Annals of Oncology

Annals of Oncology 23: 2138–2146, 2012
doi:10.1093/annonc/mdr579
Published online 6 January 2012

Identification of a novel recurrent gain on 20q13 in chronic lymphocytic leukemia by array CGH and gene expression profiling

A. E. Rodríguez[1], C. Robledo[1], J. L. García[2], M. González[3], N. C. Gutiérrez[3], J. A. Hernández[4], V. Sandoval[5], A. García de Coca[6], I. Recio[7], A. Risueño[8], G. Martín-Núñez[9], E. García[10], R. Fisac[11], J. Conde[12], J. de las Rivas[8] & J. M. Hernández[1,3*]

[1]IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [2]Instituto de Estudios de Ciencias de la Salud de Castilla y León (IECSCYL)–HUSAL, Castill y León; [3]Department of Hematology, Hospital Clínico Universitario de Salamanca, Salamanca; [4]Department of Hematology, Hospital Infanta Leonor, Madrid; [5]Department of Hematology, Hospital Virgen Blanca, León; [6]Department of Hematology, Hospital Clínico Universitario, Valladolid; [7]Department of Hematology, Hospital Nuestra Señora de Sonsoles, Ávila; [8]Bioinformatics and Functional Genomics, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [9]Department of Hematology, Hospital Virgen del Puerto, Plasencia; [10]Genomics and Proteomics Unit, Centro de Investigación del Cáncer, Universidad de Salamanca-CSIC, Salamanca; [11]Department of Hematology, Hospital General de Segovia, Segovia; [12]Department of Hematology, Hospital del Río Hortega, Valladolid, Spain

Received 3 August 2011; revised 11 November 2011; accepted 16 November 2011

Background: The presence of genetic changes is a hallmark of chronic lymphocytic leukemia (CLL). The most common cytogenetic abnormalities with independent prognostic significance in CLL are 13q14, *ATM* and *TP53* deletions and trisomy 12. However, CLL displays a great genetic and biological heterogeneity. The aim of this study was to analyze the genomic imbalances in CLL cytogenetic subsets from both genomic and gene expression perspectives to identify new recurrent alterations.

Patients and methods: The genomic imbalances and expression levels of 67 patients were analyzed. The novel recurrent abnormalities detected with bacterial artificial chromosome array were confirmed by FISH and oligonucleotide microarrays. In all cases, gene expression profiling was assessed.

Results: Copy number alterations were identified in 75% of cases. Overall, the results confirmed FISH studies for the regions frequently involved in CLL and also defined a new recurrent gain on chromosome 20q13.12, in 19% (13/67) of the CLL patients. Oligonucleotide expression correlated with the regions of loss or gain of genomic material, suggesting that the changes in gene expression are related to alterations in copy number.

Conclusion: Our study demonstrates the presence of a recurrent gain in 20q13.12 associated with overexpression of the genes located in this region, in CLL cytogenetic subgroups.

Key words: CLL, cytogenetic aberrations, gene expression profile, genomic arrays

introduction

Chronic lymphocytic leukemia (CLL) is the most common leukemia in the western world and is characterized by a highly variable clinical course with survival times ranging from months to decades despite a remarkable phenotypic homogeneity [1, 2]. This clinical heterogeneity reflects its biological diversity [3]. Our understanding of the biology of CLL has helped to identify several markers of prognostic significance, delineating CLL into several distinct diseases. These markers include cytogenetic abnormalities, the mutational status of the immunoglobulin heavy chain variable (*IGHV*) and ZAP-70, CD38 and CD49d expression [2, 4–6]. Conventional cytogenetic analyses have revealed chromosomal aberrations in 40%–50% of patients, but detection of abnormalities is limited by the low mitotic activity of CLL cells. By contrast, interphase FISH (iFISH) has identified chromosomal changes in ∼80% of patients with CLL, the presence of specific chromosomal abnormalities being a prognostic indicator of disease progression and survival [2, 7]. Thus, half of the CLL patients carry deletions of 13q, which is correlated with an indolent disease course in patients with this abnormality as their sole aberration. In contrast, deletions of 11q and 17p (which cover the *ATM* and *TP53* genes, respectively) have a poorer outcome. Furthermore, trisomy 12 is related to an intermediate prognosis, whereas deletion of 6q has been identified as a recurrent CLL progression marker [8]. In addition, great genomic complexity has been associated with

*Correspondence to: Prof. J. M. Hernández, Hematology Unit, Department of Medicine, Hospital Universitario de Salamanca, Paseo San Vicente 58, 37007 Salamanca, Spain. Tel: +34-923-291-100; Fax: +34-923-294-624; E-mail: jmhr@usal.es

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology.
All rights reserved. For permissions, please email: journals.permissions@oup.com

worse survival and is also closely related to markers of poor prognosis [9–11].

Considering the great heterogeneity of CLL from both genetic and prognostic points of view, microarray technology is a powerful tool for the analysis of genetic alterations in CLL. Thus, comparative genomic hybridization using high-density arrays, array comparative genomic hybridization (aCGH), allows high-resolution genome-wide scan for detection of copy number alterations in a single hybridization and aCGH using bacterial artificial chromosome (BAC) clones has been widely applied in the analysis of hematological malignancies [12–15]. Regarding oligonucleotide microarrays, the study of the gene expression profile (GEP) in CLL has given us insights into the molecular mechanisms involved in its pathogenesis by analyzing the impact of genomic aberrations on the expression of genes located on the corresponding loci [16–18].

