

Strukturelle Genomvarianten

Ausmaß, Entstehung und phänotypische Konsequenzen

Unsere Vorstellung über die Variabilität des menschlichen Genoms wird in beachtlichem Maß von den technischen Eckdaten der jeweils verfügbaren Analysemethoden geprägt. Vor der Einführung hochauflösender DNA-Arrays und moderner Sequenzierverfahren waren genomische Strukturvarianten (SV) in Form von balancierten und unbalancierten Chromosomenveränderungen in der gesunden Bevölkerung ein vergleichsweise seltener Befund. Die phänotypische Variabilität des Menschen wurde damals größtenteils den millionenfach vorkommenden Einzelnukleotidpolymorphismen (SNP) zugeschrieben. Erst der Einsatz von Techniken zur genomweiten Analyse von submikroskopischen Veränderungen gab Hinweise auf die Häufigkeit von SV und rückte sie in den Mittelpunkt des wissenschaftlichen Interesses.

Obwohl im engeren Sinne alle chromosomalen Rearrangements größer als eine Base SV darstellen, assoziiert man mit diesem Begriff heute v. a. submikroskopische Veränderungen mit einer Größe von über 1 kb [5]. Diese Größeneinschränkung ist etwas willkürlich gewählt, erleichtert aber u. a. die Abgrenzung zu simplen Repeats. Den mit Abstand größten Anteil an den SV haben DNA-Kopienzahlvarianten (CNV).

Einfluss der Analysemethode auf CNV-Detektionsrate und -Größenvorstellung

Auflösungsvermögen und Detektionsrate von CNV

Ein Parameter, der die Detektionshäufigkeit von CNV maßgeblich beeinflusst, ist die Mindestgröße, die eine Veränderung aufweisen muss, um von einer bestimmten Methode noch erkannt zu werden. Dieses Auflösungsvermögen wird bei DNA-Arrays v. a. von der Art, Zahl und Verteilung der einzelnen Sondenpunkte auf dem Array bestimmt. Eine frühe Generation von DNA Arrays, welche für die Analyse von CNV verwendet wurden, basierte auf BAC („bacterial artificial chromosome“)-Klonen, deren Inserts das Genom mit einem Raster von ungefähr 1 Mb abdeckten. Trotz dieser relativ geringen Auflösung konnte Iafrate et al. [8] mit dieser Plattform bereits 2004 in 39 gesunden und nicht verwandten Individuen 255 Loci mit variabler DNA-Kopienzahl identifizieren. Davon wurden 24 in mehr als 10% der Probanden gefunden. In weiterer Folge wurden BAC-Klon-Sets genutzt, welche ausgesuchte genomische Hot Spots mit hoher Dichte repräsentierten [20] oder das gesamte Genom in überlappender Weise abdeckten („whole genome tiling path“; [18]). Mit der Verwendung solcher „whole genome tiling path“-Arrays erhöhte sich die Zahl der detektierten CNV in gesunden Individuen sprunghaft [18]. Dabei er-

wiesen sich 2 Eigenschaften von BAC-Arrays für die CNV-Analyse von großem Vorteil, zum einen die damit erzielbare Datenqualität, zum anderen aber v. a. die sehr gute Abdeckung von segmentalen Duplikationen, welche bei der Entstehung von CNV eine besondere Rolle einnehmen (Erläuterungen dazu s. unten). Als klarer Nachteil von BAC-Arrays stellte sich die nur begrenzt steigerbare Auflösung aufgrund der durchschnittlichen Insertgröße eines BAC-Klons (150 kb) heraus. Das erschwerte die genaue Definition der Bruchpunkte von CNV und führte zu einer beträchtlichen Überschätzung der Größe der einzelnen CNV.

