Genomic arrays identify high-risk chronic lymphocytic leukemia with genomic complexity: a multicenter study

Alexander C. Leeksm\textsuperscript{1,2,3} Panagiotis Baliakas,\textsuperscript{4} Theodoros Myssiadios,\textsuperscript{5,6} Anna Puiggros,\textsuperscript{7,8} Karla Plevova,\textsuperscript{9,10} Anne-Marie van der Keve-Kersemaekers,\textsuperscript{11} Hidde Posthuma,\textsuperscript{11} Ana E. Rodríguez-Vicente,\textsuperscript{12} Anh Nhi Tran,\textsuperscript{6} Gisela Barbany,\textsuperscript{6} Larry Mansouri,\textsuperscript{6} Rebeqa Gunnarsson,\textsuperscript{13} Helen Parker,\textsuperscript{14} Eva van den Berg,\textsuperscript{15} Mar Bellido,\textsuperscript{16} Zadie Davis,\textsuperscript{17} Meaghan Wall,\textsuperscript{18} Ilaria Scarpelli,\textsuperscript{19} Anders Österborg,\textsuperscript{20,21} Lotta Hansson,\textsuperscript{20,21} Marie Jarosova,\textsuperscript{9,10} Paolo Ghia,\textsuperscript{22} Blanca Espinet,\textsuperscript{7,8} Sarka Pospisilova,\textsuperscript{9,10} Constantine Tam,\textsuperscript{24} Loic Ysebaert,\textsuperscript{25} Florence Nguyen-Khac,\textsuperscript{26} David Ossier,\textsuperscript{17} Claudia Haferlah,\textsuperscript{27} Jacqueline Schoumans,\textsuperscript{19} Marian Stevens-Kroef,\textsuperscript{28} Eric Eldering,\textsuperscript{2,3} Kostas Stamatopoulos,\textsuperscript{8,6} Richard Rosenquist,\textsuperscript{6} Jonathan C. Strefford,\textsuperscript{14} Clemens Mellink\textsuperscript{15} and Arnon P. Kater\textsuperscript{1,3} on behalf of ERIC, the European Research Initiative on CLL

\textsuperscript{1}Department of Hematology and Lymphoma and Myeloma Center Amsterdam (LYM M CARE), Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands; \textsuperscript{2}Department of Experimental Immunology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands; \textsuperscript{3}Cancer Center Amsterdam and Amsterdam Infection and Immunity Institute, Amsterdam, the Netherlands; \textsuperscript{4}Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden; \textsuperscript{5}Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece; \textsuperscript{6}Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; \textsuperscript{7}Laboratori de Citogenètica Molecular, Servei de Patologia, Hospital del Mar, Barcelona, Spain; \textsuperscript{8}Grup de Recerca Translacional en Neoplàsies Hematològiques, Programa de Recerca en Cancer, Institut Hospital del Mar d’Investigacions Mèdiques (IMIM), Barcelona, Spain; \textsuperscript{9}Department of Internal Medicine Hematology and Oncology, University Hospital Brno and Faculty of Medicine, Masaryk University, Brno, Czech Republic; \textsuperscript{10}Central European Institute of Technology (CEITEC), Masaryk University, Brno, Czech Republic; \textsuperscript{11}Department of Clinical Genetics, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands; \textsuperscript{12}Department of Hematology, IBSAL, IBMCC, University of Salamanca, CSIC, Cancer Research Center University Hospital of Salamanca, Salamanca, Spain; \textsuperscript{13}Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Lund, Sweden; \textsuperscript{14}Cancer Genomics, Academic Unit of Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; \textsuperscript{15}Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; \textsuperscript{16}Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; \textsuperscript{17}Department of Molecular Pathology, Royal Bournemouth Hospital, Bournemouth, UK; \textsuperscript{18}Cytogenetics Department, St Vincent Hospital, Melbourne, Victoria, Australia; \textsuperscript{19}Oncogenomic Laboratory, Department of Hematology, Lausanne University Hospital (CHUV), Lausanne, Switzerland; \textsuperscript{20}Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden; \textsuperscript{21}Department of Hematology, Karolinska University Hospital, Stockholm, Sweden; \textsuperscript{22}Division of Experimental Oncology, IRCCS Ospedale San Raffaele e Università Vita-Salute San Raffaele, Milan, Italy; \textsuperscript{23}Department of Clinical Genetics, Amsterdam University Medical Centers, Vrije Universiteit, Amsterdam, the Netherlands; \textsuperscript{24}Department of Haematology, St Vincent Hospital Melbourne and Peter MacCallum Cancer Center; University of Melbourne, Melbourne, Victoria, Australia; \textsuperscript{25}Institut Universitaire du Cancer de Toulouse-Oncopôle, Toulouse, France; \textsuperscript{26}INSERM U1138; Université Pierre et Marie Curie-Paris; Service d’Hématologie Biologique, Hôpital Pitié-Salpêtrière, APHP, Paris, France; \textsuperscript{27}MLL Munich Leukemia Laboratory, Munich, Germany and \textsuperscript{28}Radboud University Medical Center, Department of Human Genetics, Nijmegen, the Netherlands