Although the application of microarray technology in CLL has provided additional knowledge of the known recurrent aberrations as well as enabling novel aberrations, such as gain of 2p and deletion of 22q to be identified [19–23], to date, few studies have investigated genomic aberrations specifically in relation to CLL cytogenetic subsets. Therefore, the aim of this study was not only to screen and identify new genomic events in CLL patients but also to compare the prevalence of these genomic aberrations in cytogenetic CLL subsets. Furthermore, our data revealed an association between altered transcription levels and genomic imbalances in the genetic subsets of CLL, indicating that gene dosage might have pathogenic effects in CLL and delineate a new gained region, on 20q13, in CLL patients.

methods

patients

Peripheral blood samples from 67 patients with CLL were analyzed. The diagnoses were confirmed by standardized clinical, morphological and immunological data according to the World Health Organization classification and the criteria of the Working Group of the National Cancer Institute [24]. FISH studies and *IGH* mutational status were determined in all patients. The study protocol was approved by the local ethical committees and prior written informed consent was obtained from the patients. All patients were untreated and most of them were studied at the moment of diagnosis (Table 1). The main characteristics of the 67 CLL patients included in the study are reported in Table 2.

FISH analysis

Interphase FISH was carried on all the samples using commercially available probes for the following regions: 13q14, 12q13, 11q22/*ATM*, 17p13/*TP53* and 14q32/*IGH* (Abbott Co., Downers Grove, IL) using the previously described methods [25].

To confirm the gains and losses assessed by aCGH, FISH analysis was done using Vysis LSI ZNF217, the commercially available probe for 20q13.2 (Abbott Co.) and the BAC clones dJ1028D15–dJ781B1, mapping to 20q13.12, as previously described [25]. The clones were located in the same region of gain as detected by aCGH and were selected from the aCGH BAC clone library (Wellcome Trust Sanger Institute), whereas the commercial probe was located in 20q13.2 (breast tumor amplicon). DNA from the BAC clones was isolated, labeled and hybridized, as previously described [26]. The changes were validated in fixed cells from the same diagnostic samples as used for aCGH ($n = 20$).

FISH analysis was carried out on 400 interphase cells using standard fluorescent microscopy.

mutation status of IGVH genes

IGVH genes were amplified and sequenced according to the ERIC recommendations on *IGHV* gene mutational status analysis in CLL [27].

Table 1. Status of disease in the total series ($n = 67$) and in + 20q CLL patients ($n = 13$)

	CLL patients ($n = 67$) *n* (%)	+ 20q CLL patients ($n = 13$) *n*/CLL (%)
At diagnosis	50 (75)	13/50 (26)
Progressive	17 (25)	4/17 (24)

CLL, chronic lymphocytic leukemia.

Table 2. Clinical and molecular characteristics of the CLL patients

Characteristics	(%)
Median age in years (range)	68 (35–90)
Male/Female (ratio)	73/23 (2.7)
White blood cells, range/ml	39 000 (7600–175 000)
Lymphocytes/ml (range)	32 000 (5000–160 000)
Hemoglobin, g/dl (range)	13.6 (7.1–16.3)
Platelet count/ml (range)	167 000 (59 000–306 000)
LDH	
Normal	82
High	12
β_2-microglobulin	
Normal	52
High	48
Status of the disease	
At diagnosis	71.6
Progressive	28.4
Binet stage	
A	66
B	26
C	9
ZAP-70 expression	
Positive	44
Negative	56
CD38 expression	
Positive	26
Negative	74
IgVH mutational status	
Mutated	41
Unmutated	59
Interphase FISH analysis	
Normal karyotype	22
13q deletion	37
Trisomy 12	16
11q deletion	7
17p deletion	9
IGH translocation	9
20q13.12 gain	19

CLL, chronic lymphocytic leukemia; LDH, lactate dehydrogenase.

array comparative genomic hybridization

BAC arrays. DNA samples were analyzed using a BAC array containing 3523 sequence-validated BACs covering the genome with a mean resolution of 1 Mb, as previously described [26].

oligonucleotide microarrays. In order to confirm the results of the BAC aCGH analysis, a subset of 35 patients were analyzed using a NimbleGen Human CGH 4 × 72K Whole Genome v2.0 array (Roche Diagnostics, Mannheim, Germany).

The complete description of BAC and oligonucleotide microarrays experiments is available as supplementary Material (available at *Annals of Oncology* online).

GEP analysis

RNA isolation, labeling and microarray hybridization were carried out, as previously reported [28]. The GEP was analyzed in all cases with Human Genome U133A microarray (Affymetrix, Santa Clara, CA). Data analysis is available as supplementary material (available at *Annals of Oncology* online).

comparative analysis of CGH and expression arrays

In order to achieve a comparative analysis of the copy number changes, from the CGH arrays, and the gene changes, from the expression arrays, for the same patients, we select the patients who showed significant gains in the aCGH data and their corresponding expression data. We normalized the expression dataset using the R package GeneMapper [29] that allows an accurate assignment to ENSEMBL genes (instead to Affymetrix probesets) including their location in the genome. Following this, we selected the three regions in chromosome 20q where the gains detected by aCGH were significative. For such regions, we calculate, in the corresponding samples, the mean and median expression signal based on the genes included. On these expression numbers, we carried out a statistical one-tail *t*-test (using R) to check if there was a significant correlation between the aCGH gain observed in the 20q regions and the overexpression of the genes included in such regions.