Parallel zu den BAC-Arrays wurden schon sehr früh kommerziell erhältliche Oligonukleotidarrays (an anderer Stelle in diesem Heft erläutert) zur Analyse von CNV in gesunden Personen verwendet. Die erste dieser Studien aus dem Jahr 2004 nutzte einen 85.000 Oligonukleotide umfassenden Chip und eine Technik namens ROMA („representational oligonucleotide microarray analysis“), die v. a. darauf abzielt, die Hybridisierungseffizienz durch Reduktion der DNA-Komplexität zu verbessern. Mit der aus heutiger Sicht relativ geringen funktionellen Auflösung wurde ein Durchschnitt von 11-CNV/Person ermittelt [19]. In späteren Studien führte die Nutzung von hochdichten Oligonukleotidarrays mit bis zu 1 Mio. Sondenpunkte zu einer Vervielfachung der gefundenen CNV. Diese beachtliche Zunahme allein durch Steige-

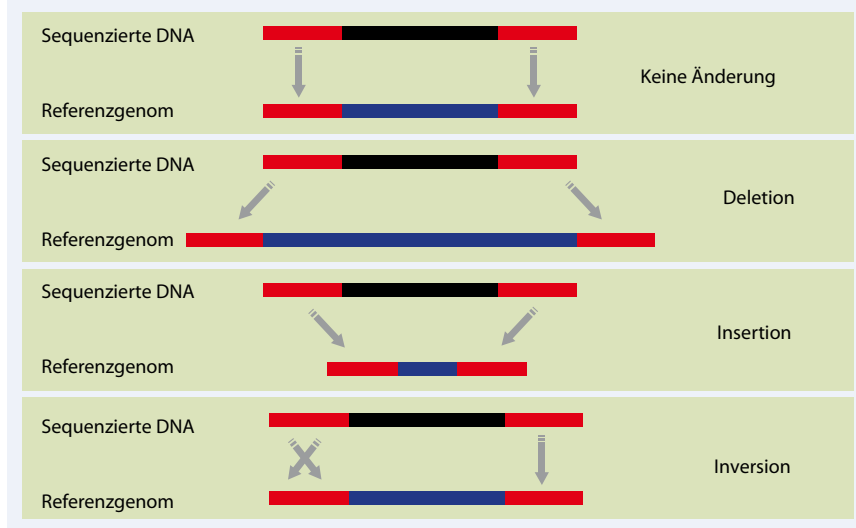


Abb. 1 ▲ Analyse von SV mittels Hochdurchsatzsequenzierung, Enden (rote Balken) von DNA-Fragmenten mit ungefähr bekannter Länge (schwarze Balken); Verortung (graue Pfeile) im humanen Referenzgenom (blaue Balken), gekreuzte Pfeile verkehrte Orientierung eines der beiden Enden bei Inversion, Erläuterung s. Text

rung der Auflösung ist Folge des reziproken Verhältnisses von Größe und Häufigkeit von CNV. Noch deutlicher wird der Zusammenhang von Auflösungsvermögen und Detektionshäufigkeit von CNV, wenn man die Ergebnisse der auf Sequenzierung beruhenden Untersuchungen betrachtet: Die kürzlich publizierten Analysen mit einem gesamtgenomischen Auflösungsvermögen von etwa 3 kb [13] und 8 kb [10] sowie den direkten Vergleich des Celera-Genoms (CR27c) mit dem NCBI Build 35 [9].

Einfluss des Referenzgenoms auf die Detektionsrate von CNV

Array-CGH („comparative genomic hybridization“) beruht auf einem Vergleich einer Test- mit einer Referenz-DNA. Als Ergebnis erhält man relative Unterschiede in der DNA-Kopienzahl, die sich aus den CNV des Test- und des Referenzgenoms zusammensetzen. Die Frequenz sehr häufiger CNV kann durch dieses vergleichende Prinzip unterschätzt werden, weil CNV, die sowohl in der Referenz- als auch in der Test-DNA vorhanden sind, nicht detektierbar sind. Besteht die Referenz aus einem Pool von DNA-Molekülen, werden sich seltene Varianten aufgrund ihres prozentuell geringen Anteils am DNA Gemisch kaum auf die gemittelte Kopienzahl auswirken. Bei Verwendung eines DNA-Pools erscheint die Anzahl von CNV da-

her oft geringer und spiegelt eher die im Testgenom vorhandenen Änderungen wider. Ähnlich ist die Situation auch in jenen Studien, die CNV durch Hybridisierung auf SNP („single nucleotide polymorphism“)-Arrays detektieren und zunächst ohne Referenz-DNA auskommen. Dort werden die Primärdaten in silico mit einem Pool von Referenzdaten abgeglichen, um das durch die oligospezifischen Hybridisierungseigenschaften verursachte Rauschen zu mindern.