©2021 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2019.239947

Received: October 9, 2019.
Accepted: January 22, 2020.
Pre-published: January 23, 2020.
Correspondence: ARNON P. KATER - a.p.kater@amc.nl
Supplementary Methods

TP53 mutation analysis and IGHV determination
In brief, exons 4-8 (in some centers also exons 1-3 and 9-10) of the TP53 gene were sequenced in 1266/2293 patients. Sanger sequencing was used in most cases (>80%; Supplemental Table 2) while the remaining cases were evaluated by targeted next generation sequencing with a variant allele frequency cutoff of 10%. Patients carrying IGHV genes with <98% germline identity were classified as IG-mutated CLL (M-CLL); those with ≥98% as IG-unmutated CLL (U-CLL).

Genomic array analysis
DNA was extracted from whole blood samples or CD19-purified cells. DNA integrity and purity were routinely verified by gel electrophoresis and a260/a280 ratios, respectively. Features of array platforms included in this study are summarized in Haraksingh et al.(1) and array processing was performed according to the manufacturer’s protocols. In general, the specific resolution of a particular platform is defined by the number and genomic distribution of the arrayed elements. Minimal resolution and sensitivity for platforms in this study are depicted in Supplemental Table 3. CNAs positioned in/overlapping with regions containing known germline copy-number variations (CNV, Database of Genomic Variants (DGV), http://projects.tcag.ca/variation) were discarded.(2) Any CNA greater than 5 Mb was included regardless of annotation in the DGV. CNAs were annotated against NCBI build GRCh37/hg19. Each genomic profile provided by the contributing centers was collated centrally and CNAs were classified as chromosomal aberrations related to a specific chromosome (loss or gain of the entire chromosome) or chromosome arm (e.g. loss 1p, gain 1p, loss 1q, gain 1q, etc). Putative chromothripsis was defined as ≥10 oscillating copy numbers involving 2 or 3 copy number states on one chromosome.(3)

ROC analysis
Receiver Operating Characteristic (ROC) curve analysis was used to assess the diagnostic accuracy of the total number of CNAs, measured at baseline(4) (date of array analysis), on overall survival. In order to detect the most appropriate threshold(s) reflecting genomic complexity, and to accommodate the time effect, time-dependent ROC analysis was applied by evaluating different time points from date of array analysis. In particular the years 5, 10, and 15 were considered and the most appropriate threshold for genomic complexity was detected in each case. The threshold/cutoff selection was based both on the (a) minimum distance criterion, and (b) the Youden index.(5, 6) The analysis was performed in R based on the
package “tdROC”, which calculates the time-dependent sensitivity, specificity and area under the curve using a nonparametric weighting adjustment. (7)

Maximally selected rank statistic
An alternative approach was applied in order to assess the diagnostic power of the total number of CNAs on overall survival, based on maximizing selected rank statistics (8). The most appropriate threshold was determined, resulting in two distinct groups. The maximally selected rank statistic approach was applied based on the R package “maxstat”.

