Oncogene (2010) 29, 2509–2516

www.nature.com/onc

ORIGINAL ARTICLE

Genomic segmental duplications on the basis of the t(9;22) rearrangement in chronic myeloid leukemia

F Albano[1,3], L Anelli[1,3], A Zagaria[1,3], N Coccaro[1], P D'Addabbo[2], V Liso[1], M Rocchi[2,4] and G Specchia[1,4]

[1]Department of Hematology, University of Bari, Bari, Italy and [2]Department of Genetics and Microbiology, University of Bari, Bari, Italy

A crucial role of segmental duplications (SDs) of the human genome has been shown in chromosomal rearrangements associated with several genomic disorders. Limited knowledge is yet available on the molecular processes resulting in chromosomal rearrangements in tumors. The t(9;22)(q34;q11) rearrangement causing the *5′BCR/3′ABL* gene formation has been detected in more than 90% of cases with chronic myeloid leukemia (CML). In 10–18% of patients with CML, genomic deletions were detected on der(9) chromosome next to translocation breakpoints. The molecular mechanism triggering the t(9;22) and deletions on der(9) is still speculative. Here we report a molecular cytogenetic analysis of a large series of patients with CML with der(9) deletions, revealing an evident breakpoint clustering in two regions located proximally to *ABL* and distally to *BCR*, containing an interchromosomal duplication block (SD_9/22). The deletions breakpoints distribution appeared to be strictly related to the distance from the SD_9/22. Moreover, bioinformatic analyses of the regions surrounding the SD_9/22 revealed a high Alu frequency and a poor gene density, reflecting genomic instability and susceptibility to rearrangements. On the basis of our results, we propose a three-step model for t(9;22) formation consisting of alignment of chromosomes 9 and 22 mediated by SD_9/22, spontaneous chromosome breakages and misjoining of DNA broken ends.
Oncogene (2010) **29,** 2509–2516; doi:10.1038/onc.2009.524; published online 25 January 2010

Keywords: segmental duplications; chronic myeloid leukemia; microdeletions

Introduction

During the last few years, genome analyses have revealed the crucial role of segmental duplications (SDs) in triggering constitutional and also tumor chromosomal abnormalities (Sharp *et al*., 2006; Gibcus *et al*., 2007; Darai-Ramqvist *et al*., 2008; Gu *et al*., 2008; Mefford and Eichler, 2009). Several rearrangements have been described so far to explain the occurrence of genomic disorders: recurrent, sharing a common size and showing clustering of breakpoints inside the SDs, and nonrecurrent rearrangements, involving regions of different sizes and showing breakpoints scattering in large regions (Gu *et al*., 2008). Most of the recurrent rearrangements result from a nonallelic homologous recombination between closely located SDs (Gu *et al*., 2008). Although a great deal of information has accumulated on the mechanisms underlying constitutional DNA rearrangements associated with inherited disorders, limited knowledge is yet available on the molecular processes resulting in chromosomal, somatic rearrangements in tumors. Recently, a role for SDs in the genesis of i(17q) in cancer has been established, strengthening the assumption that somatic rearrangements associated with human neoplasia are not random events but rather reflect susceptibilities resulting from the genomic structure (Barbouti *et al*., 2004). To date nothing is known about the molecular processes at the basis of the t(9;22)(q34;q11) rearrangement associated with chronic myeloid leukemia (CML). Some years ago, a 76-kb duplicon was reported, located close to both the *ABL* and *BCR* genes that are involved in the t(9;22)(q34;q11) translocation associated with CML (Saglio *et al*., 2002). However, the exact role of this duplicon in mediating the t(9;22) rearrangement remained mostly speculative. Moreover, it is well known that in 10–18% of patients with CML the t(9;22) is an unbalanced rearrangement, because genomic sequences of chromosomes 9 and 22 are lost during the translocation (Sinclair *et al*., 2000; Huntly *et al*., 2001; Kolomietz *et al*., 2001; Storlazzi *et al*., 2002). The mechanism on the basis of these genomic microdeletions still remains unclear.

In this paper, we present experimental evidence of the involvement of SDs in the genesis of the t(9;22) translocation in CML and in the occurrence of genomic deletions on the der(9) chromosome. We report a fine-mapping of der(9) deletions by fluorescence *in situ* hybridization (FISH) analysis with bacterial artificial chromosome (BAC) and Phage P1-derived artificial chromosome (PAC) contigs in 71 patients with CML.

Correspondence: Dr F Albano, Department of Hematology, University of Bari, Azienda Ospedaliera Universitaria Policlinico, Piazza Giulio Cesare 11, 70124 Bari, Italy.
E-mail: f.albano@ematba.uniba.it
[3]These authors contributed equally to this work.
[4]These senior authors equally oversaw the design and conduction of the research.
Received 29 August 2009; revised 12 November 2009; accepted 19 December 2009; published online 25 January 2010

This study showed breakpoints clustering on two specific genomic segments, including interchromosomal SDs and characterized by structural features making DNA susceptible to double-strand breaks.