Die in **Abb. 1** dargestellte Analyse von Strukturvarianten mittels Fosmid-ESP („end-sequence pairs“) [10] und PEM („paired-end mapping“) [13] beruht auf der Sequenzierung der beiden Enden von DNA-Fragmenten mit ungefähr bekannter Länge (bei Fosmid-Inserts etwa 40 kb und bei PEM etwa 3 kb). Die Enden werden im humanen Referenzgenom verortet. Ist ihr Abstand im Referenzgenom signifikant größer als in der sequenzierten Probe, liegt eine Deletion vor. Bei einer Insertion ist es umgekehrt. Inversionen lassen sich anhand der verkehrten Orientierung eines der beiden Enden erkennen. Auch Methoden, die auf der Sequenzierung von Fosmid-ESP oder dem „next generation sequencing“ von DNA-Fragmenten basieren, vergleichen somit ihre Daten mit einem Referenzgenom, mit teils überraschenden Einsichten, wie eine kürzlich erschienene Arbeit zeigte [10]. In dieser Studie wurden Fosmid-Bibliotheken von

8 HapMap-Individuen hergestellt und die Enden von insgesamt 6,1 Mio. Klonen im humanen Referenzgenom (Build 35) verortet. Dabei entdeckten die Autoren 525 Sequenzcontigs mit einer Größe von bis zu 130 kb, welche keine Entsprechung im Referenzgenom hatten. Array-CGH mittels eines spezifischen Oligonukleotidarrays, welcher diese Regionen mit hoher Dichte abdeckte, zeigte, dass 40% dieser Regionen variable Kopienzahlen innerhalb der 8 untersuchten Individuen aufwiesen [10]. Diese Ergebnisse legen nahe, dass Array-CGH und verwandte Methoden bislang viele CNV gar nicht entdecken konnten, da die Array-Designs auf jenem offensichtlich noch immer unvollständigen Referenzgenom basieren, welches einige polymorphe Regionen nicht repräsentiert.

Größe von CNV und CNV-Regionen

Abgesehen von der Häufigkeit ist v. a. die präzise Erfassung der Größe einer Kopienzahlvariante wichtig für eine Abschätzung der gesamtgenomischen Variabilität. Im bislang umfassendsten Vergleich von „whole genome tiling path“-BAC und Oligonukleotidarrays auf Grundlage von 270 HapMap-Individuen reduzierte sich mit dem 500k-Early-Access(EA)-Array von Affymetrix der CNV-Größenmedian von 228 kb (Mittelwert: 341 kb) bei BAC-Arrays auf 81 kb (Mittelwert: 206 kb). Dementsprechend nahm damit der Durchschnitt des variablen Genoms pro Individuum von 24 Mb (BAC-Array) auf 5 Mb (Affymetrix) ab [18]. Um den variablen Anteil des humanen Genoms besser definieren zu können, fassten die Autoren überlappende CNV der 270 Individuen zu Regionen zusammen (CNVR). Die Kombination beider Datensets ergab insgesamt 1447 CNVR, welche 12% (360 Mb) des human Genoms repräsentierten [18]. Das ist weniger als z. B. jene 18%, die auf Basis der Datenbankeinträge in der Database of Genomic Variants (<http://projects.tcag.ca/variation/>) errechnet wurden [17], aber möglicherweise trotzdem noch zu hoch gegriffen [16]. Wie bereits an anderer Stelle erwähnt, neigen BAC-Arrays dazu, CNV größer darzustellen, als sie in Wirklichkeit sind, und ein großer Teil der

CNV-Studien wurde bislang mit BAC-Arrays durchgeführt.

Zum jetzigen Zeitpunkt ist es daher äußerst schwierig, das Ausmaß genomischer Variabilität abzuschätzen. Während verbesserte Technologien die Detektion einer großen Zahl bislang unbekannter, meist kleiner Varianten ermöglichen, reduzierte sich gleichzeitig durch die erhöhte Präzision bei deren Identifizierung die Größe vieler bekannter CNV.

