Large genomic aberrations detected by SNP array are independent prognosticators of a shorter time to first treatment in chronic lymphocytic leukemia patients with normal FISH

M. Mian1,2, A. Rinaldi1, A. A. Mensah1, D. Rossi4, M. Ladetto4, F. Forconi5,6, R. Marasca7, M. Uhr8, G. Stussi8, I. Kwee9,10, F. Cavalli8, G. Gaidano3, E. Zucca8 & F. Bertoni1,8*

1Lymphoma and Genomics Research Program, IOR Institute of Oncology Research, Bellinzona, Switzerland; 2Division of Hematology & CBMT, Hospital of Bolzano, Bolzano; 3Division of Hematology, Department of Translational Medicine, Amedeo Arogadro University of Eastern Piedmont, Novara; 4Division of Hematology, Department of Experimental Medicine and Oncology, University of Turin, Turin, Italy; 5Cancer Sciences Unit, CRUK Clinical Centre, University of Southampton, Southampton, UK; 6Division of Hematology, University of Siena, Siena; 7Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy; 8Lymphoma Unit and Hematology, IOSI Oncology Institute of Southern Switzerland, Bellinzona; 9IDSIA Dalle Molle Institute for Artificial Intelligence, Manno; 10SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

Received 21 July 2012; revised 25 October 2012; accepted 3 December 2012

Background: Genomic complexity can predict the clinical course of patients affected by chronic lymphocytic leukemia (CLL) with a normal FISH. However, large studies are still lacking. Here, we analyzed a large series of CLL patients and also carried out the so far largest comparison of FISH versus single-nucleotide polymorphism (SNP) array in this disease.

Patients and methods: SNP-array data were derived from a previously reported dataset.

Results: Seventy-seven of 329 CLL patients (23%) presented with a normal FISH. At least one large (>5 Mb) genomic aberration was detected by SNP array in 17 of 77 patients (22%); this finding significantly affected TTT. There was no correlation with the presence of TP53 mutations. In multivariate analysis, including age, Binet stage, IGHV genes mutational status and large genomic lesion, the latter three factors emerged as independent prognosticators. The concordance between FISH and SNP array varied between 84 and 97%, depending on the specific genomic locus investigated.

Conclusions: SNP array detected additional large genomic aberrations not covered by the standard FISH panel predicting the outcome of CLL patients.

Key words: affymetrix, chronic lymphocytic leukemia, FISH, microarray, prognosis, TP53

introduction

The clinical course of patients affected by chronic lymphocytic leukemia (CLL) is very heterogeneous and ranges from an indolent and chronic disease to a rapidly progressing leukemia or lymphoma necessitating aggressive treatment [1–3]. Several parameters, such as staging, lymphocyte doubling time, CD38, ZAP70, the mutational status of the variable region of the immunoglobulin heavy chain (IGHV) and the presence of specific genomic aberrations, are often used to stratify patients based upon their expected clinical outcome [2–5]. A prognostic model based on genomic aberrations (del 13q14.3, del 11q22.3, del 17p13.1, trisomy 12) detected by fluorescent in situ hybridization (FISH) is also largely used [3, 6].

In the last few years, high density, single-nucleotide polymorphism (SNP) genomic arrays (SNP arrays) proved to be efficient techniques to identify genetic aberrations in malignant cells [7]. The major advantage of this approach in comparison to FISH is that the whole genome can be scanned for genomic aberrations at once and not only at a limited number of regions. Moreover, recent data have shown that the prediction of the outcome of CLL patients might be improved by the use of SNP arrays. The detection of individual genetic aberrations other than the ones detected by the standard FISH panel [8], or of an overall increased genomic complexity [9, 10] or of the capability to measure the size of the genomic aberrations [11–13], all appear to have clinical implications, better defining the clinical outcome of CLL patients. Very recently, by carrying out cytogenetic analysis with DSP30/IL2.
stimulation of CLL cells, Rigolin et al. [14] have shown that the presence of an abnormal karyotype correlates with a shorter time to first treatment (TTT) and shorter overall survival (OS). Although, in the last few years several investigators have described the application of microarrays in CLL [8, 10, 15–22], only a few of them have reported a detailed comparison with FISH, including the clinical implication of using a SNP array instead of the FISH panel to predict the outcome at diagnosis. In order to evaluate whether SNP array can better define the outcome of CLL patients with a normal FISH and can replace FISH, we analyzed a large group of CLL patients, characterized with both SNP array and the standard FISH panel.