Results

By FISH screening of 416 patients with CML at diagnosis, we identified 71 (17%) cases with der(9) deletions. Fine-mapping of the deletions was performed using appropriate BAC/PAC clones. Deletions were either on both chromosomes (47 cases, 66%), or on chromosome 9 only (13 cases, 18%) or on chromosome 22 only (11 cases, 16%). The deletions sizes were heterogeneous, ranging from 230 kb to 12.9 Mb on chromosome 22 and from 260 kb to 41.8 Mb on chromosome 9. The mapping of all breakpoints is graphically reported in Supplementary Figures 1a and b. The analysis revealed an evident breakpoint clustering on both chromosomes 9 and 22, in two regions of about 2 Mb in size. Indeed, these regions (red rectangles in Supplementary Figures 1a and b) contained the breakpoints detected in 54 out of 60 (90%) patients bearing chromosome 9 deletions and in 51 out of 58 (88%) patients with chromosome 22 sequences loss. In detail, these breakpoints clusters were delimited by RP11-379C10 (chr9:129,792,681-129,988,611) and RP11-618A20 (chr9:132,212,882-132,385,904) on chromosome 9 and by RP11-1112A23 (chr22:22,217,931-22,384,780) and RP11-1143M16 (chr22:24,055,620-24,209,526) on chromosome 22 (Supplementary Figures 1a and b).

Bioinformatic analysis of chromosome 9 and 22 genomic regions involved in the deletions was performed to search for features that could be correlated with the breakpoints clustering. To this aim, we subdivided the breakpoint regions into 250 kb intervals. We evaluated the distribution of (1) SDs, (2) Alu sequences, (3) LINE sequences, (4) GC content, (5) genes, (6) topoisomerase II recognition sites, (7) matrix association regions and (8) scaffold attachment regions. Relevant findings are graphically reported in Figure 1.

The most striking result was the fact that both clusters contain the above reported 76-kb duplicon, shared by chromosomes 9 and 22 (SD_9/22). The SD_9/22 shows an average 90% identity and maps at a distance of 1.2 Mb proximally to the *ABL* gene and at 175 kb distally to the *BCR* gene. The SD_9/22 is the only duplication located inside the breakpoints clustering region on chromosome 9, whereas the chromosome 22 clustering region harbors several duplications (Figures 1a and b). On chromosome 9, the SD_9 is located in the middle of the breakpoints clustering region whereas on chromosome 22 the SD_22 is at the border (Figures 1a and b; Supplementary Figures 1a and b). A remarkable feature of the chromosome 9 clustering region was the high frequency of Alu repeats (Figure 1c). The mean Alu frequency overall on chromosome 9 is 10.8%, whereas the average Alu content on this cluster is 31.3%. Accordingly, as expected, the content in LINE sequences of the region was relatively low (average overall on chromosome 9: 21.2%, as opposed to 8.7% on the cluster region) (Figure 1d).

SD_9/22 consists of five different segments showing an average size of 12 kb, separated by single-copy DNA sequences of variable size. In total, SD_9 and SD_22 cover 188 kb (chr9:131,188,486-131,376,512) and 79 kb (chr22:22,165,774-22,244,483) regions, respectively (Figure 2). The duplication block is arranged in a complex manner, as one out of five segments lies in the same direction whereas the remaining four show an opposite orientation. Gene distribution analysis of chromosomes 9 and 22 showed that both SD_9/22 map inside gene-poor regions, of about 460 and 250 kb in size, respectively, as clearly shown in Figure 3. No significant association between the distribution of deletions breakpoints and the GC content, topoisomerase II recognition sites, matrix association regions and scaffold attachment regions was detected on either chromosome 9 or 22 (data not shown).

To corroborate the observations on the distribution of SDs and Alu/LINE repeats, we once more divided the chromosome 9 and 22 regions surrounding the SD_9/22 into 250 kb segments, taking SD_9 and SD_22 as landmarks. A statistically significant negative association was observed between the number of breaks and the distance from SD_9/22, on both chromosomes 9 ($P=0.01$) and 22 ($P=0.006$) (Figures 4a and b), respectively. The relationship between the breaks and the interspersed repeats revealed, on chromosome 9, a positive linear regression with Alu repeats ($P=0.04$), and a negative one with LINEs ($P=0.04$) (Figures 4c and d). Very similar conclusions were obtained comparing the distance from the SD_9 and the Alu ($P=0.03$, positive) and LINE distribution ($P=0.02$, negative) (Figures 4e and f). No statistically significant relationship was observed on chromosome 22 (data not shown).

Bioinformatic analysis of breakpoint regions on additional chromosomes involved in variant t(9;22) showed the presence of several SDs, without a specific association with chromosomes 9 and 22.

Discussion

SDs cover about 10% of the human genome and are involved in numerous genomic diseases or cancer (Bailey and Eichler, 2006; Sharp *et al*., 2006; Gibcus *et al*., 2007; Darai-Ramqvist *et al*., 2008; Gu *et al*., 2008; Mefford and Eichler, 2009). In this paper, the involvement of SDs was proposed to explain the recurrent t(9;22) translocation in CML and the genomic deletions that could accompany the rearrangement. In fact, several groups have previously identified der(9) deletions next to translocation breakpoints as a frequent feature of patients with CML (Sinclair *et al*., 2000; Huntly *et al*., 2001; Kolomietz *et al*., 2001; Storlazzi *et al*., 2002; Specchia *et al*., 2004; Albano *et al*., 2007). To date, the characterization of deletions breakpoints has shown that breaks are scattered throughout large genomic regions,