Entstehung von strukturellen Varianten

Reparatur von DNA-Doppelstrangbrüchen und CNV

Die am häufigsten im Zusammenhang mit der Entstehung von CNV diskutierten Mechanismen sind Retrotransposition, die nichtallelische homologe Rekombination („non-allelic homologous recombination“; NAHR) und die Verbindung von nichthomologen Enden („non-homologous end joining“; NHEJ). NAHR und NHEJ sind beides Möglichkeiten zur Reparatur von DNA-Doppelstrangbrüchen. Die Wahrscheinlichkeit von NAHR steigt mit der Anwesenheit, Länge und Sequenzsimilarität von segmentalen Duplikationen. NAHR kann innerhalb einer Chromatide auftreten und führt dann immer zu einer Deletion. Die Rekombination zwischen 2 Chromatiden oder Chromosomen resultiert in einer Deletion und der reziproken Duplikation. Bei invertierter Orientierung der Duplikons kommt es zur Inversion [21]. Segmentale Duplikationen sind nicht gleichmäßig im Genom verteilt, sondern akkumulieren in bestimmten Regionen. Sequenzen, welche von segmentalen Duplikationen flankiert werden, haben ein erhöhtes Risiko, eine DNA-Kopienzahländerung durch NAHR zu erfahren. Das passt gut zur Beobachtung, dass auch CNV nicht gleichmäßig im Genom verteilt sind, sondern dass es offensichtlich Regionen mit erhöhter Instabilität gibt [2].

NAHR ist auch Ursache für einen großen Teil jener rekurrenten CNV, die zu so genannten „genomic disorders“ [15] führen, wie dem Williams-Beuren- oder dem 22q11-Deletions-/Duplikations-Syndrom. Dieser Umstand verdeutlicht, dass NAHR

für einen Organismus mit einem hohen Anteil an repetitiver DNA und segmentalen Duplikationen ein beträchtliches Risiko darstellt. Deshalb wurde über Mechanismen spekuliert, die NAHR aktiv unterdrücken (Zusammenfassung in [4]). Tatsächlich ist NHEJ der vorherrschende Mechanismus zur Reparatur von Doppelstrangbrüchen in vielen Eukaryonten, da NHEJ, unter Ausnutzung von Mikrohomologien von wenigen Nukleotiden Länge, die Bruchenden meist sehr präzise zusammensetzt. Oft kommt es aber dabei auch zum Verlust einiger weniger Basenpaare.

Sequenzierung von Bruchpunkten

Viele Aussagen zu den Entstehungsmechanismen von CNV und deren Häufigkeit sind spekulativ. Wenn beide chromosomalen Bruchpunkte mittels Array-CGH innerhalb von homologen segmentalen Duplikationen lokalisiert werden konnten, wurde das bislang zumeist als Hinweis für NAHR interpretiert. Mittlerweile ist bekannt, dass das nicht immer zutreffen muss. Umgekehrt wurde das Fehlen von segmentalen Duplikationen an einem oder beiden Bruchpunkten als Argument zum Ausschluss von NAHR genommen. Dies kann ebenso trügerisch sein, da segmentale Duplikationen nicht nur für die Array-CGH problematisch sind, sondern auch für das Assemblieren von Sequenzen. Es muss daher davon ausgegangen werden, dass viele segmentale Duplikationen noch gar nicht in der aktuellen Sequenzversion des humanen Genoms präsent sind. Erst die Sequenzierung von Bruchpunkten ermöglicht eine fundierte Aussage über die Entstehungsmechanismen von CNV.

Kürzlich erschienen 3 Studien, welche zusammen die Sequenz von 486 Bruchpunkten bestimmten [10, 13, 16]. Während Kidd et al. [10] NAHR als häufigste wahrscheinliche Ursache von SV ausmachten (in 47,5% der SV), war der Hauptentstehungsmechanismus bei den beiden anderen Studien NHEJ (56% bzw. bis zu 91%). Diese Divergenz bezüglich der Häufigkeit der verschiedenen Mechanismen beruht wahrscheinlich auf der methodisch bedingten tendenziellen Selektion der se-

Zusammenfassung · Abstract

medgen 2008 · 20:401–406
DOI 10.1007/s11825-008-0137-4
© Springer Medizin Verlag 2008

R. Ullmann

Strukturelle Genomvarianten. Ausmaß, Entstehung und phänotypische Konsequenzen

Zusammenfassung

Unter strukturellen Genomvarianten (SV) versteht man chromosomale Veränderungen größer als 1 kb. Obwohl es dank des technischen Fortschritts immer besser gelingt, die qualitativen und quantitativen Eigenschaften von SV zu definieren, sind deren phänotypischen Konsequenzen meist nur unzureichend verstanden. Die Unterscheidung zwischen neutraler Variante, prädisponierendem Faktor und krankheitsverursachender Veränderung stellt heute eine der großen Herausforderungen in der humangenetischen Diagnostik dar.