Bootstrap
A bootstrapping procedure was applied to validate the stability of the detected thresholds. Particularly, 100 bootstrap samples, which were equal in size to the originally selected population, were randomly generated with replacement from the originally selected CLL population. Subsequently, for each bootstrap sample, the same procedure was applied, including the application of the time-dependent ROC analysis and the maximally selected rank statistic approach. The derived thresholds in each case were recorded resulting in the threshold distribution, which enabled us to evaluate the thresholds detected in the originally selected CLL population. The percentages observed for the original thresholds exhibited an average of 79% signifying their prevalence and validating the original selection.

Concordance index
The Harrell’s concordance index (9, 10) was calculated for each multivariable Cox model to assess the discriminatory ability of the Cox model. (11)

References
1. Haraksingh RR, Abyzov A, Urban AE. Comprehensive performance comparison of high-resolution array platforms for genome-wide Copy Number Variation (CNV) analysis in humans. BMC Genomics. 2017 Apr 24;18(1):321.
2. Simons A, Sikkema-Raddatz B, de Leeuw N, Konrad NC, Hastings RJ, Schoumans J. Genome-wide arrays in routine diagnostics of hematological malignancies. Hum Mutat. 2012 Jun;33(6):941-8.
3. Rausch T, Jones DT, Zapatka M, Stutz AM, Zichner T, Weischenfeldt J, et al. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell. 2012 Jan 20;148(1-2):59-71.
4. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000 Jun;56(2):337-44.
5. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J. 2005 Aug;47(4):458-72.
6. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. Biom J. 2008 Jun;50(3):419-30.
7. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform. 2005 Oct;38(5):404-15.
8. Hothorn T, Zeileis A. Generalized maximally selected statistics. Biometrics. 2008 Dec;64(4):1263-9.
9. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982 May 14;247(18):2543-6.
10. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984 Apr-Jun;3(2):143-52.
11. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011 May 10;30(10):1105-17.
Table S1. Demographics and biological features of the patients included in the study

|                | Whole cohort (N,%) | Untreated‡ (N,%) | Treated§ (N,%) |
|----------------|-------------------|-----------------|---------------|
| **Male**       | 1419, 67.9%       | 620, 66.2%      | 225, 69.0%    |
| **Female**     | 672, 32.1%        | 316, 33.8%      | 101, 31.0%    |
| **Median age diagnosis** | 62.5 years | 63.0 years | 60.5 years |
| **<55**        | 363, 23.9%        | 222, 22.7%      | 102, 28.3%    |
| **>70**        | 360, 23.7%        | 236, 24.1%      | 75, 20.8%     |
| **Binet A**    | 794, 58.3%        | 597, 64.2%      | 146, 45.8%    |
| **Binet B**    | 387, 28.4%        | 239, 25.7%      | 106, 33.2%    |
| **Binet C**    | 181, 13.3%        | 94, 10.1%       | 67, 21.0%     |
| **M-CLL**      | 509, 50%          | 345, 54.6%      | 56, 36.4%     |
| **TP53 abn†**  | 238, 17.7%        | 82, 10.8%       | 66, 28%       |
| **del(11)(q22.3)** | 395, 17.2%  | 164, 16.8%      | 77, 21.3%     |
| **trisomy 12** | 293, 12.8%        | 118, 12.1%      | 49, 13.6%     |
| **del(13)(q14)** | 1184, 51.6% | 528, 53.9%      | 195, 54.0%    |
| **Median follow up** | 33 months | 44 months | 15 months |