statistical analysis

Two-tailed Chi-square and Fisher's exact tests were used to analyze the associations between variables. For all tests, values of $P < 0.05$ were considered to indicate statistical significance. The calculations were carried out using SPSS 17.0 for Windows (SPSS Inc., Chicago, IL).

results

FISH and mutational status

FISH analyses revealed that 25 of the 67 cases analyzed (37%) carried the 13q14 deletion and that this was the only abnormality in 20 patients (30%). Overall, the 11q22.3 and 17p13.1 deletions were present in 5 (7%) and 6 (9%) patients, respectively, while trisomy 12 was present in 11 (16%) and t (14q32) in 9 (13%) patients. The remaining 21 (31%) patients did not show aberrations by FISH. To better characterize the 20q13.12 gain, FISH analysis was also carried out in a validation series of 58 patients: 17% patients showed this alteration in ≥4.5% cells (ranging from 4.5% to 12%). In relation to mutational status, 49% of cases had unmutated *IgVH* gene.

aCGH showed recurrent genomic imbalances in CLL

Fifty of 67 patients (75%) displayed genomic changes with aCGH. In addition to the regions detected by FISH abnormalities, aCGH enabled the presence of novel recurrent genomic imbalances to be demonstrated. In order to rule out previously described single nucleotide polymorphisms, the minimal regions of overlap for all the recurrent lesions were

Table 3. Recurrent aberrations identified by aCGH in CLL

Chromosome	Cytoband	Start position	End position	Size	Frequency of gains	Frequency of losses	Number of genes covered by the aCGH
1	q21.3–q22	151208615	155645185	4.44	22		31
1	q31.1–q31.2	189804292	192389810	2.59	13		1
5	q13.3–q14.1	74766678	77235911	2.47		6	18
5	q31.1	131568885	132631297	1.06		6	21
6	p21.31–p21.1	33722842	43897461	10.17	19		146
7	q22.1	98853924	101547956	2.69		6	78
10	q22.3	80565766	81502183	0.94	7.5		2
11	q13.1	64268965	65470124	1.2	22		51
11	q13.3–q13.4	69974549	73001497	3.03	21	16	37
11	q22.3–q23.1	108518932	112964950	3.73		9	24
12	p13.33–q24.33	1	133851895	133.9	15		
13	q14.2–q14.3	48863579	54941189	6.08		21	27
16	q23.2–q24.2	80972437	88692209	7.72	21		56
17	p13.2–p13.1	4571828	7483888	2.91		10	104
17	q25.3	75993754	80470659	4.48	19		105
18	q21.2	49844734	50270501	0.43	9	7.5	5
20	q13.12	42188467	44495323	2.31	19		52

Positions and sizes are expressed in base pairs. Bp locations according to GRCh37, February 2009 (hg 19).
aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

compared with the frequencies of known copy number variations. A total of 443 altered chromosomal regions were found, of which 237 (53%) were deletions. The median number of changes per patient was five (range 0–14). The most commonly recurring alterations (observed in >5% of cases), their boundaries and frequencies are shown in Table 3. Losses in 13q14.2–q14.3 (21%), 11q13.3 (16%), 17p13.2–p13.1 (10%), 11q22.3–q23.1 (9%) and 5q13.3–q14.1, 5q31.1 and 7q22 (6% each) as well as gains in 1q21.3–q22 (22%), 11q13.3 (21%), 16q23.2–q24.2 (21%), 6p21.31–p21.1 (19%) and 10q22.3 (7.5%) were the most frequent changes revealed by aCGH (Table 3). Interestingly, a critical segment of gain was delineated on chromosome 20q in 13 patients (19%). The analysis identified a minimal region of gain on 20q13.12 of ∼2.31 Mb involving three clones at linear positions (42 188 467–44 495 323), as shown in Figure 1. Most of these cases (75%) were studied at the time of diagnosis (Table 1). Changes detected in diagnostic and progression groups are shown in Table 4.

genomic abnormalities in the cytogenetic subgroups of CLL

Overall, correlation between FISH and aCGH was observed for + 12, 11q- or 17p- cases (100%, 91% and 83%, respectively). However, this was not the case for 13q- subgroup. Interestingly, most of these cases displayed <30% of 13q deletions and in the 14 cases with a deletion of 13q revealed by aCGH, losses were located in 13q14.2–q21.1, with heterogeneous breakpoints. It should be noted that the 21 CLL samples showing no aberrations with FISH also appeared to be normal for the CLL FISH regions when analyzed by aCGH. However, novel recurrent alterations by aCGH were detected in this group of CLL patients (Table 5).

The analysis of the relationship between the recurrent abnormalities revealed by FISH and the presence of novel chromosomal imbalances detected by aCGH showed a significant association between the loss of 13q and the loss of 5q13.3–q14.1 and 5q31.1 ($P < 0.05$). No other additional abnormalities were observed in any of the cytogenetic subgroups.