Schlüsselwörter

Array-CGH · Kopienzahlvarianten · Phänotypische Variabilität · Prädisposition · Interpretation von CNV

Structural genome variants. Extent, genesis and phenotypic consequences

Abstract

Structural variants (SV) are defined as chromosomal changes larger than 1 kb. Although technical progress has enabled improved characterization of the qualitative and quantitative features of SV, their phenotypic consequences remain poorly understood. Distinguishing between a neutral variant, a predisposing factor and a disease-causing aberration represents one of the major challenges in today's human genetic diagnostics.

Keywords

Array-CGH · Copy number variants · Phenotypic variability · Predisposition · Interpretation of CNV

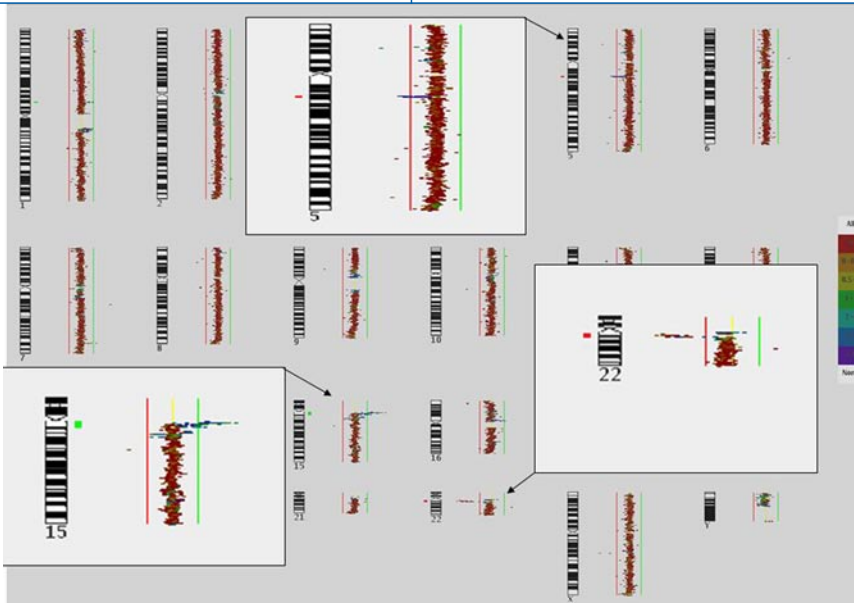


Abb. 2 ▲ Interpretation des CNV-Resultats einer BAC-Array-CGH-Analyse eines Patienten mit Schizophrenie, Cy3/Cy5-log₂ Ratios jedes einzelnen Klon entsprechend ihrer chromosomalen Lokalisation dargestellt (links negative, rechts positive Ratios), Färbung der einzelnen Klone Gehalt an segmentalen Duplikationen, Farbkodierung s. Farbbalken am rechten Bildrand, rot keine segmentalen Duplikationen. (Nach [11])

quenzierten Bruchpunkte in den unterschiedlichen Studien.

Interessanterweise beschrieben Perry et al. [16], dass in 9 der sequenzierten Deletionsbruchpunkte z. T. invertierte Insertionen mit Homologie zum deletierten Abschnitt gefunden wurden. Dieser Befund könnte auf andere, mit der DNA-Replikation assoziierte Mechanismen, hindeuten (z. B. „replication fork stalling and template switching“, FoSTeS [14]).

Häufigkeit der De-novo-Entstehung von CNV

Ein großer Teil der genomischen Varianten scheint konserviert zu sein. SNP-Daten zeigen ein starkes Linkage-Disäquilibrium in vielen CNV und deren flankierenden Sequenzen, und eine vergleichende Analyse verschiedener Ethnien konnte nachweisen, dass diese CNV mit denselben SNP-Allelen assoziiert waren, was für einen gemeinsamen Ursprung spricht ([21] und darin enthaltene Referenzen).