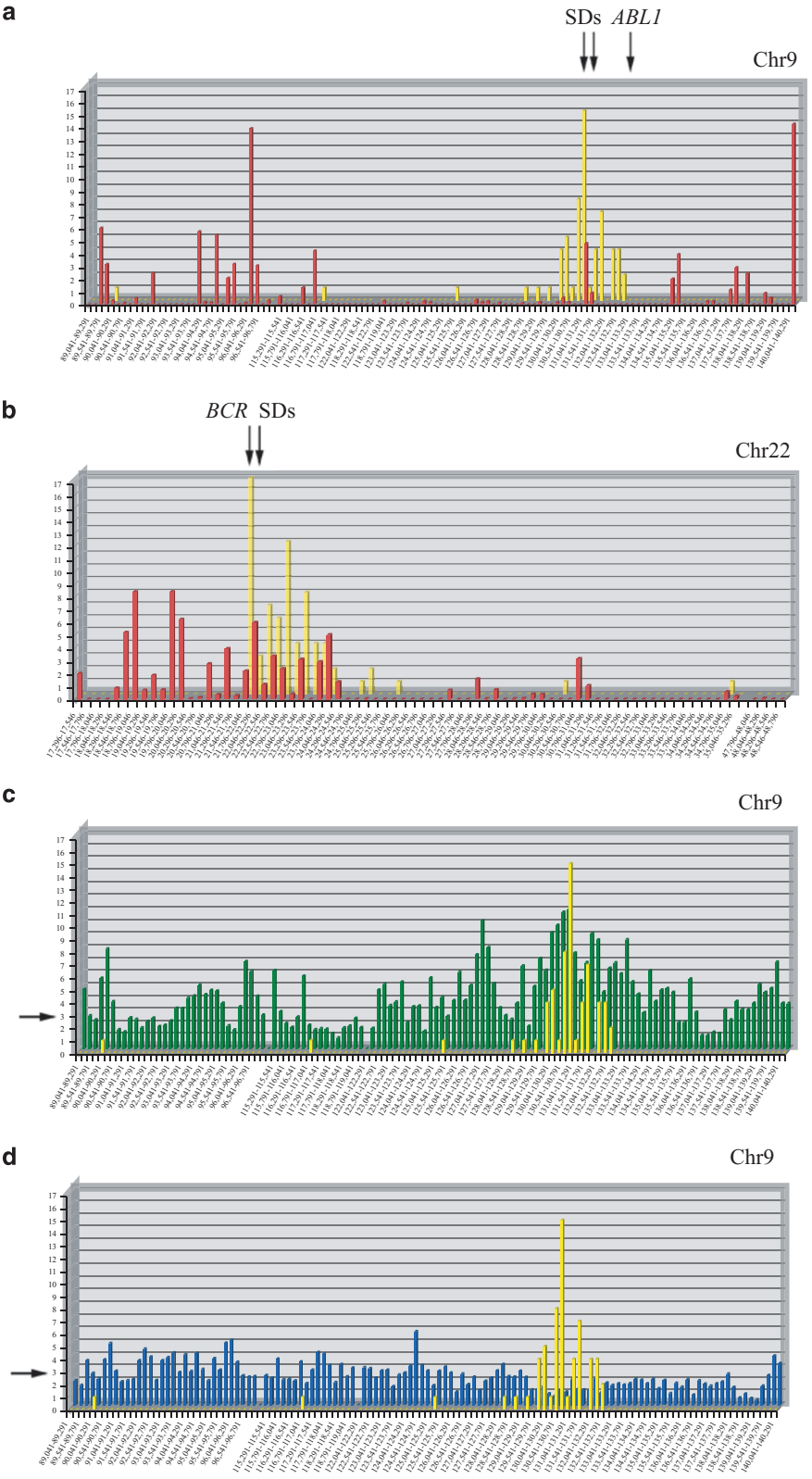

Figure 1 Bioinformatic studies performed on the chromosomes 9 and 22 analyzed regions. Histograms showing the distribution of deletions breakpoints (yellow) as compared to the frequency of segmental duplications (red) on chromosomes 9 (**a**) and 22 (**b**). In (**c**) and (**d**) the frequency of Alu (green) and LINE (blue) repeats on chromosome 9 is reported with respect to the number of deletions breakpoints (yellow). Below each chart the 250 kb size intervals are shown. The vertical arrows show the mapping position of the SD_9/22 and genes. Alu and LINE frequencies were both increased three times to better appreciate their distribution patterns (horizontal arrows).

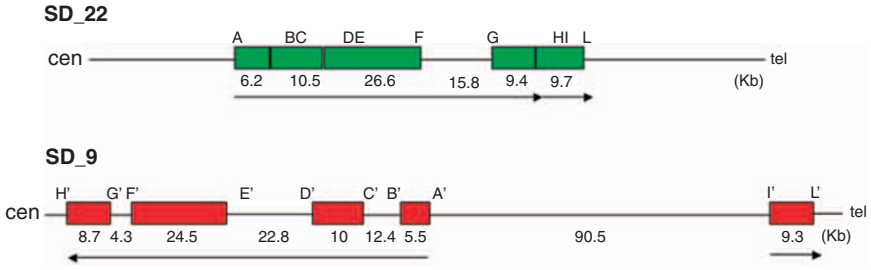

Figure 2 Genomic organization of SD_9/22. The ends of the five segments that constitute each duplication are indicated by capital letters whereas the horizontal black line represents nonduplicated genomic regions. Four out of five segments delimited by A-H and A′-H′ are arranged in an opposite orientation whereas the two fragments I-L and I′-L′ lie in the same direction. The size of each segment and of single-copy sequences is reported in kb.