In order to assess the genomic complexity in the cytogenetic CLL subsets, a comparison between the number of genetic changes ascertained by aCGH and the FISH CLL subgroup was carried out. As the median of changes per patient was five (range 0–14), we defined two groups to analyze the number of changes with respect to the FISH categories: ≤5 (low genomic complexity) and >5 (high genomic complexity) (Table 6). Interestingly, an association between the presence of a large number of changes detected by aCGH and ATM deletion ($P = 0.026$) by FISH was observed. The presence of gains on 20q12.13 ($P = 0.002$) was also associated with a high frequency of changes as revealed by aCGH (Table 6).

oligonucleotide and FISH studies validated the changes observed by aCGH

Oligonucleotide aCGH was carried out in 35 cases to confirm the BAC array results. Genomic patterns of gains and losses representative of the probe sizes (∼150 kb) were compared

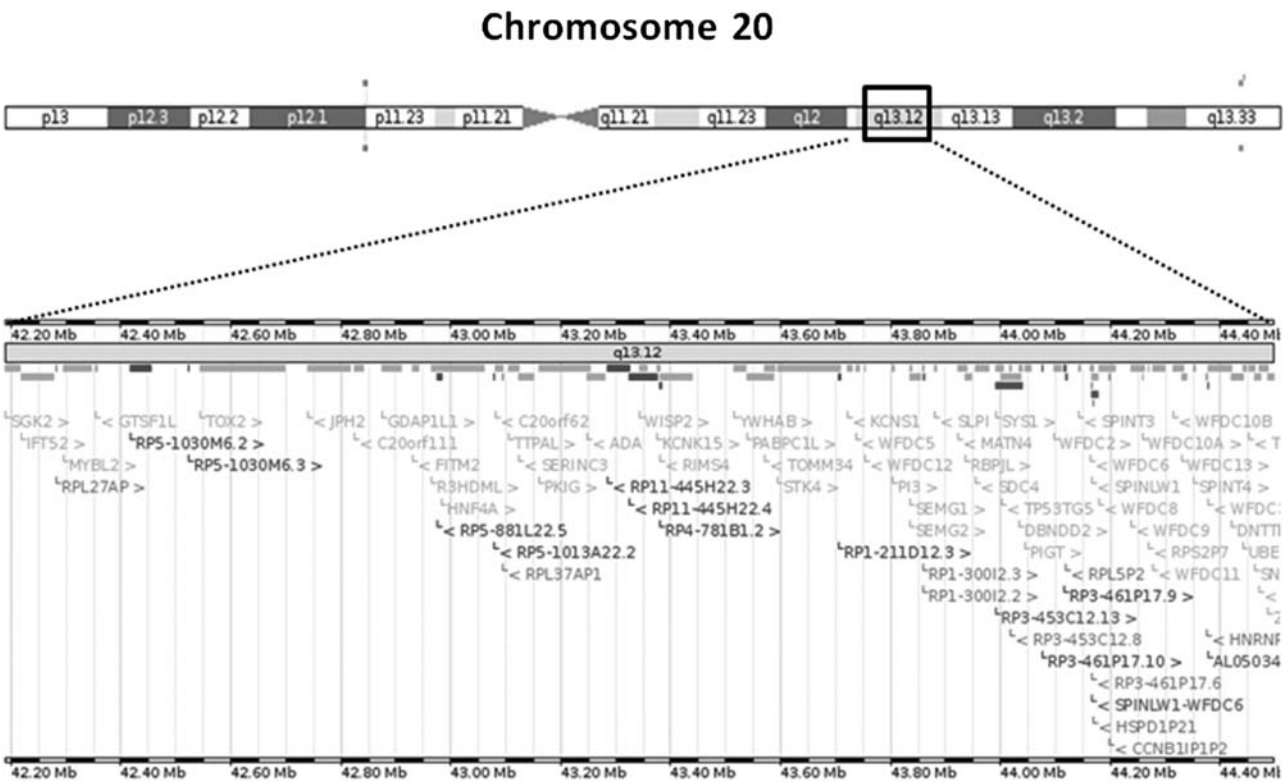

Figure 1. Integration of annotated genomic sequence with array comparative genomic hybridization data: common region of gain (CRG) on 20q (42188467–44495323 bp) showing the candidate genes (GRCh37, February 2009, hg 19).

Table 4. Characteristics of the CLL series (IgVH mutational status, number of aberrations, FISH subgroup and frequency of recurrent alterations detected by aCGH) according to the status of disease (at diagnosis versus progression)

Characteristics	Status of disease	
	At diagnosis (%)	Progression (%)
IgVH mutational status		
Mutated	63.6	46.7
Unmutated	36.4	53.3
Number of aberrations		
≤5	64	52.9
>5	36	47.1
FISH subgroup		
Normal FISH	38[a]	11.8
13q deletion	40	29.4
Trisomy 12	8	41.2[a]
17p deletion	6	17.6
11q deletion	6	11.8
t(14q32)	16	5.9
Recurrent alteration by aCGH		
Gains		
1q21.3–q22	24	17.6
1q31.2	8	29.4
6p21.31–p21.1	18	23.5
10q22.3	6	11.8
11q13.1	24	17.6
11q13.3	20	23.5
12	6	41.2[a]
16q23.2–q24.2	22	17.6
17q25.3	20	17.6
18q21.2	6	17.6
20q13.12	20	17.6
Losses		
5q13.3–q14.1	8	0
5q31.1	8	0
7q22	4	11.8
11q13.3	16	17.6
11q22.3–q23.1	8	11.8
13q14.2–q14.3	24	11.8
17p13.2–p13.1	10	11.8
18q21.2	8	5.9

[a]Statistically significant associations ($P < 0.05$).
aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

with those obtained by BAC aCGH and found to be 100% concordant.

