmerle M, Eissmann M, Hsu J, Kim Y, Hung G, Revenko A, Arun G, Stentrup M, Gross M, Zörnig M, MacLeod AR, Spector DL, Diederichs S (2013) The Noncoding RNA MALAT1 is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Res* 73(3):1180–1189. doi: 10.1158/0008-5472.CAN-12-2850
- Hanahan D, Weinberg RA (2000) The Hallmarks of Cancer. *Cell* 100(1):57–70. doi: 10.1016/S0092-8674(00)81683-9
- Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M, Attardi LD, Regev A, Lander ES, Jacks T, Rinn JL (2010) A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. *Cell* 142(3):409–419. doi: 10.1016/j.cell.2010.06.040
- International HapMap Project (2009) Raw Data Download Affymetrix 6.0. http://hapmap.ncbi.nlm.nih.gov/downloads/raw_data/hapmap3_affy6.0. Accessed 31 Jul 2014
- Karimpour-Fard A, Dumas L, Phang T, Sikela JM, Hunter LE (2010) A survey of analysis software for array-comparative genomic hybridisation studies to detect copy number variation. *Hum Genomics* 4(6):421–427. doi: 10.1186/1479-7364-4-6-421
- Lane DP (1992) Cancer. p53, guardian of the genome. *Nature* 358(6381):15–16. doi: 10.1038/358015a0
- Lauriat TL, Dracheva S, Kremerskothen J, Duning K, Haroutunian V, Buxbaum JD, Hyde TM, Kleinman JE, McInnes LA (2006) Characterization of KIAA0513, a novel signaling molecule that interacts with modulators of neuroplasticity, apoptosis, and the cytoskeleton. *Brain Res* 1121(1):1–11. doi: 10.1016/j.brainres.2006.08.099
- Le Scouarnec S, Gribble SM (2012) Characterising chromosome rearrangements: recent technical advances in molecular cytogenetics. *Heredity* 108:75–85. doi: 10.1038/hdy.2011.100
- Li W, Olivier M (2013) Current analysis platforms and methods for detecting copy number variation. *Physiol Genom* 45(1):1–16. doi: 10.1152/physiolgenomics.00082.2012
- Li Y, Zhang L, Ball RL, Liang X, Li J, Lin Z, Liang H (2012) Comparative analysis of somatic copy-number alterations across different human cancer types reveals two distinct classes of breakpoint hotspots. *Hum Mol Genet* 21(22):4957–4965. doi: 10.1093/hmg/dd5340
- Marczok S (2014) Genomweite Identifizierung von Kopienzahlvariationen in humanen malignen Tumoren. Master Thesis, Technische Hochschule Wildau
- Marczok S, Bortz B, Wang C, Pospisil H (2015) Comprehensive Analysis of Genome Rearrangements in Eight Human Malignant Tumor Tissues. *PLoS ONE* (submitted)
- McAllister H, Fenoglio J (1978) Tumors of the Cardiovascular System. In: Hartmann WH, Cowan WR (eds) *Atlas of Tumor Pathology*, 2nd ser, fasc 15. Armed Forces Institute of Pathology, Washington DC, pp 1–20
- McCarroll SA, Kuruville FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker, Paul I W, Maller JB, Kirby A, Elliott AL, Parkin M, Hubbell E, Webster T, Mei R, Veitch J, Collins PJ, Handsaker R, Lincoln S, Nizari M, Blume J, Jones KW, Rava R, Daly MJ, Gabriel SB, Altshuler D (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 40(10):1166–1174. doi: 10.1038/ng.238
- National Center for Biotechnology Information (2014) Gene Expression Omnibus. <http://www.ncbi.nlm.nih.gov/geo>. Accessed 20 Jan 2015
- Neuville P, Bengtsson H, Speed TP (2011) Statistical Analysis of Single Nucleotide Polymorphism Microarrays in Cancer Studies. In: Lu HH, Schölkopf B, Zhao H (eds) *Handbook of Statistical Bioinformatics*. Springer Handbooks of Computational Statistics. Springer, Berlin/Heidelberg, ISBN: 978-3-642-16344-9, pp 225–255. doi: 10.1007/978-3-642-16345-6_11