**materials and methods**

**tumor panel**

CLL samples from a previously published series were investigated [8]. All samples were collected at diagnosis and had an immunophenotypic score of $>3$ according to Matutes et al. [23]. Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The study was approved by the Ethical Committee of Bellinzona. CLL diagnosis and management were based on the National Cancer Institute working group criteria [3]. Cases of monoclonal B-cell lymphocytosis were discarded. Baseline clinical and molecular characteristics have been collected as previously described [8]. The FISH panel interrogated deletions at 13q14.3, 11q22, 17p13 and trisomy 12 [8].

**SNP arrays**

Copy-number data, obtained from high-molecular-weight genomic DNA of CLL patients using the Genome-Wide Human SNP Array Version 6.0 (Affymetrix, Santa Clara, CA), were derived from a previously reported dataset [8]. The bioinformatics pipeline used for the identification of copy-number alterations is described in detail by Rinaldi et al. [8]. In addition, raw and smoothed copy-number values of each sample were plotted for quality assessment. Copy-number aberrations $>5$ Mb were identified by filtering on segment size and confirmed by visual inspection of the plotted genomic profiles. The cut-off was stabilized according to a previously published experience [24]. Validation of 13q14.3 losses was carried out using real-time PCR, as previously described [25]. TP53 gene mutations were evaluated as previously described [26].

**statistical analysis**

TTT was measured from the date of diagnosis until first treatment or, for untreated patients, to last follow-up. OS was measured from the date of diagnosis to the last follow-up or death by any cause. The cumulative probability of OS and TTT were plotted as curves according to the Kaplan–Meier [27] method. A log-rank test was used to investigate the impact on OS and TTT of categorical variables. Cox’s proportional hazard model [28] was applied to evaluate the impact on OS and TTT of known clinical and biological risk factors as well as the large genomic aberrations identified by SNP array. A P-value of $<0.05$ was considered as statistically significant. Statistical analyses were carried out with the SPSS v.17.0.1 (SPSS, Chicago, IL) and with Stata/SE v.12.1 (StataCorp, College Station, TX).

**results**

**patient characteristics**

The main patient characteristics are summarized in Table 1. The median age at the time of diagnosis was 66 years (range 27–94). The majority of the patients (60%) were male and only a few presented with advanced disease (28%). Unfavorable prognostic factors, namely unmutated IGHV-genes, CD38 $>30\%$, ZAP 70 $>20\%$, beta2-microglobulin above the upper normal value (UNV), lymphocyte doubling time <12 months were registered in 39%, 36%, 39%, 61% and 18%, respectively.

**genomic lesions affect the outcome of patients with a ‘normal’ FISH**

We investigated whether the presence of large genomic aberrations (>5 megabases; Mb; Figure 1), not covered by the probes used in the FISH panel, would affect the outcome in patients with a ‘normal’ FISH, who are known to have an intermediate outcome [6]. Seventy-seven of the 329 (23%) did not present any lesion by FISH. Seventeen (21%) of these patients with a normal FISH had additional DNA gains or losses in the SNP array affecting different genomic regions (Table 2). No TP53 mutations were detected in 13 of 17 cases with genomic lesions that were analyzed for mutations. The presence of at least one large genomic aberration determined a significantly shorter TTT ($P = 0.0010$; hazard ratio (HR) 2.8; 95% CI 1.5–5.5) ($P = 0.001$, but not OS ($P = 0.098$; HR 2.3; 95% CI 0.83–6.6) (Figure 2).

In multivariate analysis, including age $>60$ years, advanced Binet stage disease, unmutated IGHV genes and the presence of at least one large genomic lesion different from those described by Döhner et al. [29], the latter three factors emerged as independent prognosticators for TTT with a significance of $P < 0.001$, $P < 0.001$ and $P = 0.036$, respectively. None of the examined variables maintained their significance in the same model according to OS (data not shown). In particular, the presence of at least one large genomic aberration identified patients with a shorter TTT among those with mutated IGHV genes ($P = 0.0002$; HR 5.0; 95% CI 2.0–12.4) (Figure 3, left panel) and with early stage disease ($P = 0.005$; HR 3.22; 95% CI 1.36–7.62) (Figure 3, right panel).

