

dbCRID: a database of chromosomal rearrangements in human diseases

Fanlou Kong¹, Jing Zhu², Jun Wu¹, Jianjian Peng¹, Ying Wang¹, Qing Wang¹, Songbin Fu², Li-Lian Yuan³ and Tongbin Li^{1,*}

¹Biolead.org Research Group, LC Sciences, Houston, TX 77054, USA, ²Laboratory of Medical Genetics, Harbin Medical University, Harbin, Heilongjiang, 150081, China and ³Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA

Received August 15, 2010; Revised September 25, 2010; Accepted October 11, 2010

ABSTRACT

Chromosomal rearrangement (CR) events result from abnormal breaking and rejoining of the DNA molecules, or from crossing-over between repetitive DNA sequences, and they are involved in many tumor and non-tumor diseases. Investigations of disease-associated CR events can not only lead to important discoveries about DNA breakage and repair mechanisms, but also offer important clues about the pathologic causes and the diagnostic/therapeutic targets of these diseases. We have developed a database of Chromosomal Rearrangements In Diseases (dbCRID, <http://dbCRID.biolead.org>), a comprehensive database of human CR events and their associated diseases. For each reported CR event, dbCRID documents the type of the event, the disease or symptoms associated, and—when possible—detailed information about the CR event including precise breakpoint positions, junction sequences, genes and gene regions disrupted and experimental techniques applied to discover/analyze the CR event. With 2643 records of disease-associated CR events curated from 1172 original studies, dbCRID is a comprehensive and dynamic resource useful for studying DNA breakage and repair mechanisms, and for analyzing the genetic basis of human tumor and non-tumor diseases.

INTRODUCTION

A chromosomal rearrangement (CR) event occurs as a consequence of double-strand breaks (DSBs) of the

DNA, followed by abnormal rejoining of the non-homologous ends, producing a new chromosomal arrangement (1). Alternatively, a CR event can result from crossing-over between repetitive DNA sequences (2). CR events may lead to disruption of genes and other functional structures. They have been implicated in many tumor and non-tumor human diseases (3,4), and are frequently examined in clinical diagnosis, treatment and prognosis (5–7). Analyses of junction sequences at CR breakpoints have led to the discovery of multiple short sequence motifs, including polypurine/polypyrimidine sequences (8), Ig heptamers (9), translin-binding sites (10) and minisatellite core sequences (11). These sequence motifs may play important roles in DNA breakage and repair mechanisms (12). CR events can be categorized into the following seven types: deletion, duplication, insertion, inversion, reciprocal translocation, ring chromosome and translocation.

Cytogenetics techniques commonly applied to identify and analyze CR events include high-resolution chromosome banding (13), fluorescence *in situ* hybridization (FISH) (14), spectral karyotyping (SKY) (15), comparative genomic hybridization (CGH) (16), and polymerase chain reaction (PCR). In recent years, advances in DNA sequencing technologies and the completion of the human genome have led to accelerated accumulation of experimentally identified CR events. Information about CR events and their associated diseases and/or clinical symptoms can provide important clues about the chromosomal breakage and DNA repair mechanisms. Moreover, this information, if fully exploited, can bring about improved understanding about how CR events lead to these important diseases, as well as how other factors (e.g. features in junction sequences) influence their occurrence and progression, which in turn will lead to improved prevention of and clinical intervention for these diseases.

*To whom correspondence should be addressed. Tel: +1 651 340 1974; Fax: +1 651 315 8196; Email: tolib@biolead.org

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

© The Author(s) 2010. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.5>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a first step towards this goal, we have developed dbCRID, a manually curated database of human CR events and their associated diseases. dbCRID includes 2643 individually curated entries of experimentally tested CR events, their associated diseases and/or clinical symptoms and detailed information about the CR events, including the precise locations of the breakpoints, genes involved, junction sequences, as well as the experimental techniques applied, and links to the original studies. Curated from 1172 original studies, dbCRID is a comprehensive resource of human CR events, breakpoints/junctions and associated diseases.