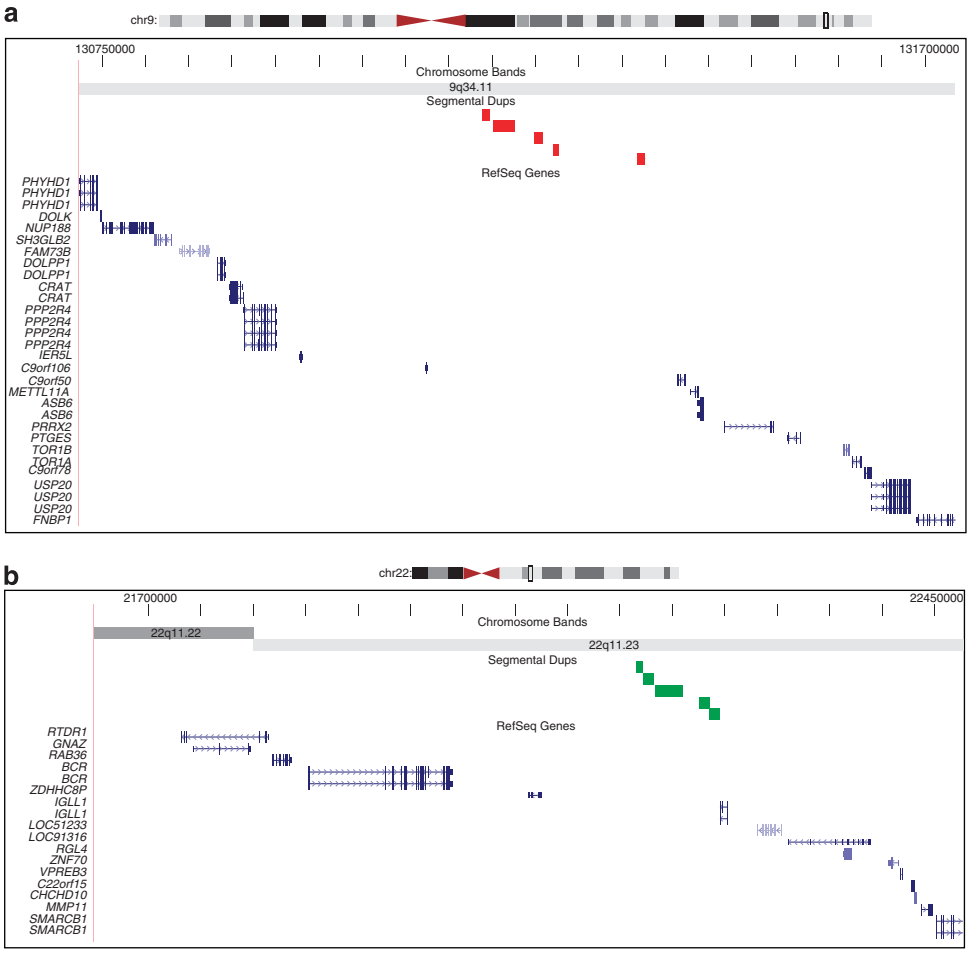

Figure 3 Genes map next to the SD_9/22. The image generated querying the UCSC database shows the localization of the RefSeq Genes in correspondence with SD_9 (**a**) and SD_22 (**b**). It is noteworthy that both segmental duplications (SDs) are included in gene-poor regions as compared with the chromosomal flanking sequences.

resulting in variable deletions size as a consequence of a nonrecurrent event (Sinclair *et al*., 2000; Storlazzi *et al*., 2002; Kolomietz *et al*., 2003; Fourouclas *et al*., 2006; Kreil *et al*., 2007).

The deletions on der(9) have been associated with an adverse prognosis in relation to the efficacy of interferon-α therapy, whereas controversial data are available about their influence on the response to imatinib. Several studies have suggested that the prognostic impact of der(9) sequences loss could depend on the deletions size or on their extension on one or both sides of the *5′ABL/3′BCR* fusion gene (Fourouclas *et al*., 2006; Kreil *et al*., 2007; Vaz de Campos *et al*., 2007).

In our study, structural analysis of deletions breakpoints in a very large cohort of patients with CML has been a crucial way to identify genomic regions with a