FISH experiments were carried out on 20 patients to confirm the gains on 20q13.12 observed with aCGH (supplemental Figure S1, available at *Annals of Oncology* online). All but one of the cases (95%) was concordant with the aCGH results. The median of cells showing this aberration was 20% (range 16%–25%). In addition, the cases were analyzed with a probe covering 51 992 266–52 410 801 bp (Vysis LSI ZNF217, breast tumor amplicon at 20q13.2). The results failed to show any involvement of this region, delineating the commonly gained region at 20q between 42 188 467–44 495 323 bp (2.31 Mb) (Figure 1).

gene expression profile confirmed the dosage effect of aCGH changes

In order to assess the relevance of the genomic imbalances in gene expression, a gene expression profile study was carried out. For this purpose, we grouped the cases by aCGH findings. The group of patients displaying trisomy 12 showed deregulation of 89 genes when compared with the rest of patients. A total of 76 of the 89 genes were overexpressed in relation to the other patients and 56% of them were located on chromosome 12.

It should be noted that overexpression of the 52 genes located in 20q13.12 (Figure 1), the 20q region gained by aCGH, was also observed ($P = 0.01$). Among these genes, we found well-known protein-coding cancer-related genes (supplemental Table S1, available at *Annals of Oncology* online) such as *PI3* (elafin), *SLPI* (secretory leukocyte peptidase inhibitor) and *WFDC2* [whey acidic protein (WAP) four-disulfide core domain 2], members of the WAP family; *PIGT* (phosphatidylinosotol glycan anchor biosynthesis, class T), a component of the glycosylphosphatidylinositol (GPI) glycan transamidase complex; *HNF4A* (hepatocyte nuclear factor 4, alpha) and *YWHAB* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide), members of the SMAD and Ras signal transduction pathways, respectively. In addition, *ADA* (adenosine deaminase), a regulator of B-cell proliferation, overexpression was also present in CLL cases with gains on 20q.

Moreover, in patients with the 17p13 deletion, a significant proportion (83%) of the differentially underexpressed genes clustered in this region ($P < 0.05$). Among the downregulated genes were *GPS2* (G protein pathway suppressor 2)/*AMF1*, *SGSM2* (small G protein signaling modulator 2), *DRG2* (developmentally regulated GTP-binding protein 2), *SAT2* (spermidine/spermine N1-acetyltransferase family member 2) and *C17orf49* (chromosome 17 open reading frame 49). This gene dosage effect was also observed in CLL patients showing 11q-. Thus, all the genes located in the minimal region of deletion observed with aCGH on 11q22.3–q23.2 (108518932–112964950 bp) were downregulated when compared with the rest of patients ($P < 0.01$).

discussion

The presence of cytogenetic abnormalities is a hallmark of CLL. Indeed, these abnormalities have been associated with the prognosis or progression of the disease and for this reason the genetic changes have been extensively studied in CLL [30, 31]. The present study integrates genomic and gene expression profile analyses in a cohort of 67 CLL patients. Overall, the results enable us to detect hitherto undescribed recurrent alterations in CLL, such as gains on chromosome 20 and confirm the dosage effect, not only for the common cytogenetic abnormalities but also for this new genetic abnormality.

The present study found genomic copy number changes in 75% of the CLL patients. Our findings are similar to those previously reported in this disease [19, 22, 23, 32]. Detection rates of genomic alterations involving loci known to be

Table 5. Correlation between the most frequent chromosomal imbalances identified by aCGH and the CLL cytogenetic subgroups

Aberration/%	Cytogenetic/FISH subgroup						Total (%)
	13q14.3	Trisomy 12	11q22.3	17p13.1	t(14q32)	Normal FISH	
1q21.3–q22							
Gain	28	27	20		11	24	22
1q31.1–q31.2							
Gain	8	27	40	17	22		13
5q13.3–q14.1							
Loss	16[a]			17	11		6
5q31.1							
Loss	16[a]			17	11		6
6p21.31–p21.1							
Gain	24	27	20		22	14	19
7q22.1							
Loss	12	9	20		11		6
10q22.3							
Gain	16	9	20		11		7.5
11q13.1							
Gain	28	18	20	33	22	19	22
11q13.3–q13.4							
Loss	12	18	60[a]		22	14	16
Gain	28	18		17	22	19	21
11q22.3–q23.1							
Loss		18	100			5	9
12							
Gain		91					15
13q14.2–q14.3							
Loss	56			17	11		21
16q23.2–q24.2							
Gain	24	27	20	17	22	19	21
17p13.2–p13.1							
Loss	16			83			10
17q25.3							
Gain	24	9	20	17	22	14	19
18q21.2							
Loss	16			33			7.5
Gain	12	18		17		9	12
20q13.12							
Gain	24	27	20		22	14	19

Results are expressed as percentages.

[a]Statistically significant associations ($P < 0.05$).

aCGH, array comparative genomic hybridization; CLL, chronic lymphocytic leukemia.

associated with CLL occurred at expected frequencies [33] and overall, correlation between FISH and aCGH was observed except in the 13q- subgroup. Both FISH and aCGH revealed that 13q- was an heterogeneous group in size of the deletion and percentage of cells displaying the abnormality. Interestingly, when aCGH failed to demonstrate the presence of 13q deletion, FISH data revealed that most of these cases had <30%. This could justify, at least in part, the lack of correlation between both the techniques. We also confirmed that deletions are more abundant than gains in CLL: deletions in chromosomes 5, 7, 11, 13, 17 and 18 and gains in chromosomes 1, 6, 10, 11, 12, 16, 17, 18 and 20 were present in this series. Regarding other recently reported alterations, we observed gain on 2p [34] in one case.