UTILITY

Data content

All data were manually extracted from original, peer-reviewed published studies about CR events and associated human diseases. A consistent procedure was established to identify studies likely to contain useful data by examining the abstracts and full-text articles through the PubMed database. For each reported CR event, we documented the CR type, breakpoint information, junction sequences, genes involved in the CR event, names of the diseases (tumor or non-tumor) and/or clinical symptoms, as well as the experimental technique applied in the study. The breakpoint information includes the position of the breakpoint, disrupted gene (when available) and gene region (which exon or intron), as well as a 'precision code', which describes the level of precision determined for the breakpoint. The precision code 'A' indicates that the CR event is determined at the highest precision—to an individual nucleotide level. Precision

code 'B' indicates that the CR event is determined to an individual gene level. Precision code 'C' indicates the lowest precision: the breakpoints of CR events with precision code 'C' can only be localized to a broad region within a chromosome.

For all events with precision code 'A', the exact breakpoint positions were consistently mapped to the human genome hg19 (NCBI Build 37.1 Feb 2009). For events with precision code 'A' or 'B', the chromosomal locations of the breakpoints (e.g. 2p24.1) and precise locations of involved genes (consistently identified using RefSeq accessions) were mapped to the same build of the human genome using the Ensembl BLAST tool (17).

Junction sequences at CR breakpoints may provide important clues as to how the CR events occurred (18), and analyses of these junction sequences have led to interesting discoveries about the DNA repair mechanisms (19). We have carefully documented the junction sequences, whenever possible. Frequently, the junction sequence discovered in the original study does not perfectly match the joining of the two non-homologous sequences prior to the CR event. We labeled these junction sequences as 'reported' in dbCRID. If the junction sequence discovered is the same as the joining of the non-homologous sequences, we labeled it as 'confirmed'. For certain CR events with precision code 'A', the junction sequences were not explicitly provided in the original study. We inferred the junction sequences by joining 100 nt from both sides of the junction following the CR events, and labeled these junction sequences as 'inferred'.

The current release of dbCRID hosts the records of 2643 CR events curated from 1172 original studies. Detailed statistics of the records is summarized in Figure 1.

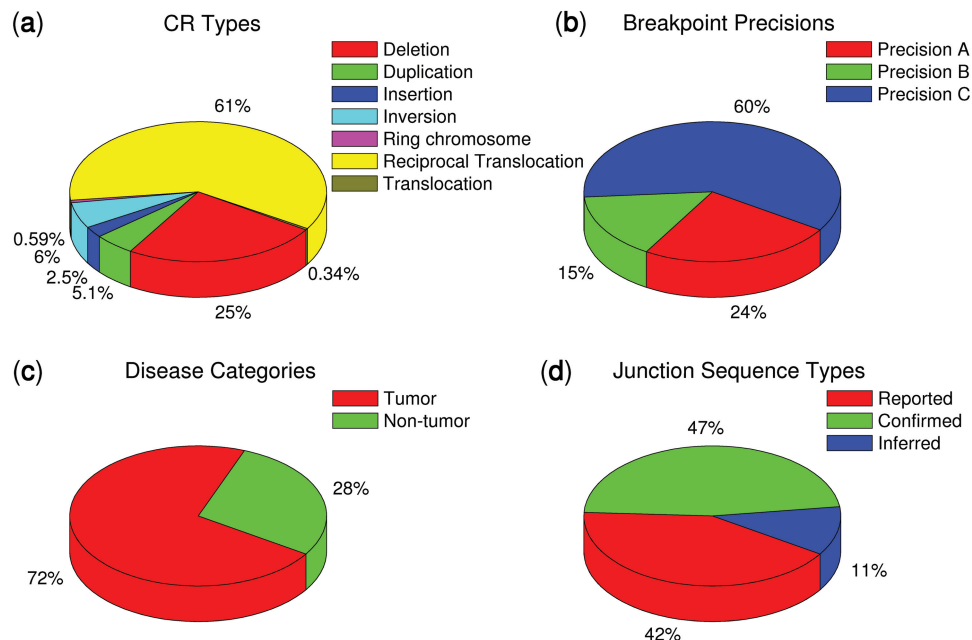


Figure 1. Statistical summaries of CR events documented in dbCRID. (a) Summary of CR types. (b) Summary for breakpoint precisions. (c) Summary for disease categories. (d) Summary for junction sequence types.