Über die De-novo-Rate von CNV ist noch wenig bekannt. Eine Studie untersuchte das Auftreten von NAHR in Regionen, welche in die Entstehung ausgesuchter kongenitaler Krankheiten mit rekurrenten Bruchpunkten involviert sind.

Aufgrund ihrer Untersuchungen in Spermien und Lymphozyten folgerten die Autoren, dass

- NAHR häufig Ursache von De-novo-CNV ist,
- dass die Rekombinationsereignisse auf die Meiose beschränkt sind und
- dass Deletionen doppelt so häufig wie Duplikationen vorkommen [22]; ein Hinweis auf die Dominanz intrachromatidaler Rekombination (s. oben).

Tatsächlich ist aber die heute angenommene Rate von Duplikationen und Deletionen bei CNV ungefähr gleich, was die Autoren mit der Möglichkeit negativer Selektion und/oder anderen Entstehungsmechanismen von CNV begründeten, welche vorwiegend Duplikationen verursachen [22].

Für die Abschätzung der De-novo-Rate von CNV in somatischen Zellen müssen 2 Aspekte berücksichtigt werden:

1. wie oft entstehen CNV im Rahmen der somatischen Zellteilung und
2. welche davon entkommen der negativen Selektion und zeigen klonale Expansion?

Von Studien zu krankheitsassoziierten CNV ist bekannt, dass mitotische Rekombination zu somatischen Mosaiken füh-

ren kann (Beispiele dazu in [21]) und v. a. in der Tumorigenese eine wichtige Rolle spielt. Schnell wird klar, dass die derzeit üblichen Strategien zur genomweiten Detektion von CNV De-novo-Ereignisse in somatischen Zellen nicht entsprechend erfassen können. Erst umfangreiche Analysen auf dem Niveau von Einzelzellen werden hier Klarheit bringen können.

Phänotypische Konsequenzen von strukturellen Varianten

Selektion von CNV

Die ungleichmäßige Verteilung von CNV im Genom und die Unterrepräsentation von funktionellen Sequenzen lassen vermuten, dass sich ein großer Teil der CNV mit phänotypischen Auswirkungen aufgrund negativer Selektion nicht durchsetzen kann [1, 18]. Dabei scheinen Duplikationen eher toleriert zu werden als Deletionen. Letztere sind im Schnitt auch beträchtlich kleiner und haben im Vergleich auch weniger Gene [1, 2, 18]. (Kontrovers wird das Verhältnis von CNV und Gendichte von Cooper et al. [2] diskutiert, die eine Häufung von mit segmentalen Duplikationen assoziierten CNV in genreichen Regionen feststellten). Übereinstimmung herrscht bei der Aussage, dass besonders funktionelle Kategorien von Genen in CNV angereichert sind, die die Reaktion auf Umwelteinflüsse beeinflussen und damit die Adaptionsfähigkeit fördern, wie Gene des Geruchssinns oder der Immunantwort.

Genotyp-Phänotyp-Korrelation

Für die Abschätzung möglicher phänotypischer Konsequenzen einer CNV ist eine genaue Charakterisierung unabdingbar. Denn scheinbar identische CNV in verschiedenen Personen können bei einer Analyse mit präziseren Methoden als unabhängig entstandene Veränderungen mit unterschiedlicher Größe und Relevanz erkannt werden. Eindrucksvoll konnte dies kürzlich mittels Sequenzierung des variablen *SIRPB1*-Locus demonstriert werden. Nur eine der beiden entdeckten Deletionsvarianten betraf ein Exon des Gens, während sich die andere Variante auf intronische Sequenzen beschränkte [10].

Aufgrund der Präsenz von segmentalen Duplikationen an den Bruchpunkten wäre die Unterschiedlichkeit der Deletionen am *SIRPB1*-Locus bei einer Array-CGH-Analyse wahrscheinlich nicht aufgefallen [10].