Abbreviations: \(^*\)M-CLL= CLL with mutated IGHV, \(^†\)TP53abn= del(17)(p13.1) and/or TP53 mutation, \(^‡\)“untreated”=untreated at date of sampling, \(^§\)“treated”=treated at date of sampling. Percentages were calculated with respect to the number of patients with available data for the respective parameter in each of the 3 groups and not with respect to the total number of patients in the respective groups (whole cohort, untreated or treated).
### Table S2. Sequencing methods used by participating centers

| Center | Total nr | TP53 sequencing method |
|--------|----------|-------------------------|
| Groningen UMC | NA | NA |
| Hospital del Mar Barcelona | NA | NA |
| University Hospital Brno | 67 | Sanger |
| Uppsala | 364 | Sanger |
| Karolinska Institute | 216 | Sanger |
| Southampton (Royal Bournemouth Hospital) | 152 | Sanger |
| Radboud UMC | 206 | Sanger/NGS |
| CHUV | NA | NA |
| Amsterdam UMC | 56 | Sanger |
| Pitie-Salpetriere | 150 | Sanger |
| MLL | 38 | Sanger/NGS |
| IUCT-Oncopole | 14 | NGS |
| SVHM | 3 | NGS |

### Table S3. Platform characteristics for genomic arrays used in this study

| Center | Agilent oligonucleotides 180K | SurePrint G3 ISCA | Affymetrix 250K SNP-array | Affymetrix SNP6.0 array | Whole-Genome 2.7M | sensitivity | resolution |
|--------|-------------------------------|-------------------|-------------------------|-----------------------|-------------------|------------|-----------|
| Groningen UMC | 138 | 10-20% | 1 kb |
| Hospital del Mar Barcelona | 74 | 10-25% | 1 kb |
| University Hospital Brno | 46 | 21 | 10-20% | 1 kb |
| Uppsala | 368 | 10-15% | 1-10 kb |
| Karolinska Institute | 216 | 10-15% | 13 kb |
| Southampton (Bournemouth Hospital) | 190 | 10-15% | 1 kb |
| Radboud UMC | 221 | 10-20% | 1 kb |
| CHUV | 480 | 10-15% | 1 kb |
| Amsterdam UMC (AMC) | 179 | 10-15% | 13 kb |
| Pitie-Salpetriere | 161 | 10-15% | 13 kb |
| Amsterdam UMC (VUMC) | 124 | 10-20% | 1 kb |
| MLL | 41 | 10-15% | 25 kb |
| IUCT-Oncopole | 22 | 10-20% | 1 kb |
| SVHM | 12 | 10-20% | 1 kb |
## Table S4. Copy number alterations associated with del(11)(q22.3) (ATM), del(13)(q14), trisomy 12 and del(17)(p13.1) (TP53) and different CLL subgroups

| Subgroup | Del | OR | Relative risk | OR (95 % CI) | OR ratio | P value |
|----------|-----|----|--------------|--------------|----------|--------|
| Normal   |     |    |              |              |          |        |
| Binet A  |     |    |              |              |          |        |
| Binet B  |     |    |              |              |          |        |
| Binet C  |     |    |              |              |          |        |

### Subgroups

- **Binet A**
- **Binet B**
- **Binet C**
- **Normal**

### OR and Relative Risk

- **OR**
- **Relative Risk**
- **OR (95 % CI)**
- **OR ratio**
- **P value**

### Notes

- OR refers to odds ratio, and the 95 % CI represents the 95 % confidence interval.
### Table S5. Overview of detected CNAs captured by FISH vs. genomic arrays for patients with simultaneous FISH and genomic array data

|                 | del(11)(q22.3) (ATM) | trisomy 12 | del(13)(q14) | del(17)(p13.1) (TP53) |
|-----------------|----------------------|------------|--------------|-----------------------|
| **array**       |                      |            |              |                       |
| total           | 249                  | 1          | 249          | 1                     |
| No              | 186                  | 0.747      | 223          | 0.896                 |
| Yes             | 63                   | 0.253      | 26           | 0.104                 |
| **FISH**        |                      |            |              |                       |
| total           | 249                  | 1          | 249          | 1                     |
| No              | 171                  | 0.687      | 217          | 0.871                 |
| Yes             | 78                   | 0.313      | 32           | 0.129                 |