Our study identifies a previously undescribed recurrent region of gain in CLL, located on 20q13 in 19% of CLL patients. This frequency is similar to other well-characterized abnormalities in CLL (+ 12, 11q- and 17p-). It should be noted that gains in 20q were not associated with any other cytogenetic abnormality, although no patients with loss on 17p displayed 20q gains. The presence of 20q gains was not associated with mutational status either. Abnormalities of chromosome 20 are frequently observed aberrations in cancer [35–37]. In addition, the presence of gains on 20q has been associated with aggressive tumor behavior and poor clinical prognosis [38]. By contrast, deletions of the long arm of chromosome 20 are a common chromosomal abnormality associated with myeloid malignancies and are rarely seen in lymphoid malignancies [39]. A detailed analysis of 20q gains in cancer revealed that the size and location of the alteration are both variable. A region of gain at 20q13 was identified in CGH studies in human breast tumors [40]. The region has been analyzed at higher resolution,

Table 6. Number of changes per patient in FISH groups and 20q13.12 cases

FISH	Number of aberrations (% cases)		Median of changes
	≤5	>5	
13q14.3	10 (40)	15 (60)	6
Trisomy 12	5 (45)	6 (55)	3
11q22.3	0	5[a] (100)	9
17p13.1	3 (50)	3 (50)	4
t(14q32)	5 (56)	4 (44)	3
Normal FISH	20[a] (95)	1 (5)	3
20q13.12 gain	3 (23)	10[a] (77)	7
Total (%)	4	33	

[a]Statistically significant associations ($P < 0.05$).

enabling three independently amplified regions to be characterized, with 20q13.2 being the most common region of gain in breast cancer. In the present study, FISH studies identified a minimal region of gain on 20q13.12 of ~2.31 Mb. This region is located close to the 20q breast cancer amplicon but is not included in it.

The gain on 20q13 in CLL could be relevant to the pathogenesis and evolution of CLL because 11 protein-coding cancer-related genes have been identified in this region (supplemental Table S1, available at *Annals of Oncology* online). It should be noted that all of these genes were upregulated in the CLL patients showing 20q gains in comparison with the other CLL cases. Thus, *PIGT*, *PI3*, *SLPI* and *WFDC2* could be potential candidate genes since they have been previously related to progression or tumor invasion. Phosphatidylinositol glycan (PIG) class T (*PIG-T*) is a component of the GPI transamidase complex and is amplified and overexpressed in human breast cancer cell lines and primary tumors [41]. Previous studies suggested that activation of the GPI transamidase complex could be a molecular mechanism underlying the progression of various human cancers [41, 42]. Interestingly, *GIP-S*, another GPI subunit, is located on 17p13.2, a region frequently deleted in cancer and in CLL. Therefore, further studies of these genes and their biological effects of all GPI transamidase complex subunits could be relevant in CLL. *PI3*, *SLPI* and *WFDC2* are members of the WAP family, a group of genes coding for proteins with a WAP motif. All of them have been identified as molecular markers for cancer and are clustered on chromosome 20q12–13.1. These genes are amplified and upregulated in several cancers [43]. The expression levels of all these genes were significantly higher in CLL cases with gains on 20q. Therefore, we suggest that 20q13.12 overexpressed genes may also be important in the evolution of CLL and warrant detailed study.

The present study also revealed a gene dosage effect in other chromosomal regions. Thus, CLL patients with trisomy 12 overexpressed genes located on chromosome 12, while patients with losses on 17p underexpressed genes located on 17p, as previously reported [16–18].

Gains in 20q13 in CLL did not occur as a single aberration because all CLL patients with gains in this region also had additional genetic changes. In fact, gains on 20q were associated with genomic complexity (Table 6). It is of note that genomic complexity has a significant impact on cancer prognosis and a number of studies have described the presence of several genomic changes as being predictors of disease progression and chemosensitivity in CLL [9, 44]. A significantly high level of genomic complexity in patients with loss on 11q was also observed. However, the CLL patients with losses on 17p did not have a large number of genomic alterations. This observation may indicate that the poor prognosis of patients with CLL exhibiting loss on 17p is unrelated to their genomic complexity [9]. The presence of a large number of genomic alterations in 20q13-gain patients suggest that this new genetic entity could be associated with a more advanced disease in CLL, as has been suggested in non-Hodgkin's lymphomas [45]. Genomic instability could therefore be another molecular feature of CLL progression, as has recently been suggested [46]. In order better to assess the clinical value of gain on 20q, a prospective study in a large series of CLL patients needs to be carried out.

Our results failed to demonstrate the presence of recurrent secondary genetic imbalances in the cytogenetic subgroups. In fact, only the group of patients with losses in 13q showed an association with losses in 5q13.3–q14.1 and 5q31. These changes had not been previously reported and could be examined further in subsequent studies.