Web interface

The web interface of dbCRID database can be accessed through the URL <http://dbCRID.biobase.org/>. In the main page, the user can click any of the human chromosomes underneath the introduction text to obtain a list of all CR events related to the chromosome. This list can be sorted by Event ID, Karyotype, Precision code, Case ID or Disease name by clicking any of these attribute names. The user can click an Event ID to display the ‘detailed information’ about this event. On the top of the ‘detailed information’ page is the name of the disease associated with this event, the DOID of the disease, and a link to the corresponding disease in the Ontology Lookup Service (OLS) (20). The user can obtain a list of all events associated with this disease by clicking the link labeled ‘All events related to this disease’. The information about the CR event, including the precision code, event type, chromosome and karyotype and breakpoint information is displayed below. The junction sequences (when available), the types of the junction sequences (‘confirmed’, ‘reported’ or ‘inferred’), the experimental techniques applied to identify/analyze the CR events, as well as the reference to the original study are provided below. On the bottom of the ‘detailed information’ page is a reduced-size genome browser image. The user can click this image to access the genome browser view of the CR event (described below).

The user can click the ‘Search’ tab below the dbCRID project logo to reach the ‘event search page’. This page allows the user to search CR events by chromosome, chromosome arm, CR type, by the gene involved (using a RefSeq accession), and by the name of an author or a keyword of the original study. The list of CR events returned by the search function is of the same format as the CR event list for individual chromosomes (described above), with links to ‘detailed information’ pages of the events listed.

The ‘Disease List’ tab below the dbCRID project logo allows the user to view the entire list of diseases associated with any CR events documented in dbCRID, displayed in a hierarchical order based on the DOID structure. The user can click any of the disease names to view a list of cases and related events associated with the disease of interest.

The ‘Genome Browser’ tab below the project logo allows the user to view the CR events in the browser mode. The user can specify the chromosome and region of interest, or select a region of interest using by mouse operations, and the browser will display all CR events (or the ones of selected types) occurring in the selected regions. Mouse-over a region with a CR event in the browser will trigger the display of a schematic graph that intuitively presents the CR event. Information about known RefSeq genes in the region is also displayed. Optionally, the user can upload his/her own track files that contain other types of data for comparison purposes.

The ‘Download’ tab below the project logo points to a data download page, where the current release of the entire dataset can be downloaded as a tab-delimited text file.

The ‘Glossary’ tab below the project logo points to the glossary page where terms and abbreviations used in the web site are explained.

Data access

The dbCRID web site is publicly accessible through the URL <http://dbCRID.biobase.org/>. The current released dataset can be downloaded from the web site. Additional requests can be made by emailing to dbCRID@biobase.org.

Implementation

dbCRID was constructed as a relational database implemented with MySQL on a Fedora Core 10 Linux system. The front-end web interface was implemented as a PHP project (enhanced with JavaScript) running under Apache 2.0. The Genome Browser function was implemented using Gbrowse 2.0 (21).

COMPARISON WITH RELATED DATABASES

dbCRID is a new, comprehensive database of human CR events and associated diseases (both tumor and non-tumor) with detailed documentation of the CR events—including precise breakpoint positions and junction sequences—for all seven types of CR events. A comparison with related databases demonstrates the unique features of this new resource (Table 1).

Atlas of Genetics and Cytogenetics in Oncology and Haematology (22) is an integrative database of human cancers, cancer-prone diseases and involved genes. For a given cancer type, the breakpoints of related CR events can be inferred from the hybrid genes or chromosomal aberrations listed. ChimerDB (1.0 & 2.0) is a knowledgebase for fusion genes (23). Transcript sequences were obtained and integrated from Genbank and other public resources. Fusion genes resulting from interchromosomal translocations, intrachromosomal deletions or inversions are classified in ChimerDB 2.0. The breakpoints on the genes can be viewed in the custom genome browser viewer. CytoD 1.0 (Cytogenetics Database, <http://www.changbioscience.com/cytogenetics/cyto.htm>) is a large database of literature abstracts of reported cytogenetic abnormalities. It does not provide detailed information about the breakpoints or the junctions, nor does it provide any information about related diseases or experimental techniques applied. DARCO (<https://www1.hgu.mrc.ac.uk/Softdata/Translocation/>) is an unpublished database of CR events and associated abnormal phenotypes that focuses mainly on translocation events. Breakpoint locations are provided on chromosomal region level (equivalent to precision code ‘C’ in dbCRID), and no information about disrupted genes, junction sequences or experimental techniques applied is provided. Decipher (24) is a database of chromosomal imbalance, inversion and translocation events (derived from array CGH data) and associated symptoms/phenotypes. The chromosomal locations of submicroscopic copy number changes—mainly microdeletion and microduplications—as well as the genes involved can be