**SNP array and FISH give overall similar results**

We compared the copy-number estimate obtained with genome-wide DNA SNP array with the standard FISH, for the

Table 1. Clinical parameters assessed at the time of diagnosis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of patients</th>
<th>Valid cases</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age $&gt;$60</td>
<td>232</td>
<td>327</td>
<td>71</td>
</tr>
<tr>
<td>Gender (m:f)</td>
<td>196:133</td>
<td>329</td>
<td>60:40</td>
</tr>
<tr>
<td>Binet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage A</td>
<td>234</td>
<td>327</td>
<td>72</td>
</tr>
<tr>
<td>Stage B or C</td>
<td>93</td>
<td>327</td>
<td>72</td>
</tr>
<tr>
<td>IGHV mutational status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGHV unmutated</td>
<td>127</td>
<td>326</td>
<td>39</td>
</tr>
<tr>
<td>IGHV mutated</td>
<td>199</td>
<td>326</td>
<td>61</td>
</tr>
<tr>
<td>CD38 $&gt;$30%</td>
<td>117</td>
<td>321</td>
<td>36</td>
</tr>
<tr>
<td>Zap 70 $&gt;$20%</td>
<td>112</td>
<td>284</td>
<td>39</td>
</tr>
<tr>
<td>Beta2-microglobulin $&gt;$UNV</td>
<td>137</td>
<td>223</td>
<td>61</td>
</tr>
<tr>
<td>LDT $&lt;$12 months</td>
<td>33</td>
<td>145</td>
<td>18</td>
</tr>
</tbody>
</table>

m, male; f, female; UNV, upper normal value; LDT, lymphocyte doubling time.
four loci investigated by the latter technique, in 329 CLL patients. Since FISH is considered the gold standard for the detection of prognostically relevant genomic aberrations in CLL samples, cases with a negative SNP array and a positive FISH were considered as false negative. As a whole, the percentage of SNP array ‘false negative’ was relatively low, affecting only between 2% and 8% of the whole CLL series (Figure 3).

Among the 18 cases with 17p loss by FISH, all the eight samples with a loss detected by SNP array in over 40% of the nuclei were also detected up by SNP array while half of the 10 cases with less than 40% of nuclei carrying the lesion were classified as normal by genomic profiling (supplementary Figure S1B, available at Annals of Oncology online).

Table 2. Large genomic aberrations spanning >5 Mb in patients with a normal FISH

<table>
<thead>
<tr>
<th>Region</th>
<th>Cytoband</th>
<th>Type</th>
<th>Candidate cancer genes</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1:225,493,922-235,530,710</td>
<td>1q42.12–q42.3</td>
<td>Loss</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>chr3:145,475,764-176,735,018</td>
<td>3q24–q26.32</td>
<td>Gain</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>chr6:87,049,197-125,406,247</td>
<td>6q14.3–q22.31</td>
<td>Loss</td>
<td>PRDM1</td>
<td>1</td>
</tr>
<tr>
<td>chr6:136,647,604-142,392,181</td>
<td>6q23.3–q24.1</td>
<td>Loss</td>
<td>TNFAIP3</td>
<td>1</td>
</tr>
<tr>
<td>chr7:125,307,489-141,917,586</td>
<td>7q31.33–q34</td>
<td>Loss</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>chr8:106,322,626-146,131,926</td>
<td>8q23.1–qter</td>
<td>Gain</td>
<td>MYC</td>
<td>1</td>
</tr>
<tr>
<td>chr8:182,272,38,003,609</td>
<td>8p23.3–p11.23</td>
<td>Loss</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>chr10:64,476,057-123,550,977</td>
<td>10q21.2–q26.13</td>
<td>Loss</td>
<td>PTEN</td>
<td>1</td>
</tr>
<tr>
<td>chr14:70,726,012-79,783,628</td>
<td>14q24.2–q31.1</td>
<td>Loss</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>chr18:51,454,691-77,917,301</td>
<td>18q21.2–q23</td>
<td>Gain</td>
<td>BCL2, NFATC1</td>
<td>2</td>
</tr>
<tr>
<td>chr19:37,187,908-59,107,865</td>
<td>19q13.12–q13.43</td>
<td>Gain</td>
<td>SPIB</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 1. Examples of three genomic profiles with genomic aberrations detected by SNP array in CLL patients with a normal FISH. Cases 08-515 presented multiple lesions on chromosomes 4q, 6p, 8q, 9p, 12q, 13q, 14q, 17q, 22q. Case 09-414 presented a single gain at 3q. Cases 08-560 presented a single-interstitial deletion at 4q. Black, raw copy-number values; red, smoothed copy-number values. X-axis, genomic mapping; Y-axis, log2 copy-number values.
have been examined by SNP array and by FISH or even due to false-positive FISH results. Real-time PCR carried out on a series of samples discordant in terms of the 13q14.3 status largely confirmed the SNP-array results (supplementary Figure S1C, available at *Annals of Oncology* online).