In summary, our results demonstrated that submicroscopic 20q13.12 gains are common in CLL and confirmed that these gains result in an overexpression of the genes located on 20q13 [Figure 1, supplemental Table S1 (available at *Annals of Oncology* online)]. Furthermore, 20q gain is associated with great genomic complexity. These results suggest that the diversity of genomic aberrations in CLL is much greater than previously suggested. Further studies are needed to assess the prognostic significance of these alterations and how the genes located in these loci could contribute to the pathogenesis of CLL.

acknowledgements

We thank N Carter and H Fiegler (Sanger Center, Cambridge, UK) for providing us with the BACs library. We thank Irene Rodríguez, Sara González, Teresa Prieto, M[a] Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo and Vanesa Gutiérrez of the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance.

funding

This work was partially supported by grants from the Spanish Fondo de Investigaciones Sanitarias (02/1041 and FIS 09/01543); Fondo Social Caja de Burgos de Investigación Clínica, Proyectos de investigación del SACYL (106/A/06) and by the 'Acción Transversal del Cáncer' project, through an agreement between the Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Science and Innovation and the Cancer Research Foundation of Salamanca University and the Redes de Investigación RTIIC (FIS). AER is fully supported by an

'Ayuda predoctoral FIS de formación en investigación' by the Spanish Fondo de Investigaciones Sanitarias.

disclosure

The authors declare no conflicts of interest.

references

1. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005; 352: 804–815.
2. Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000; 343: 1910–1916.
3. Rozman C, Montserrat E. Chronic lymphocytic leukemia. N Engl J Med 1995; 333: 1052–1057.
4. Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999; 94: 1840–1847.
5. Orchard JA, Ibbotson RE, Davis Z et al. ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. Lancet 2004; 363: 105–111.
6. Rassenti LZ, Huynh L, Toy TL et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. N Engl J Med 2004; 351: 893–901.
7. Hernandez JA, Rodriguez AE, Gonzalez M et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. Haematologica 2009; 94: 364–371.
8. Finn WG, Kay NE, Kroft SH et al. Secondary abnormalities of chromosome 6q in B-cell chronic lymphocytic leukemia: a sequential study of karyotypic instability in 51 patients. Am J Hematol 1998; 59: 223–229.
9. Kujawski L, Ouillette P, Erba H et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. Blood 2008; 112: 1993–2003.
10. Kipps TJ. Genomic complexity in chronic lymphocytic leukemia. Blood 2008; 112: 1550.
11. Grubor V, Krasnitz A, Troge JE et al. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). Blood 2009; 113: 1294–1303.
12. de Leeuw RJ, Davies JJ, Rosenwald A et al. Comprehensive whole genome array CGH profiling of mantle cell lymphoma model genomes. Hum Mol Genet 2004; 13: 1827–1837.
13. Kohlhammer H, Schwaenen C, Wessendorf S et al. Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. Blood 2004; 104: 795–801.
14. Rubio-Moscardo F, Climent J, Siebert R et al. Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome. Blood 2005; 105: 4445–4454.
15. Tyybakinoja A, Saarinen-Pihkala U, Elonen E et al. Amplified, lost, and fused genes in 11q23-25 amplicon in acute myeloid leukemia, an array-CGH study. Genes Chromosomes Cancer 2006; 45: 257–264.
16. Haslinger C, Schweifer N, Stilgenbauer S et al. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. J Clin Oncol 2004; 22: 3937–3949.
17. Porpaczy E, Bilban M, Heinze G et al. Gene expression signature of chronic lymphocytic leukaemia with trisomy 12. Eur J Clin Invest 2009; 39: 568–575.
18. Dickinson JD, Joshi A, Iqbal J et al. Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. Int J Mol Med 2006; 17: 769–778.
19. Pfeifer D, Pantic M, Skatulla I et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood 2007; 109: 1202–1210.
20. Gunn SR, Bolla AR, Barron LL et al. Array CGH analysis of chronic lymphocytic leukemia reveals frequent cryptic monoallelic and biallelic deletions of chromosome 22q11 that include the PRAME gene. Leuk Res 2009; 33: 1276–1281.
21. Patel A, Kang SH, Lennon PA et al. Validation of a targeted DNA microarray for the clinical evaluation of recurrent abnormalities in chronic lymphocytic leukemia. Am J Hematol 2008; 83: 540–546.
22. Schwaenen C, Nessling M, Wessendorf S et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A 2004; 101: 1039–1044.
23. Tyybakinoja A, Vilpo J, Knuutila S. High-resolution oligonucleotide array-CGH pinpoints genes involved in cryptic losses in chronic lymphocytic leukemia. Cytogenet Genome Res 2007; 118: 8–12.
24. Binet JL, Caligaris-Cappio F, Catovsky D et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. Blood 2006; 107: 859–861.
25. Gonzalez MB, Hernandez JM, Garcia JL et al. The value of fluorescence in situ hybridization for the detection of 11q in multiple myeloma. Haematologica 2004; 89: 1213–1218.
26. Robledo C, Garcia JL, Caballero D et al. Array comparative genomic hybridization identifies genetic regions associated with outcome in aggressive diffuse large B-cell lymphomas. Cancer 2009; 115: 3728–3737.
27. Ghia P, Stamatopoulos K, Belessi C et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia 2007; 21: 1–3.
28. Gutierrez NC, Lopez-Perez R, Hernandez JM et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005; 19: 402–409.
29. Risueno A, Fontanillo C, Dinger ME et al. GATExplorer: genomic and transcriptomic explorer; mapping expression probes to gene loci, transcripts, exons and ncRNAs. BMC Bioinformatics 2010; 11: 221.
30. Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet 2008; 40: 1204–1210.
31. Crowther-Swanepoel D, Broderick P, Di Bernardo MC et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet 2010; 42: 132–136.
32. Ouillette P, Erba H, Kujawski L et al. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14. Cancer Res 2008; 68: 1012–1021.
33. Gunn SR, Mohammed MS, Gorre ME et al. Whole-genome scanning by array comparative genomic hybridization as a clinical tool for risk assessment in chronic lymphocytic leukemia. J Mol Diagn 2008; 10: 442–451.
34. Jarosova M, Urbankova H, Plachy R et al. Gain of chromosome 2p in chronic lymphocytic leukemia: significant heterogeneity and a new recurrent dicentric rearrangement. Leuk Lymphoma 2010; 51: 304–313.
35. Yang SH, Seo MY, Jeong HJ et al. Gene copy number change events at chromosome 20 and their association with recurrence in gastric cancer patients. Clin Cancer Res 2005; 11: 612–620.
36. Zhu H, Lam DC, Han KC et al. High resolution analysis of genomic aberrations by metaphase and array comparative genomic hybridization identifies candidate tumour genes in lung cancer cell lines. Cancer Lett 2007; 245: 303–314.
37. Lassmann S, Weis R, Makowiec F et al. Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas. J Mol Med (Berl) 2007; 85: 293–304.
38. Bar-Shira A, Pinthus JH, Rozovsky U et al. Multiple genes in human 20q13 chromosomal region are involved in an advanced prostate cancer xenograft. Cancer Res 2002; 62: 6803–6807.
39. Bench AJ, Nacheva EP, Hood TL et al. Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. UK Cancer Cytogenetics Group (UKCCG). Oncogene 2000; 19: 3902–3913.
40. Tanner MM, Tirkkonen M, Kallioniemi A et al. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. Cancer Res 1996; 56: 3441–3445.