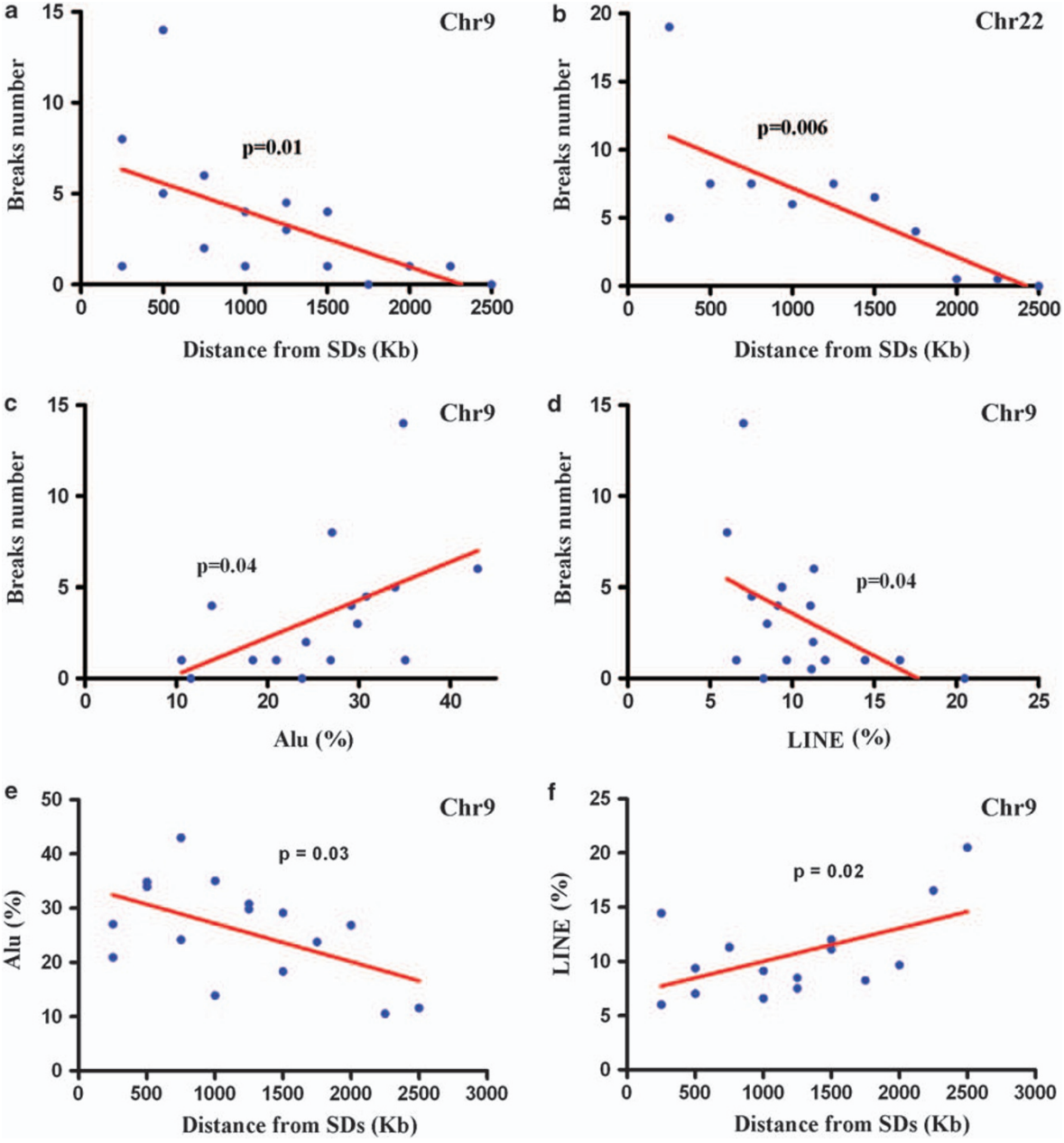

Figure 4 Statistical analysis of the examined genomic features. Linear regression analysis shows a statistically significant negative relationship between the number of breaks and the distance from SD_9 (**a**) to SD_22 (**b**), a positive and a negative association between the number of breaks and Alu (**c**) and LINE (**d**) frequency on chromosome 9. In (**e**) and (**f**) the distance from SD_9 is compared with the Alu and LINE frequencies, respectively.

pivotal role in the t(9;22) rearrangement and in the occurrence of sequences loss on der(9). Patients with CML bearing der(9) deletions share common overlapping genomic regions encompassing the *ABL* and *BCR* genes. The deletions breakpoints are heterogeneous on both chromosomes 9 and 22, as they represent secondary molecular events not under the selective pressure. However, most of them fall inside two clustering regions sharing an interchromosomal SD_9/22 (Figure 5). The location of SD_22 at the border of the breakpoints clustering region on chromosome 22 may be explained by its strict proximity to the *BCR* gene; deletions breakpoints located centromerically to *BCR* could never occur in CML cells bearing the *5′BCR/3′ABL* fusion gene.

It is noteworthy that the presence of SD_9/22 near the translocation breakpoints on both chromosomes is highly unlikely to happen by chance. Because the entire genome contains about 3.1 thousand sequenced Mb, the probability of finding the duplication partner for a chromosome 22 segment within a 2 Mb region surrounding the chromosome 9 breakpoint by chance can be estimated as very low ($2/3100 = 0.0006$). In this respect, our study revealed a statistically significant correlation

between the breakpoints distribution and the distance from the duplicons, as the number of breaks decreases moving away from the SDs.

The complex arrangement of SD_9/22 and the genomic features identified by our analyses, such as the Alu content and the poor gene density, highlighted the structural complexity and the genomic instability of the chromosomal regions surrounding the t(9;22) breakpoints. The content of Alu repeats is high in proximity to the SD_9 and decreases gradually at increasing distances from it. It is well known that Alu elements accelerate SDs formation and the association of SDs with Alu elements decreases with the decreasing age of the SDs (Kim *et al*., 2008). In this respect, our data showed a significant association between Alu repeats and SD_9, which represents the oldest duplication according to the sequence evolutionary history (Saglio *et al*., 2002). Interspersed repetitive elements such as Alu and LINE are commonly present at or near genomic breakpoints. Such repetitive elements may have an important role by providing substrates with a specific DNA secondary structure that stabilizes broken chromosomes, increasing the probability of rearrangements (Yatsenko *et al*., 2009).