To ensure that the small discordances between the two detection methods did not influence the prognostic value of the assay, we applied the FISH-based prognostic model developed by Döhner et al. [29] to our cohort according to each method (Figure 4). The Kaplan–Meier curves were comparable between FISH and SNP array, both for OS as for TTT and the log-rank test was highly significant. The results were evaluated using a Cox multivariate regression model adjusted for Binet stage, age at the time of diagnosis >60 years, IGHV mutational status and the information regarding the aberrations identified either by FISH or by SNP array. In both the models and for both OS and TTT, the clinical characteristics and del17p emerged as independent prognosticators.

**discussion**

FISH is a powerful tool to predict the clinical course of patients affected by CLL [2, 3, 6]. However, this technique has some limitations and the SNP array provides additional prognostic information [8–13] and could represent an interesting comparison or, possibly, even an alternative technique to be implemented in the initial workup of CLL patients. In this article, in the so far largest comparison of FISH versus SNP array in CLL, we showed that SNP array was able to better define the prognosis of patients with a ‘normal’ FISH. Although SNP array provided results largely overlapping with FISH, it had the drawback of a reduced sensitivity in cases with aberrations present in only a low percentage of neoplastic cells, which can be easily seen by FISH.

An important benefit of SNP array over FISH is the ability to detect chromosomal aberrations not investigated by the standard FISH panel or even not targeted by any commercially available probes, but with a clinical relevance. In the present
study, large genomic aberrations spanning >5 Mb were detected by SNP array in 20% of the patients with a ‘normal’ FISH, which is in the absence of any of the four main genomic aberrations targeted by the probes included in the standard FISH panel (deletions at 13q14.3, 11q22, 17p13 and trisomy 12). Very importantly, the presence of at least one genomic lesion exclusively detected by SNP array did not appear to be associated with the presence of a mutated TP53 gene and determined a significantly shorter TTT, as also confirmed by multivariate analysis. Presumably due to the relatively indolent disease course, no statistically significant difference was noted in OS. With this regard, some limitations of our study have to be mentioned. Although based upon a very large series of cases, the number of patients actually bearing a normal FISH was relatively limited. Also, we have not validated the findings in an independent series, which would be ideally represented by a series of samples obtained from CLL patients enrolled in a prospective clinical trial.

Of note, the presence of such large aberrations in patients with a normal FISH was able to identify patients with a more aggressive clinical course among those with mutated IGHV genes. Interestingly, it has just been reported that also conventional karyotyping, when carried out using DSP30/IL2 stimulation, is able to detect high-risk patients among this class of individuals classified as intermediate risk by FISH [14]. These findings are in agreement with the prognostic role of genomic complexity in CLL, in which additional genomic aberrations would represent a clonal evolution possibly leading to a shorter survival [9, 10, 14].

TP53 gene resulted unmutated in cases with a normal FISH indicating that the phenotype we observed was unlikely due to TP53 inactivation. Only integration with large-scale mutational screening programs will reveal the relationship between these lesions and the presence of the recently reported somatic mutations targeting, among others, BIRC3, SF3B1, NOTCH1 and MYD88 [30–32].

The additional genomic lesions we detected did not occur at specific sites but were found across the whole genome. Nevertheless, some of the genes located in these regions, namely TNF, PRDM1 and TNFAIP3, MYC, CDKN2A, PTEN, BCL2 and NFATC1, are known or seem to play a role in lymphoid neoplasms, including CLL. Other lesions did not contain known candidate cancer genes (7q loss, 3q gain, 14q loss), but were recurrent and already previously reported [20]. It is possible that some of these regions have a yet unknown role in the pathogenesis of CLL, but their individually low frequency makes it difficult to identify a minimal common region and define candidate genes (as well as choosing appropriate FISH probes).