Table 1. Comparison between dbCRID and related databases

Database	URL	Disease type covered		CR types covered	Breakpoint positions provided	Junction sequences provided	Methods applied documented	Genes involved in breakpoints documented	Genome Browser view provided	Number of events documented
		Tumor	Non-tumor							
Atlas of Genetics and Cytogenetics in Oncology and Haematology	http://www.infobiogen.fr/services/chromcancer	Yes	Yes ^a	Six types: deletion, duplication, insertion, inversion, reciprocal translocation and translocation	No	No	No	No	No	719
ChimerDB	http://eresb.ewha.ac.kr/fusiongene	Yes	No	Four types: deletion, duplication, reciprocal translocation and translocation	Nucleotide level	Yes	No	Yes	Yes	_b
CytoD	http://www.changbioscience.com/cytogenetics/cyto.htm	Yes	Yes	All seven types	No	No	No	No	No	_c
DACRO	https://www1.hgu.mrc.ac.uk/Softdata/Translocation/	Yes	Yes	Four types: insertion, inversion, reciprocal translocation and translocation	Chromosome level	No	No	No	No	1087
Decipher	https://decipher.sanger.ac.uk	No	Yes	Six types: deletion, duplication, insertion, inversion, reciprocal translocation, translocation	Nucleotide level	No	Array CGH	No	Yes	_d
HYBRIDdb	http://www.primare.or.kr/hybriddb/	Yes	No	Two types: reciprocal translocation, translocation	Nucleotide level	Yes	No	Yes	No	_e
Mitelman	http://egap.nci.nih.gov/Chromosomes/Mitelman	Yes	No	All seven types	No	No	No	No	No	2044 ^f
SKY/M-FISH and CGH	http://www.ncbi.nlm.nih.gov/sky/	Yes	No	Five types: deletion, duplication, insertion, reciprocal translocation, translocation	No	No	SKY/FISH/CGH	No	No	1700
TICdb	http://www.unav.es/genetica/TICdb/	Yes	No	Reciprocal translocation	Nucleotide level	Yes	No	Yes	No	795
dbCRID	http://dbcrd.biolead.org	Yes	Yes	All seven types	Three levels: nu-cleotide, gene and chromosome	Yes	Yes	Yes	Yes	2643

^aCancer-prone diseases only.

^bA total of 11 747 gene fusion events are documented, however, their association with diseases are not known.

^cA total of 11 570 records of PubMed abstracts are documented, but no detailed information about the CR events are provided.

^dA total of 4200 array CGH reports associated with symptoms/phenotypes are documented.

^eA total of 1060 gene fusion events are documented, however, their association with diseases are not known.

^fThese events are associated with cancer characteristics.

visualized with the integrated Ensembl genome browser. HYBRIDdb (25), like ChimerDB, is a database of hybrid genes in human genome. More than 1000 genes pairs resulting from chromosome translocation were identified, and the breakpoint on each gene was located. The Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>) is a popular database of chromosomal aberrations in human tumors and associated genes. Karyotype abnormalities associated with a particular tumor can be searched through the Cases Searchers. Gene rearrangements resulting from cytogenetic aberrations can be searched through the Molecular Biology Associations Searcher. The Recurrent Chromosome Aberrations Searcher allows the searching of recurrent aberrations in specific tumors. Moreover, the Clinical Associations Searcher allows the searching of associations between cytogenetic aberrations and/or gene rearrangements and tumor characteristics. Currently, 2044 CR events clinically associated with cancers are documented in the Mitelman Database. The SKY/M-FISH & CGH database (26) is a database of molecular cytogenetics data from SKY/M-FISH or CGH studies in human cancers. As with the Mitelman Database, it focuses on CR events associated with tumors only, and the detailed information about the breakpoints (e.g. their precise locations, disrupted genes, or junction sequences) is not provided. TICdb (27) is database of fusion genes resulting from reciprocal translocation events associated with tumors. It documents the precise location of each breakpoint inside a fusion gene, but does not map the breakpoints to the genome.