CNV als prädisponierende Faktoren für Krankheiten

CNV können direkt über die Änderung der Gendosis für Krankheiten prädisponieren oder die Voraussetzungen für weitere, krankheitsverursachende chromosomale Rearrangements schaffen. Eine Reihe von Inversionen und CNV wurden mit der Entstehung von rekurrenten Chromosomenanomalien in Verbindung gebracht, wie dem Sotos- und dem Williams-Beuren-Syndrom (diese und weitere Beispiele in [3, 6, 21]). Es ist sehr wahrscheinlich, dass sich insbesondere die Zahl der erkannten, für chromosomale Veränderungen prädisponierenden Inversionen weiter erhöhen wird. Eine Ahnung davon gibt die bereits oben beschriebene Se-

quenzanalyse von 8 Fosmid-Bibliotheken, in deren Verlauf 3 Inversionen gefunden wurden, die mit kürzlich beschriebenen Mikrodeletionssyndromen örtlich zusammenfallen [10].

Unterschiede in der Gendosis können mit einer veränderten Verstoffwechslung von Medikamenten einhergehen. (z. B. Cytochrom P450), das Immunsystem beeinflussen (z. B. α - und β -Defensin-Cluster), das Krebsrisiko erhöhen (z. B. *GSTT1*) oder die Ausprägung von Krankheiten modifizieren (z. B. spinale Muskeldystrophie und *SMN2*; diese und weitere Beispiele finden sich in Referenz [21] zusammengefasst). Die mittlerweile wohl bekannteste Kopienzahlvariante in diesem Zusammenhang ist jene, die das Gen *CCL3L1* betrifft und die Dosis eines HIV-1 supprimierenden Chemokins verändert [7].

Interpretation von CNV in der humangenetischen Beratung

Es ist kein Zufall, dass der Begriff „CNV“ jeglichen Hinweis auf Häufigkeit und Krankheitsrelevanz vermeidet. Dieser Umstand spiegelt vielmehr die Unsicherheit wider, die mit der Interpretation der Mehrheit von CNV einhergeht. Folgerichtig – aber oft verwirrend – wird der Begriff für Varianten in der gesunden Bevölkerung genauso verwendet wie für Veränderungen mit vermutetem oder eindeutigen Krankheitsbezug (Letztere werden an anderer Stelle in diesem Heft ausführlicher behandelt). Ein Problem dabei ist, dass selbst CNV mit gleicher Größe aufgrund der oft mangelnden Korrelation von DNA-Kopienzahländerung und der Expression der darin enthaltenen Gene massiv unterschiedliche Auswirkungen haben können. Für die Abschätzung der möglichen phänotypischen Konsequenzen dürfte man eigentlich eine CNV nicht isoliert betrachten, sondern müsste sie im Kontext des genetischen

Hier steht eine Anzeige.

und epigenetischen Hintergrunds der jeweiligen Person beurteilen. Dies ist allerdings mit dem derzeitigen Wissensstand kaum möglich, zu komplex und ungenügend definiert ist das Zusammenspiel der verschiedenen Faktoren.

Ein Beispiel dazu gibt **Abb. 2**, die das Resultat einer BAC-Array-CGH-Analyse eines Patienten mit Schizophrenie zeigt, wobei die $Cy3/Cy5$ -log₂-Ratios jedes einzelnen Klons entsprechend ihrer chromosomalen Lokalisation dargestellt sind. Abgesehen von der typischen 22q11-Deletion sieht man weitere CNV, die einen hohen Anteil an segmentalen Duplikationen aufweisen. Inwieweit diese den Phänotyp beeinflussen, ist unklar.

Ein weiteres Beispiel hierzu ist die rekurrente Deletion auf 1q21, die mit dem TAR („thrombocytopenia absent radius“-) Syndrom (MIM 27400) assoziiert ist. Die Veränderung wurde in 30 von 30 Patienten gefunden, aber auch in einem Teil der gesunden Verwandtschaft [12]. Die Aberration ist ohne Zweifel für das Auftreten der Krankheit notwendig, aber nur in einem bestimmten genetischen und epigenetischen Umfeld kommt es zu deren Ausprägung. Gleichzeitig verdeutlicht dieses Beispiel, dass die Präsenz einer Veränderung in gesunden Eltern oder entsprechenden Datenbanken nur noch bedingt als Kriterium zum Ausschluss von Krankheitsassoziation dienen kann. Angesichts dieser Unsicherheiten stellt sich die Ergebnismitteilung einer Array-CGH-Analyse komplizierter dar als erwartet.