The comparison between SNP array and FISH revealed that SNP array was able to detect the four prognostically relevant FISH loci in nearly all cases with an overall concordance of 93%, in line with recent publications [10, 20, 21, 33]. In our series, the rate of samples called as normal by SNP array in the presence of a lesion detected by FISH (false negative) for 11q21 and 17p13 losses was nearly two-fold higher than for trisomy 12 or del 13q14.3. This result was expected, since these two lesions are known to be often subclonal, possibly present even in a very low percentage of tumor cells [34, 35]. In this study, the higher sensitivity of FISH was clearly evident, when looking at the cases with 17p1.3 loss defined by FISH: while all cases with a percentage of >40% positive nuclei were detected

Figure 4. OS and TTT in 329 CLL patients according to the Döhner prognostic model by FISH (A and B; P < 0.001) or by SNP array (C and D; P < 0.001).
by both FISH and SNP array, half of the cases with <40% of nuclei carrying the loss in FISH were negative by genomic profiling. In line with this, other authors have reported that array-based-karyotyping misclassified samples with a low percentage of neoplastic cells carrying a genomic aberration (up to 30%–40% of tumor cells in the sample) [15, 16, 20, 21]. This makes now impossible to fully replace the use of the standard FISH panel with SNP array in all the patients. On the other hand, it is of clinical importance that, although having a high sensitivity, FISH cannot identify all the CLL patients with a direct genetic inactivation of TP53: a percentage of patients have somatic mutations inactivating TP53 in the absence of 17p loss, and these cases have an outcome overlapping the ones with 17p deletion [26, 36]. Thus, the two techniques could be complementary in identifying poor-risk patients bearing a TP53 inactivation.

In line with previously published data [37], also false-positive FISH positive results were observed, mainly for the 13q14 region and the lesions detected by SNP array only and not by FISH were confirmed by real-time PCR.

Over the past decade, there have been important advances in the treatment of CLL [2], but the FISH-based prognostic model developed >10 years ago [6] is still largely applied. In order to validate the representativeness of our results in a clinical setting, we applied this model to our cohort using data generated by FISH or by SNP array. The Kaplan–Meier curves were comparable between FISH and SNP array, both for OS and for TTT, and the log-rank test was highly significant, also confirmed in multivariate analysis.

As the time and effort needed for both FISH and SNP array, as well their costs, are nowadays becoming less markedly different, a workflow could be suggested in which CLL samples would be first analyzed with SNP array. Due to the lower sensitivity of SNP array, in the absence of losses affecting the TP53 gene locus, an interphase FISH test, investigating only this specific locus, would be mandatory, since the inactivation of TP53 can have important clinical consequences when starting treatment [2, 3, 38]. Since DNA samples are routinely obtained for the assessment of the IGHV mutational status [2, 3], but also, more and more, for the investigation of TP53 gene somatic mutations [26, 36, 38], SNP array would not require an additional blood sampling from the patients. As currently done for FISH, genomic profiling could be repeated during the follow-up, to identify the occurrence of new lesions, which could change patients’ outcome. However, a prospective evaluation of this approach and its possible integration with the information derived from the mutational status of other genes besides TP53 (for example, BIRC3, SF3B1, NOTCH1) will be necessary. Moreover, it is important to stress that the current guidelines indicate the evaluation of chromosomal abnormalities at diagnosis or before treatment as mandatory only in the context of clinical trials and as ‘advisable’ in the general practice [2, 3, 38]. The testing for TP53 inactivation (by deletions or somatic mutations) can now be seen as the only genetic diagnostic assay that might affect the choice of the treatment, although there are still no data derived from randomized trials, demonstrating the superiority of a specific regimen for patients with inactivated TP53 [4, 38].

In conclusion, here, the identification of large genomic lesions detected by SNP array represented an independent prognostic factor in CLL patients with a normal FISH. SNP array was also able to identify the most important known prognostically relevant genomic aberrations of CLL. A validation in prospective trials is needed.

funding
This study was supported by Helmut Horton Foundation, Lugano, Switzerland; San Salvatore Foundation, Lugano, Switzerland; Nelia et Amadeo Barletta Foundation, Lausanne, Switzerland; Computational life science/Ticino in rete [no grant numbers]. MM was a recipient of fellowship from Alto Adige Bolzano-AIL Onlus.

disclosure
The authors have declared no conflicts of interest.

references