In summary, each of these existing databases has its distinct strengths and values. Compared to these resources, dbCRID is the only database that (i) documents CR events associated with both tumor and non-tumor diseases, (ii) covers all seven types of CR events, (iii) provides breakpoint position information in three different precision levels, (iv) provides detailed junction sequence information, (v) documents experimental methods applied in original studies, (vi) provides detailed information about disrupted genes and gene regions and (vii) provides an intuitive browser view for examining the CR events.

FUTURE DIRECTIONS

The development and release of the current version of dbCRID marks an important first step towards our long-term goal. Our next-step plans include: (i) we will continue and expand the dbCRID curation effort as published studies on CR events and associated diseases accumulate. The advances in new experimental techniques and strategies, especially in deep sequencing technologies will lead to accelerated accumulation of interesting data (28). Large-scale sequencing projects, in particular, the Cancer Genome Project (29,30) will continue to feed important data to the curation effort. (ii) We will try to identify potential genomic architecture features, e.g. CpG island, low copy repeats which may contribute to CR susceptibility. (iii) We will investigate patterns in disrupted genes,

gene clusters or pathways and their potential associations with disease etiology. (iv) We will develop and improve methods for identifying and analyzing junction sequence features critical for disease mechanisms (31).

ACKNOWLEDGEMENTS

The authors thank Minnesota Supercomputing Institute and its support team.

FUNDING

National Institutes of Health/National Cancer Institute (5R33 CA126209-03). Funding for open access charge: National Institutes of Health/National Cancer Institute.

Conflict of interest statement. None declared.

REFERENCES

- van Gent,D.C., Hoeijmakers,J.H. and Kanaar,R. (2001) Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.*, **2**, 196–206.
- Shrivastav,M., De Haro,L.P. and Nickoloff,J.A. (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res.*, **18**, 134–147.
- Stephens,P.J., McBride,D.J., Lin,M.L., Varela,I., Pleasance,E.D., Simpson,J.T., Stebbings,L.A., Leroy,C., Edkins,S., Mudie,L.J. *et al.* (2009) Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature*, **462**, 1005–1010.
- Muhle,R., Trentacoste,S.V. and Rapin,I. (2004) The genetics of autism. *Pediatrics*, **113**, e472–e486.
- Stilgenbauer,S., Bullinger,L., Lichter,P. and Dohner,H. (2002) Genetics of chronic lymphocytic leukemia: genomic aberrations and V(H) gene mutation status in pathogenesis and clinical course. *Leukemia*, **16**, 993–1007.
- Helmrich,A., Lee,S., O'Brien,P., Dorken,B., Lowe,S.W., Schrock,E. and Schmitt,C.A. (2005) Recurrent chromosomal aberrations in INK4a/ARF defective primary lymphomas predict drug responses in vivo. *Oncogene*, **24**, 4174–4182.
- Mullighan,C.G., Miller,C.B., Radtke,L., Phillips,L.A., Dalton,J., Ma,J., White,D., Hughes,T.P., Le Beau,M.M., Pui,C.H. *et al.* (2008) BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*, **453**, 110–114.
- Kalish,J.M., Seidman,M.M., Weeks,D.L. and Glazer,P.M. (2005) Triplex-induced recombination and repair in the pyrimidine motif. *Nucleic Acids Res.*, **33**, 3492–3502.
- Tycko,B., Reynolds,T.C., Smith,S.D. and Sklar,J. (1989) Consistent breakage between consensus recombinase heptamers of chromosome 9 DNA in a recurrent chromosomal translocation of human T cell leukemia. *J. Exp. Med.*, **169**, 369–377.
- Wei,Y., Sun,M., Nilsson,G., Dwight,T., Xie,Y., Wang,J., Hou,Y., Larsson,O., Larsson,C. and Zhu,X. (2003) Characteristic sequence motifs located at the genomic breakpoints of the translocation t(X;18) in synovial sarcomas. *Oncogene*, **22**, 2215–2222.
- Jaeger,U., Karth,G.D., Knapp,S., Friedl,J., Laczika,K. and Kusec,R. (1994) Molecular mechanism of the t(14;18)—a model for lymphoid-specific chromosomal translocations. *Leuk. Lymphoma*, **14**, 197–202.
- Chuzhanova,N., Chen,J.M., Bacolla,A., Patrinos,G.P., Ferec,C., Wells,R.D. and Cooper,D.N. (2009) Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair. *Hum. Mutat.*, **30**, 1189–1198.
- Lothe,R.A., Heimdal,K., Lier,M.E., Fossa,S.D., Moller,P. and Brogger,A. (1992) High resolution chromosome banding in search of germ line mutations applied on testicular cancer patients. *Cancer Genet. Cytogenet.*, **59**, 62–67.