Fazit

Innerhalb der letzten 4 Jahre hat unsere Vorstellung von chromosomalen Veränderungen einen beachtlichen Wandel erfahren. Waren zur Hochblüte der klassischen Zytogenetik identifizierte Veränderungen meist kausal mit einem Krankheitsbild assoziiert, hat die Interpretation von Chromosomenaberrationen durch die Masse neu entdeckter submikroskopischer Veränderungen an Komplexität gewonnen. CNV können in Abhängigkeit vom jeweiligen genetischen und epigenetischen Hintergrund den Phänotyp beeinflussen oder auch nicht. Um die Konsequenzen von CNV besser voraussagen zu können, sind umfang-

reichere Studien nötig. Dass dabei das zu untersuchende Kollektiv in seiner Größe den mit der Zahl der CNV steigenden Freiheitsgraden genügen muss, stellt eine logistische, bioinformatische und v. a. finanzielle Herausforderung dar.

Korrespondenzadresse

R. Ullmann

Max Planck Institut für Molekulare Genetik,
Inhnestraße 73, 14195 Berlin
ullmann@molgen.mpg.de

Danksagung. Ich möchte mich recht herzlich bei meinen Kollegen Vera Kalscheuer, Andreas Kuss und Andreas Tzschach für die hilfreichen Kommentare bedanken. Des Weiteren möchte ich mich bei all jenen Kollegen entschuldigen, deren Originalarbeiten ich aufgrund des limitierten Platzes nicht zitieren konnte.

Interessenkonflikt. Der korrespondierende Autor gibt an, dass kein Interessenkonflikt besteht.

Literatur

- Conrad DF, Hurler ME (2007) The population genetics of structural variation. *Nat Genet [Suppl 7]* 39: 30–36
- Cooper GM, Nickerson DA, Eichler EE (2007) Mutational and selective effects on copy-number variants in the human genome. *Nat Genet [Suppl 7]* 39: S22–29
- Cusco I, Corominas R, Bayes M et al. (2008) Copy number variation at the 7q11.23 segmental duplications is a susceptibility factor for the Williams-Beuren syndrome deletion. *Genome Res* 18: 683–694
- Erdogan F, Chen W, Kirchhoff M et al. (2006) Impact of low copy repeats on the generation of balanced and unbalanced chromosomal aberrations in mental retardation. *Cytogenet Genome Res* 115: 247–253
- Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. *Nat Rev Genet* 7: 85–97
- Freeman JL, Perry GH, Feuk L et al. (2006) Copy number variation: new insights in genome diversity. *Genome Res* 16: 949–961
- Gonzalez E, Kulkarni H, Bolivar H et al. (2005) The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 307: 1434–1440
- lafrate AJ, Feuk L, Rivera MN et al. (2004) Detection of large-scale variation in the human genome. *Nat Genet* 36: 949–951
- Khaja R, Zhang J, MacDonald JR et al. (2006) Genome assembly comparison identifies structural variants in the human genome. *Nat Genet* 38: 1413–1418
- Kidd JM, Cooper GM, Donahue WF et al. (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature* 453: 56–64
- Kirov G, Gumus D, Chen W et al. (2008) Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 17: 458–465
- Klopocki E, Schulze H, Strauss G et al. (2007) Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 80: 232–240
- Korbel JO, Urban AE, Affourtit JP et al. (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318: 420–426
- Lee JA, Carvalho CM, Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 131: 1235–1247
- Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 14: 417–422
- Perry GH, Ben-Dor A, Tsalenko A et al. (2008) The fine-scale and complex architecture of human copy-number variation. *Am J Hum Genet* 82: 685–695
- Pinto D, Marshall C, Feuk L et al. (2007) Copy-number variation in control population cohorts. *Hum Mol Genet* 16: 168–173
- Redon R, Ishikawa S, Fitch KR et al. (2006) Global variation in copy number in the human genome. *Nature* 444: 444–454
- Sebat J, Lakshmi B, Troge J et al. (2004) Large-scale copy number polymorphism in the human genome. *Science* 305: 525–528
- Sharp AJ, Locke DP, McGrath SD et al. (2005) Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* 77: 78–88
- Sharp AJ, Cheng Z, Eichler EE (2006) Structural variation of the human genome. *Annu Rev Genomics Hum Genet* 7: 407–442
- Turner DJ, Miretti M, Rajan D et al. (2008) Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet* 40: 90–95