### Table S6. Minimal common regions of deletion or amplification

| Chromosome arm [GRCh37] chromosomal regions                               |
|----------------------------------------------------------------------|
| del1q 1q21.1q21.2(144894611_149768855)                                 |
| del1q 1q23.3q23.3(160751105_161479451)                                 |
| dup2p 2p25.3p25.1(3721713_9073918)                                    |
| dup2p 2p16.1p15(60932040_62206329)                                    |
| dup2p 2p23.3p22.3(25342914_32841818)                                  |
| dup3q 3q26.3q27.2(174773031-184972301)                                |
| del4p 4p15.2p15.1(27647757_28761977)                                 |
| del6q 6q25.2q25.3(153946329_157482664)                                |
| del6q 6q21q21(107327737_110881818)                                   |
| del8p 8p21.3p21.2(19101696_23304899)                                 |
| dup8q 8q24.21q24.21(128286744_130836899)                              |
| del9p 9p24.3-p24.1(1404921-5932368)                                  |
| del9p 9p21.3p21.3(22899648_23041037)                                 |
| del9p 9p13.1p11.1(38916514_46746820)                                 |
| del13q.other* 13q33.2q33.3(105570440_108304501)                      |
| del13q.other 13q21.2q21.33(62290433_70260961)                        |
| del13q.other 13q12.11q12.12(21375669_25254198)                       |
| dup13q 13q31.3q32.2(92210001_98472541)                               |
| del14q 1q21.1(39583972_44352416)                                    |
| del14q 1q24.1q24.2(69704553_70051926)                                |
| del14q 1q32.13q32.3(95998766_104101254)                              |
| del15q 15q26.1q26.3(94308921_99056760)                               |
| del15q 15q15.1q15.1(40721923_40845473)                               |
| del15q 15q25.2q25.3(83734673_84867550)                               |
| del15q 15q21.3q21.3(54289217_54570517)                               |
| dup17q 17q22q24.3(53736288_67341400)                                 |
| del18p 18p11.2p11.31(2641858_5824910)                                |
| del20p 20p12.3p12.3(6927825_7704212)                                 |

*del13q.other are deletions on 1q not containing the 13q14 region recurrently deleted in CLL*
### Table S7. Univariable Cox regression analysis for time to first treatment (TTFT)

| Predictors                        | N=963 | HR   | 95% HR CI       | P-values |
|-----------------------------------|-------|------|-----------------|----------|
| Male                              | 920   | 1.34 | 1.11-1.62       | 0.003    |
| >70 years                         | 963   | 0.77 | 0.62-0.96       | 0.017    |
| Binet B/C                         | 915   | 4.74 | 3.95-5.69       | <0.001   |
| U-CLL∗                            | 628   | 4.68 | 3.79-5.80       | <0.001   |
| TP53abn†                          | 749   | 1.57 | 1.18-2.08       | 0.002    |
| del(11)(q22.3)                    | 963   | 2.13 | 1.73-2.61       | <0.001   |
| GC‡ (3 categories)                |       |      |                 |          |
| intermediate-GC|| vs. low-GC§      | 963   | 1.67 | 1.32-2.12       | <0.001   |
| high-GC¶ vs. low-GC               | 963   | 2.81 | 2.04-3.86       | <0.001   |
| GC≥5                              | 963   | 2.59 | 1.89-3.55       | <0.001   |