41. Wu G, Guo Z, Chatterjee A et al. Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunits phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. Cancer Res 2006; 66: 9829–9836.
42. Scotto L, Narayan G, Nandula SV et al. Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. Genes Chromosomes Cancer 2008; 47: 755–765.
43. Clauss A, Lilja H, Lundwall A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. Biochem J 2002; 368: 233–242.
44. Kay NE, Eckel-Passow JE, Braggio E et al. Progressive but previously untreated CLL patients with greater array CGH complexity exhibit a less durable response to chemoimmunotherapy. Cancer Genet Cytogenet 2010; 203: 161–168.
45. Carter SL, Eklund AC, Kohane IS et al. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006; 38: 1043–1048.
46. Stephens PJ, Greenman CD, Fu B et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011; 144: 27–40.

Annals of Oncology 23: 2146–2153, 2012
doi:10.1093/annonc/mdr576
Published online 4 January 2012

Identification and validation of gene expression models that predict clinical outcome in patients with early-stage laryngeal cancer

E. Fountzilas[1], K. Markou[2], K. Vlachtsis[2], A. Nikolaou[2], P. Arapantoni-Dadioti[3], E. Ntoula[3], G. Tassopoulos[4], M. Bobos[5], P. Konstantinopoulos[1], G. Fountzilas[6] & D. Spentzos[1]*

[1]Division of Hematology/Oncology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, USA; [2]Department of Otorhinolaryngology, "AHEPA" Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki; [3]Department of Pathology; [4]Department of Otorhinolaryngology, "Metaxa" Cancer Hospital, Piraeus; [5]Laboratory of Molecular Oncology, Hellenic Foundation for Cancer Research, Thessaloniki; [6]Department of Medical Oncology, "Papageorgiou" Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece

Received 12 February 2011; revised 16 October 2011; accepted 7 November 2011

Background: Despite improvement in therapeutic techniques, patients with early-stage laryngeal cancer still recur after treatment. Gene expression prognostic models could suggest which of these patients would be more appropriate for testing adjuvant strategies.
Materials and methods: Expression profiling using whole-genome DASL arrays was carried out on 56 formalin-fixed paraffin-embedded tumor samples of patients with early-stage laryngeal cancer. We split the samples into a training and a validation set. Using the supervised principal components survival analysis in the first cohort, we identified gene expression profiles that predict the risk of recurrence. These profiles were then validated in an independent cohort.
Results: Gene models comprising different number of genes identified a subgroup of patients who were at high risk of recurrence. Of these, the best prognostic model distinguished between a high- and a low-risk group (log-rank $P < 0.005$). The prognostic value of this model was reproduced in the validation cohort (median disease-free survival: 38 versus 161 months, log-rank $P = 0.018$), hazard ratio = 5.19 (95% confidence interval 1.14–23.57, $P < 0.05$).
Conclusions: We have identified gene expression prognostic models that can refine the estimation of a patient's risk of recurrence. These findings, if further validated, should aid in patient stratification for testing adjuvant treatment strategies.
Key words: early stage, expression profiling, laryngeal cancer, recurrence

introduction

Laryngeal cancer is one of the most common subtypes among head and neck malignancies [1, 2]. In about half of the cases, laryngeal cancer is located on the vocal cords. These patients frequently present in early stages with persistent hoarseness [3]. Early-stage (T1NOMO, T2NOMO) disease can be treated with radiotherapy or partial laryngectomy, endoscopic or open, according to the NCCN guidelines [4]. Five-year survival of patients with early-stage laryngeal cancer is 80%–85% after potentially curative treatment [5, 6]. Despite the improvement of the therapeutic techniques, some patients still recur after

*Correspondence to: Dr D. Spentzos, Division of Hematology/Oncology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Avenue, Boston, MA 02115, USA. Tel: +1-617-667-0586; Fax: +1-617-667-3915; E-mail: dspentzo@bidmc.harvard.edu

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology.
All rights reserved. For permissions, please email: journals.permissions@oup.com