Moreover, the poor gene density surrounding the SD_9/22 appeared as a favorable scenario permissive of SDs fixation (Lomiento *et al*., 2008). In fact, the absence of coding sequences renders these genomic regions free to recombine.

On the basis of our results, we propose a three-step model of t(9;22) rearrangement consisting of alignment of chromosomes 9 and 22, spontaneous chromosome breakages and misjoining of DNA broken ends (Figure 6). The alignment could be allowed by the homology with the SD_9/22, producing an abnormal chromatin conformation. The attempt to resolve these unstable DNA structures subjected to torsional stress could induce double-strand breaks. The occurrence of double-strand breaks could also be stimulated by the significant enrichment of Alu repeats around the *ABL* locus (about threefold) (Elliott and Jasin, 2002). Finally, the incorrect joining of DNA broken ends results in DNA exchange and in the 5′*BCR*/3′*ABL* fusion gene. In some cases, additional secondary events consisting of genomic deletions could occur at the junction sites on chromosome der(9).

A duplication of a 258-bp M-bcr fragment on both Ph and der(9) chromosomes was previously reported in a small group of patients with CML. Several hypotheses were made about the occurrence of these duplications and their involvement in the t(9;22) rearrangement (Litz *et al*., 1993). The hypothesis that the duplicated sequence might exist on chromosome 9 before the translocation event, creating a potential site for

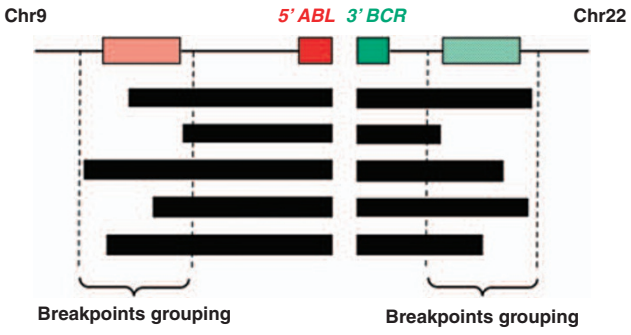

Figure 5 Double grouping of deletions breakpoints on der(9)t(9;22) chromosome. Deletions breakpoints map mainly in correspondence with SD_9 and SD_22 (dotted rectangles in red and green, respectively), defining two grouping regions. Black bars represent der(9) deletions, extending from 5′*ABL* (red square) and/or 3′ *BCR* (green square) to the breakpoints grouping regions.

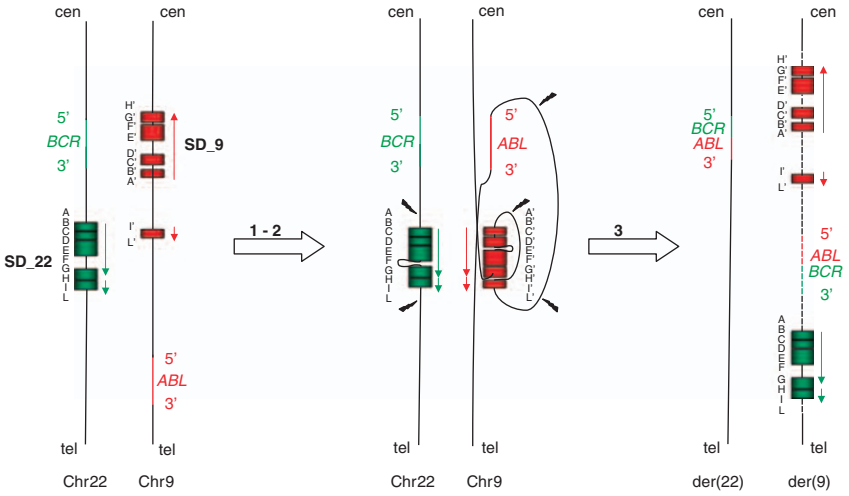

Figure 6 Model of molecular rearrangement at the basis of t(9;22)(q34;q11) and der(9) deletions formation. The SD_9/22 and the *ABL* and *BCR* genes location is reported on chromosomes 9 and 22, respectively. The letters A-L and A′-L′ mark the extremities of each repeated segment. The pairing between homologous regions generates a conformational DNA distortion as a consequence of the SD_9/22 orientation and complex arrangement (step 1). Chromosome breakages (indicated by lightning symbols) could occur to resolve the unstable DNA structures (step 2). The erroneous joining of DNA broken ends could result in the t(9;22) rearrangement whereas the extensive exonuclease activities of the DNA repair mechanism could generate sequences deletions on der(9), as represented by dotted lines (step 3). The represented genomic distances are not in scale.

homologous recombination, was in agreement with the model suggested by our findings.

In conclusion, although the chromosomes 9 and 22 breakpoints clustering regions are quite large, the strong nonrandomness of SD_9/22 location and the genomic features identified in this study suggest that the chromsomal segments near the *ABL* and *BCR* genes facilitate their alignment and recombination. In the light of these findings, the analysis of secondary nonrecurrent events could represent a new methodological approach able to identify architectural elements involved in the occurrence of recurrent primary rearrangements in human neoplasia.