Abbreviations: ‘U-CLL= CLL with unmutated IGHV, †TP53abn= del(17)(p13.1) (TP53) and/or TP53 mutation, ‡GC=genomic complexity, GC categories: §low-GC=[0-2], ||Intermediate-GC=[3-4], ¶High-GC=[≥5] CNAs detected by array

### Table S8. Univariable Cox regression analysis for overall survival (OS)

| Predictors                        | N=961 | HR   | 95% HR CI       | P-values |
|-----------------------------------|-------|------|-----------------|----------|
| Male                              | 918   | 1.38 | 1.10-1.75       | 0.006    |
| >70 years                         | 961   | 2.13 | 1.68-2.70       | <0.001   |
| Binet B/C                         | 913   | 2.17 | 1.74-2.69       | <0.001   |
| U-CLL∗                            | 628   | 4.04 | 3.16-5.17       | <0.001   |
| TP53abn†                          | 749   | 2.73 | 2.01-3.70       | <0.001   |
| del(11)(q22.3)                    | 961   | 2.04 | 1.60-2.61       | <0.001   |
| GC‡ (3 categories)                |       |      |                 |          |
| intermediate-GC|| vs. low-GC§      | 961   | 1.67 | 1.24-2.25       | 0.001    |
| high-GC¶ vs. low-GC               | 961   | 4.20 | 2.87-6.12       | <0.001   |
| GC≥5                              | 961   | 3.90 | 2.68-5.67       | <0.001   |

Abbreviations: ‘U-CLL= CLL with unmutated IGHV, †TP53abn= del(17)(p13.1) (TP53) and/or TP53 mutation, ‡GC=genomic complexity, GC categories: §low-GC=[0-2], ||Intermediate-GC=[3-4], ¶High-GC=[≥5] CNAs detected by array
### Table S9. Multivariable analysis for time to first treatment (TTFT)

|              | N=528 | HR   | 95% HR CI     | P-values |
|--------------|-------|------|---------------|----------|
| Male         | 1.05  | 0.83-1.32 | 0.682        |
| >70 years    | 1.10  | 0.83-1.46 | 0.496        |
| Binet B/C    | 3.88  | 3.04-4.94 | <0.001       |
| U-CLL<sup>∗</sup> | 3.23  | 2.50-4.19 | <0.001       |
| TP53<sup>abn</sup><sup>†</sup> | 1.16  | 0.80-1.67 | 0.435        |
| del(11)(q22.3) | 1.24  | 0.95-1.61 | 0.11         |
| GC<sup>‡</sup>≥5 | 2.00  | 1.28-3.14 | 0.002        |

Abbreviations: *U-CLL= CLL with unmutated IGHV, †TP53<sup>abn</sup>= del(17)(p13.1) and/or TP53 mutation, ‡GC≥5 =genomic complexity with ≥5 CNAs detected by array

### Table S10. Multivariable analysis for overall survival (OS)

|              | N=528 | HR   | 95% HR CI     | P-values |
|--------------|-------|------|---------------|----------|
| Male         | 1.24  | 0.95-1.63 | 0.112        |
| >70 years    | 2.49  | 1.87-3.33 | <0.001       |
| Binet B/C    | 1.49  | 1.15-1.94 | 0.003        |
| U-CLL<sup>∗</sup> | 3.85  | 2.86-5.18 | <0.001       |
| TP53<sup>abn</sup><sup>†</sup> | 1.72  | 1.18-2.51 | 0.005        |
| del(11)(q22.3) | 0.98  | 0.72-1.32 | 0.87         |
| GC<sup>‡</sup>≥5 | 2.18  | 1.35-3.54 | 0.002        |

Abbreviations: *U-CLL= CLL with unmutated IGHV, †TP53<sup>abn</sup>= del(17)(p13.1) and/or TP53 mutation, ‡GC≥5 =genomic complexity with ≥5 CNAs detected by array
Figure S1. Diagram of the patients included in this study. For survival analysis only patients untreated at date of sampling were included to exclude the effects of prior treatment on survival.