- Peifer M, Fernández-Cuesta L, Sos ML, George J, Seidel D, Kasper LH, Pfenker D, Leenders F, Sun R, Zander T, Menon R, Koker M, Dahmen I, Müller C, Di Cerbo V, Schildhaus H, Altmüller J, Baessmann I, Becker C, Wilde B de, Vandesompele J, Böhm D, Ansén S, Gabler F, Wilkening I, Heynck S, Heuckmann JM, Lu X, Carter SL, Cibulskis K, Banerji S, Getz G, Park K, Rauh D, Grütter C, Fischer M, Pasqualucci L, Wright G, Wainer Z, Russell P, Petersen I, Chen Y, Støelben E, Ludwig C, Schnabel P, Hoffmann H, Muley T, Brockmann M, Engel-Riedel W, Muscarella LA, Fazio VM, Groen H, Timens W, Sietsma H, Thunnissen E, Smit E, Heideman DAM, Snijders PJF, Cappuzzo F, Ligorio C, Damiani S, Field J, Solberg S, Brustugun OT, Lund-Iversen M, Sängler J, Clement JH, Soltermann A, Moch H, Weder W, Solomon B, Soria J, Validire P, Besse B, Brambilla E, Brambilla C, Lantuejoul S, Lorimier P, Schneider PM, Hallek M, Pao W, Meyerson M, Sage J, Shendure J, Schneider R, Büttner R, Wolf J, Nürnberg P, Perner S, Heukamp LC, Brindle PK, Haas S, Thomas RK (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* 44(10):1104–1110. doi: 10.1038/ng.2396
- R Core Team (2013) R: A Language and Environment for Statistical Computing. <http://www.R-project.org>
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacós M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurler ME (2006) Global variation in copy number in the human genome. *Nature* 444(7118):444–454. doi: 10.1038/nature05329
- Seshan VE, Olshen A (2012) DNACopy: DNA copy number data analysis. R package version 1.32.0
- Standfuß C, Pospisil H, Klein A (2012) SNP Microarray analyses reveal copy number alterations and progressive genome reorganization during tumor development in SVT/t driven mice breast cancer. *BMC Canc* 12:380. doi: 10.1186/1471-2407-12-380
- Stewart BW, Wild CP (eds) (2014) *World Cancer Report 2014*. World Health Organization: International Agency for Research on Cancer, Lyon. ISBN: 978-92-832-0429-9
- Thys RG, Lehman CE, Pierce LCT, Wang Y (2015) DNA Secondary Structure at Chromosomal Fragile Sites in Human Disease. *Curr Genomics* 16(1):60–70. doi: 10.2174/1389202916666150114223205
- Tomasetti C, Vogelstein B (2015) Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 347(6217):78–81. doi: 10.1126/science.1260825
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW (2013) Cancer Genome Landscapes. *Science* 339(6127):1546–1558. doi: 10.1126/science.1235122
- Xue W, Kitzing T, Roessler S, Zuber J, Krasnitz A, Schultz N, Revill K, Weissmueller S, Rappaport AR, Simon J, Zhang J, Luo W, Hicks J, Zender L, Wang XW, Powers S, Wigler M, Lowe SW (2012) A cluster of cooperating tumor-suppressor gene candidates in chromosomal deletions. *Proc Natl Acad Sci USA* 109(21):8212–8217. doi: 10.1073/pnas.1206062109
- Yang R, Chen B, Pfüzte K, Buch S, Steinke V, Holinski-Feder E, Stöcker S, Schönfels W von, Becker T, Schackert HK, Royer-Pokora B, Kloor M, Schmiegel WH, Büttner R, Engel C, Laszcz Puertolas J, Försti A, Kunkel N, Bugert P, Schreiber S, Krawczak M, Schafmayer C, Propping P, Hampe J, Hemminki K, Kurwinkel B (2014) Genome-wide analysis associates familial colorectal cancer with increases in copy number variations and a rare structural variation at 12p12.3. *Carcinogenesis* 35(2):315–323. doi: 10.1093/carcin/bgt344
- Zarrei M, MacDonald JR, Merico D, Scherer SW (2015) A copy number variation map of the human genome. *Nat Rev Genet* 16(3):172–183. doi: 10.1038/nrg3871
- Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen T, Girard L, Minna J, Christiani D, Leo C, Gray JW, Sellers WR, Meyerson M (2004) An Integrated View of Copy Number and Allelic Alterations in the Cancer Genome Using Single Nucleotide Polymorphism Arrays. *Cancer Res* 64(9):3060–3071. doi: 10.1158/0008-5472.CAN-03-3308

AUTOREN

Birgit Bortz
Stefanie Marczok
Prof. Dr. rer. nat. Heike Pospisil
 Technische Hochschule Wildau
 Forschungsgruppe High Performance
 Computing in Life Sciences

E-Mail für Korrespondenz:
 heike.pospisil@th-wildau.de