Materials and methods

FISH analysis

FISH analysis was carried out on bone marrow samples of 416 patients with CML at diagnosis with specific BAC and PAC clones. In detail, the *ABL* and *BCR* genes identification was performed using a pool of PAC, RP5-1132H12 (chr9: 132,534,486-132,656,096) and RP5-835J22 (chr9:132,604,903-132,774,088), and BAC RP11-164N13 (chr22:21,892,458-22,086,126), respectively (Storlazzi *et al*., 2002; Specchia *et al*., 2004; Albano *et al*., 2007). A second round of FISH cohybridization with RP11-17L7 (9q34.11) (chr9:130,371,696-130,536,601), proximal to *ABL*, and RP11-248J22 (22q11.23) (chr22:22,068,196-22,215,310), distal to *BCR* and overlapping with RP11-164N13, was performed in each case (Storlazzi *et al*., 2002; Albano *et al*., 2007). To define the microdeletions size, we selected a set of BAC/PAC probes according to the University of California Santa Cruz database (UCSC, http://www.genome.ucsc; May 2006 release) generating two contigs, covering several Mbs, centromerically and telomerically to the *ABL* and *BCR* genes, respectively (Storlazzi *et al*., 2002; Albano *et al*., 2007) (Supplementary Figures 1a and b). Among 71 patients with deletions on der(9), 18 (indicated by '*') have been previously described in terms of deletions size (Storlazzi *et al*., 2002; Albano *et al*., 2007). Further FISH experiments with appropriate clones selected by UCSC database allowed the identification of breakpoints on additional chromosomes involved in variant t(9;22). Chromosome preparations were hybridized *in situ* with probes labeled with biotin by nick translation (Lichter *et al*., 1990). Briefly, 500 ng of labeled probe was used for FISH experiments; hybridization was performed at 37 °C in 2 × standard saline citrate, 50% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, 5 mg COT1 DNA (Bethesda Research Laboratories, Gaithersburg, MD, USA), and 3 mg sonicated salmon sperm DNA in a volume of 10 ml. Posthybridization washing was at 60 °C (0.1 × standard saline citrate). Biotin-labeled DNA was detected with Cy3-conjugated avidin. In cohybridization experiments, other probes were directly labeled with fluorescein. Chromosomes were identified by 4′,6-diamidino-2-phenylindole staining. Digital images were obtained using a Leica DMRXA epifluorescence microscope equipped with a cooled CCD camera (Princeton Instruments, Boston, MA, USA). Cy3 (red; New England Nuclear, Boston, MA, USA), fluorescein (green; NEN Life Science Products, Boston, MA, USA) and 4′,6-diamidino-2-phenylindole (blue) fluorescence signals, which were detected by using specific filters, were recorded separately as grayscale images. Pseudocoloring and merging of images were performed with Adobe Photoshop software (Adobe, San Jose, CA, USA).

Bioinformatic analysis

Chromosomal regions proximally and distally to the *ABL* (chr9:89,041,000–140,273,251 bp) and *BCR* (chr22:17,296,000–35,296,000 bp) genes were divided *in silico* into 250 kb size fragments; each fragment was checked for the presence of interspersed repeats classes (Alu and LINE repeats), SDs, GC content and gene density. The UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables) was queried for summary analysis about the items belonging to the tracks 'RepeatMasker', 'Segmental Dups', 'GC Percent' and 'RefSeq Genes'.

Specific chromatin structural elements making DNA susceptible to double-strand breaks were investigated. Putative topoisomerase II consensus sites, matrix association regions and scaffold attachment regions were searched for through the fuzznuc (local installation of EMBOSS tools) and the MARSCAN (online: http://anabench.bcm.umontreal.ca/cgi-bin/emboss.pl?_action=input&_app=marscan) algorithms. Bioinformatic analysis of breakpoint regions on other chromosomes involved in variant t(9;22) was performed to verify the presence of SDs.

Statistical analysis

Linear regression analysis was performed to test the relationship between the breakpoints distribution and the distance from SDs, the Alu and LINE frequency, the SDs frequency, the GC content and the gene density. Moreover, a linear regression test was used to verify the degree of relation between the distance from duplicons and the interspersed repeats class frequency, the SDs frequency, the GC content and the gene density. Only P-values $\leqslant 0.05$ were considered significant. The analysis was performed with GraphPad Prism software v.5.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Ms MVC Pragnell for language revision of the paper. The financial support of Associazione Italiana contro le Leucemie (AIL)-BARI is gratefully acknowledged.

References

Albano F, Anelli L, Zagaria A, Archidiacono N, Liso V, Specchia G *et al*. (2007). 'Home-brew' FISH assay shows higher efficiency than BCR-ABL dual color, dual fusion probe in detecting microdeletions and complex rearrangements associated with t(9;22) in chronic myeloid leukemia. *Cancer Genet Cytogenet* **174**: 121–126.

Bailey JA, Eichler EE. (2006). Primate segmental duplications: crucibles of evolution, diversity and disease. *Nat Rev Genet* **7**: 552–564.

Barbouti A, Stankiewicz P, Nusbaum C, Cuomo C, Cook A, Hoglund M *et al*. (2004). The breakpoint region of the most common

isochromosome, i(17q), in human neoplasia is characterized by a complex genomic architecture with large, palindromic, low-copy repeats. *Am J Hum Genet* **74**: 1–10.