CLL diagnostic centers involved in this multicenter study (n=13):
Amsterdam UMC (n=303) (AMC and VUMC), Radboud UMC (n=221), Groningen UMC (n=138) (the Netherlands), Royal Bournemouth Hospital (United Kingdom) (n=190), Pitie-Salpetriere (n=161), IUCT-OncoPole (n=22) (France), Hospital del Mar Barcelona (Spain) (n=74), CHUV (Switzerland) (n=480) SVHM (Australia) (n=12), MLL (Germany) (n=41), University Hospital Brno (Czech Republic) (n=67), Uppsala (n=368), Karolinska Institute (n=216) (Sweden)
Figure S2. Overview of CNAs in different CLL subgroups. A-D) Pie charts representing the percentage of patients with a given number of CNAs detected by genomic array. Untreated and previously treated cases (A), different Binet subgroups (B), IGHV gene status (C) and TP53 status (D) are shown.
Figure S3. Kaplan-Meier plots representing the effect of putative chromothripsis events on overall survival in all evaluable (A) and in TP53abn/del(11q22.3) (ATM) cases (B). Analyses performed on all patients of which survival data were available (irrespective of treatment information; n=1432 or n=406, respectively).
Figure S4. Overview of CNAs in this study observed in at least 10 patients

| CNAs                  | % (# patients) |
|-----------------------|---------------|
| loss 13q14            | 51.6% (1184)  |
| loss 11q22.3          | 17.2% (395)   |
| tris 12               | 12.8% (293)   |
| loss 17p13.1          | 7.3% (168)    |
| gain 2p*              | 6.4% (147)    |
| gain 8q*              | 4.8% (110)    |
| loss 14q*             | 4.5% (103)    |
| loss 6q*              | 4.1% (93)     |
| loss 8p*              | 3.8% (86)     |
| loss 18p*             | 3.1% (72)     |
| loss 4p*              | 2.3% (53)     |
| loss 1q*              | 2.1% (48)     |
| loss 15q*             | 1.7% (40)     |
| gain 3q*              | 1.7% (39)     |
| loss 9p               | 1.7% (39)     |
| loss Y                | 1.7% (38)     |
| loss 13q other        | 1.3% (30)     |
| loss 4q               | 1.3% (30)     |
| loss 3p               | 1.3% (30)     |
| gain 17q              | 1.1% (26)     |
| loss 6p               | 1.1% (26)     |
| loss 7q               | 1.1% (25)     |
| loss 2q               | 1.1% (25)     |
| loss 20p              | 1.0% (24)     |
| tris 19               | 1.0% (23)     |
| gain 13q              | 0.9% (21)     |
| loss 10q              | 0.9% (21)     |
| tris 18               | 0.9% (20)     |
| loss 1p               | 0.8% (19)     |
| loss 18q              | 0.8% (18)     |
| gain 18q              | 0.8% (18)     |
| loss 9q               | 0.8% (18)     |
| loss 5q               | 0.8% (18)     |
| gain 5q               | 0.8% (18)     |
| gain 21q              | 0.7% (16)     |
| loss 11q other        | 0.7% (16)     |
| loss 8q               | 0.7% (15)     |
| loss 3q               | 0.7% (15)     |
| loss X                | 0.6% (14)     |
| gain 22q              | 0.6% (13)     |
| gain 15q              | 0.5% (12)     |
| gain 12p              | 0.5% (12)     |
| gain 11p              | 0.5% (12)     |
| loss 2p               | 0.5% (12)     |
| loss 21q              | 0.5% (11)     |
| loss 19p              | 0.5% (11)     |
| loss 12p              | 0.4% (10)     |
| loss 7p               | 0.4% (10)     |
| gain 2q               | 0.4% (10)     |
Figure S5. Correlation of CNAs detected by genomic arrays with IGHV gene status. Circos plot comparing the correlation of the 10 most frequently observed CNAs other than del(11)(q22.3) (ATM), trisomy 12, del(13)(q14) and del(17)(p13.1) (TP53) normally detected by FISH in this study, with IGHV gene status. Significant correlations with a corrected $p<0.01$ are indicated with an asterisk (*).
Figure S6. Correlation of CNAs detected by genomic array with CNAs normally analyzed by FISH. Circos plots comparing the correlation of del(11q) (A), trisomy 12 (B), del(13q) (C) and del(17p) (D) status with 10 CNAs not captured by CLL FISH probes. Significant correlations with a corrected $p<0.01$ are indicated with an asterisk (*).
Figure S7. Kaplan-Meier plots representing the effect of different CNAs on overall survival
Figure S8. Kaplan-Meier plots representing the effect of GC subgroups on time to first treatment (A) and overall survival (B) in unmutated IGHV gene (U-CLL) cases.