14. Speicher, M.R., Gwyn Ballard, S. and Ward, D.C. (1996) Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet.*, **12**, 368–375.
15. Adam, P., Steinlein, C., Schmid, M., Haralambieva, E., Stocklein, H., Leich, E., Rosenwald, A., Muller-Hermelink, H.K. and Ott, G. (2006) Characterization of chromosomal aberrations in diffuse large B-cell lymphoma (DLBL) by G-banding and spectral karyotyping (SKY). *Cytogenet. Genome Res.*, **114**, 274–278.
16. Kallioniemi, A., Kallioniemi, O.P., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F. and Pinkel, D. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*, **258**, 818–821.
17. Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T. *et al.* (2002) The Ensembl genome database project. *Nucleic Acids Res.*, **30**, 38–41.
18. Abeysinghe, S.S., Chuzhanova, N., Krawczak, M., Ball, E.V. and Cooper, D.N. (2003) Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs. *Hum. Mutat.*, **22**, 229–244.
19. Lee, J.A., Carvalho, C.M. and Lupski, J.R. (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*, **131**, 1235–1247.
20. Cote, R.G., Jones, P., Martens, L., Apweiler, R. and Hermjakob, H. (2008) The Ontology Lookup Service: more data and better tools for controlled vocabulary queries. *Nucleic Acids Res.*, **36**, W372–W376.
21. Stein, L.D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., Nickerson, E., Stajich, J.E., Harris, T.W., Arva, A. *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res.*, **12**, 1599–1610.
22. Huret, J.L., Dessen, P. and Bernheim, A. (2001) Atlas of genetics and cytogenetics in oncology and haematology, updated. *Nucleic Acids Res.*, **29**, 303–304.
23. Kim, P., Yoon, S., Kim, N., Lee, S., Ko, M., Lee, H., Kang, H., Kim, J. and Lee, S. ChimerDB 2.0—a knowledgebase for fusion genes updated. *Nucleic Acids Res.*, **38**, D81–D85.
24. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M. and Carter, N.P. (2009) DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.*, **84**, 524–533.
25. Kim, D.S., Huh, J.W. and Kim, H.S. (2007) HYBRIDdb: a database of hybrid genes in the human genome. *BMC Genomics*, **8**, 128.
26. Knutsen, T., Gobu, V., Knaus, R., Padilla-Nash, H., Augustus, M., Strausberg, R.L., Kirsch, I.R., Sirotkin, K. and Ried, T. (2005) The interactive online SKY/M-FISH & CGH database and the Entrez cancer chromosomes search database: linkage of chromosomal aberrations with the genome sequence. *Genes Chromosomes Cancer*, **44**, 52–64.
27. Novo, F.J., de Mendibil, I.O. and Vizmanos, J.L. (2007) TICdb: a collection of gene-mapped translocation breakpoints in cancer. *BMC Genomics*, **8**, 33.
28. Chen, W., Kalscheuer, V., Tzschach, A., Menzel, C., Ullmann, R., Schulz, M.H., Erdogan, F., Li, N., Kijas, Z., Arkesteijn, G. *et al.* (2008) Mapping translocation breakpoints by next-generation sequencing. *Genome Res.*, **18**, 1143–1149.
29. Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabé, R.R., Bhan, M.K., Calvo, F., Eerola, I., Gerhard, D.S. *et al.* (2010) International network of cancer genome projects. *Nature*, **464**, 993–998.
30. Ledford, H. (2010) Big science: the cancer genome challenge. *Nature*, **464**, 972–974.
31. Shaw, C.J. and Lupski, J.R. (2004) Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum. Mol. Genet.*, **13(Spec No 1)**, R57–R64.