Darai-Ramqvist E, Sandlund A, Muller S, Klein G, Imreh S, Kost-Alimova M. (2008). Segmental duplications and evolutionary plasticity at tumor chromosome break-prone regions. *Genome Res* **18**: 370–379.

Elliott B, Jasin M. (2002). Double-strand breaks and translocations in cancer. *Cell Mol Life Sci* **59**: 373–385.

Fourouclas N, Campbell PJ, Bench AJ, Swanton S, Baxter EJ, Huntly BJ *et al*. (2006). Size matters: the prognostic implications of large and small deletions of the derivative 9 chromosome in chronic myeloid leukemia. *Haematologica* **91**: 952–955.

Gibcus JH, Kok K, Menkema L, Hermsen MA, Mastik M, Kluin PM *et al*. (2007). High-resolution mapping identifies a commonly amplified 11q13.3 region containing multiple genes flanked by segmental duplications. *Hum Genet* **121**: 187–201.

Gu W, Zhang F, Lupski JR. (2008). Mechanisms for human genomic rearrangements. *Pathogenetics* **1**: 4.

Huntly BJ, Reid AG, Bench AJ, Campbell LJ, Telford N, Shepherd P *et al*. (2001). Deletions of the derivative chromosome 9 occur at the time of the Philadelphia translocation and provide a powerful and independent prognostic indicator in chronic myeloid leukemia. *Blood* **98**: 1732–1738.

Kim PM, Lam HY, Urban AE, Korbel JO, Affourtit J, Grubert F *et al*. (2008). Analysis of copy number variants and segmental duplications in the human genome: evidence for a change in the process of formation in recent evolutionary history. *Genome Res* **18**: 1865–1874.

Kolomietz E, Al-Maghrabi J, Brennan S, Karaskova J, Minkin S, Lipton J *et al*. (2001). Primary chromosomal rearrangements of leukemia are frequently accompanied by extensive submicroscopic deletions and may lead to altered prognosis. *Blood* **97**: 3581–3588.

Kolomietz E, Marrano P, Yee K, Thai B, Braude I, Kolomietz A *et al*. (2003). Quantitative PCR identifies a minimal deleted region of 120 kb extending from the Philadelphia chromosome ABL translocation breakpoint in chronic myeloid leukemia with poor outcome. *Leukemia* **17**: 1313–1323.

Kreil S, Pfirrmann M, Haferlach C, Waghorn K, Chase A, Hehlmann R *et al*. (2007). Heterogeneous prognostic impact of derivative chromosome 9 deletions in chronic myelogenous leukemia. *Blood* **110**: 1283–1290.

Lichter P, Tang CJ, Call K, Hermanson G, Evans GA, Housman D *et al*. (1990). High-resolution mapping of human chromosome 11 by *in situ* hybridization with cosmid clones. *Science* **247**: 64–69.

Litz CE, McClure JS, Copenhaver CM, Brunning RD. (1993). Duplication of small segments within the major breakpoint cluster region in chronic myelogenous leukemia. *Blood* **81**: 1567–1572.

Lomiento M, Jiang Z, D'Addabbo P, Eichler EE, Rocchi M. (2008). Evolutionary-new centromeres preferentially emerge within gene deserts. *Genome Biol* **9**: R173.

Mefford HC, Eichler EE. (2009). Duplication hotspots, rare genomic disorders, and common disease. *Curr Opin Genet Dev* **19**: 196–204.

Saglio G, Storlazzi CT, Giugliano E, Surace C, Anelli L, Rege-Cambrin G *et al*. (2002). A 76-kb duplicon maps close to the BCR gene on chromosome 22 and the ABL gene on chromosome 9: possible involvement in the genesis of the Philadelphia chromosome translocation. *Proc Natl Acad Sci USA* **99**: 9882–9887.

Sharp AJ, Hansen S, Selzer RR, Cheng Z, Regan R, Hurst JA *et al*. (2006). Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**: 1038–1042.

Sinclair PB, Nacheva EP, Leversha M, Telford N, Chang J, Reid A *et al*. (2000). Large deletions at the t(9;22) breakpoint are common and may identify a poor-prognosis subgroup of patients with chronic myeloid leukemia. *Blood* **95**: 738–743.

Specchia G, Albano F, Anelli L, Storlazzi CT, Zagaria A, Liso A *et al*. (2004). Derivative chromosome 9 deletions in chronic myeloid leukemia are associated with loss of tumor suppressor genes. *Leuk Lymphoma* **45**: 689–694.

Storlazzi CT, Specchia G, Anelli L, Albano F, Pastore D, Zagaria A *et al*. (2002). Breakpoint characterization of der(9) deletions in chronic myeloid leukemia patients. *Genes Chromosomes Cancer* **35**: 271–276.

Vaz de Campos MG, Montesano FT, Rodrigues MM, Chauffaille Mde L. (2007). Clinical implications of der(9q) deletions detected through dual-fusion fluorescence *in situ* hybridization in patients with chronic myeloid leukemia. *Cancer Genet Cytogenet* **178**: 49–56.

Yatsenko SA, Brundage EK, Roney EK, Cheung SW, Chinault AC, Lupski JR. (2009). Molecular mechanisms for subtelomeric rearrangements associated with the 9q34.3 microdeletion syndrome. *Hum Mol Genet* **18**: 1924–1936.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)