A Time to first treatment

![Kaplan-Meier plot for time to first treatment](image)

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 194/211 | 0.11 (0.05-0.2) |
| intermediate-GC [3-4 CNAs] | 45/50 | 0.20 (0.03-0.4) |
| high-GC [≥5 CNAs] | 14/25 | 0.01 (0.0-0.02) |

'B' vs. 'intermediate' p=0.925
'low' vs. 'high' p<0.001
'intermediate' vs. 'high' p<0.001

B Overall Survival

![Kaplan-Meier plot for overall survival](image)

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 164/212 | 6.4 (5.8-7.0) |
| intermediate-GC [3-4 CNAs] | 34/50 | 4.6 (1.0-8.2) |
| high-GC [≥5 CNAs] | 18/25 | 4.1 (1.9-6.4) |

'low' vs. 'intermediate' p=0.39
'low' vs. 'high' p<0.001
'intermediate' vs. 'high' p=0.047
Figure S9. Kaplan-Meier plots representing the effect of GC subgroups on time to first treatment (A) and overall survival (B) in mutated IGHV gene (M-CLL) cases.

A Time to first treatment

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 129/303 | 13.2 (10.5-15.9) |
| intermediate-GC [3-4 CNAs] | 22/31 | 2.9 (1.7-4.2) |
| high-GC [≥5 CNAs] | 8/8 | 0.02 (0.0-0.2) |

B Overall Survival

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 91/303 | 14.6 (13.2-16.1) |
| intermediate-GC [3-4 CNAs] | 13/31 | 12.5 (7.9-17.2) |
| high-GC [≥5 CNAs] | 6/8 | 1.4 (0.0-3.7) |
Figure S10. Kaplan-Meier plots representing the effect of GC subgroups on time to first treatment (A) and overall survival (B) in TP53abn/del(11q) positive cases.

A Time to first treatment

TP53abn/del(11q)

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 75/110 | 0.23 (0.0-0.7) |
| intermediate-GC [3-4 CNAs] | 55/72 | 0.36 (0.05-0.7) |
| high-GC [≥5 CNAs] | 36/45 | 0.06 (0.0-0.4) |

B Overall survival

TP53abn/del(11q)

| GC Subgroup | No. of events | Median (95% CI) |
|-------------|---------------|-----------------|
| low-GC [0-2 CNAs] | 60/111 | 6.3 (4.5-8.1) |
| intermediate-GC [3-4 CNAs] | 38/74 | 4.7 (2.1-7.3) |
| high-GC [≥5 CNAs] | 27/46 | 3.1 (1.0-5.1) |
Figure S11. Distribution of chromosomal abnormalities detected by CBA (A) and genomic arrays (B) in patients with simultaneous CBA and genomic array analyses available.

A CBA

B Genomic arrays
**Supplemental excel file.** A list of curated array profiles is provided online in the Supplemental excel file, separately uploaded.