Oncol Ther (2021) 9:621–634
https://doi.org/10.1007/s40487-021-00172-2

ORIGINAL RESEARCH

Differential Levels of mRNAs in Normal B Lymphocytes, Monoclonal B Lymphocytosis and Chronic Lymphocytic Leukemia Cells from the Same Family Identify Susceptibility Genes

Abdullah Alshahrani · Kristen K. Skarratt · Kristy P. Robledo · Maryam Hassanvand · Benjamin Tang · Stephen J. Fuller

Received: July 27, 2021 / Accepted: September 23, 2021 / Published online: October 7, 2021
© The Author(s) 2021

ABSTRACT

Introduction: People with a family history of chronic lymphocytic leukemia (F-CLL) have an increased risk of monoclonal B lymphocytosis (F-MBL), which is found in up to 18% of first-degree relatives of patients compared to 5% of the total population. This may indicate that the presence of an F-MBL in the relative of a F-CLL patient is due to genetic susceptibility. In this study, we hypothesized that progressive changes in gene expression result in malignant transformation of B lymphocytes to F-MBL, and subsequent alterations in gene expression occur before overt F-CLL develops. The aim of this study of affected and unaffected individuals from a family with multiple CLL cases was to compare mRNA expression levels in control B-lymphocytes, pre-malignant F-MBL and malignant F-CLL cells.

Methods: To identify inherited changes in gene expression, a high-resolution DNA microarray was used to identify differentially abundant mRNAs in age-matched cases of F-MBL (n = 4), F-CLL (n = 2) and unaffected family relatives (F-Controls, n = 3) within one family. These were then compared to non-kindred controls (NK-Controls, n = 3) and sporadic CLL (S-CLL) cases (n = 6).

Results: Seven differentially abundant mRNAs were identified against similar genetic backgrounds of the family: *GRASP* and *AC016745.3* were decreased in F-MBL and further decreased in F-CLL compared to F-Controls, whereas *C11orf80* and *METTL8* were progressively increased. *PARP3* was increased in F-MBL compared to F-Controls but was decreased in F-CLL compared to F-MBL. Compared to F-Controls, levels of *ROR1* and *LEF1* were similarly increased in F-MBL and F-CLL. For six of the genes, there were no differences in mRNA levels between S-CLL and F-CLL; however *PARP3* was higher in S-CLL.

Conclusion: These results are consistent with the hypothesis that changes in expression of specific genes contribute to transformation from normal lymphocytes to MBL and CLL.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40487-021-00172-2.

A. Alshahrani · K. K. Skarratt · M. Hassanvand · B. Tang · S. J. Fuller (✉)
Department of Medicine, Sydney Medical School Nepean, Nepean Hospital, University of Sydney, Penrith, NSW 2750, Australia
e-mail: stephen.fuller@sydney.edu.au

A. Alshahrani
College of Applied Medical Sciences, King Khalid University, Guraiger, Abha 62529, Kingdom of Saudi Arabia

K. P. Robledo
NHMRC Clinical Trials Centre, University of Sydney, Camperdown, NSW 2006, Australia

Keywords: Chronic lymphocytic leukemia; Familial; Gene expression; Monoclonal B lymphocytosis

Key Summary Points
Why carry out this study?
CLL is the most common form of leukemia in Western countries and remains incurable.
There have been major advances in development of ‘small molecule’ targeted drugs; however, treatment failures and resistance to new targeted therapies are common.
Our hypothesis is that expression of potential target genes changes with progression of normal B-lymphocytes through pre-malignant MBL cells to malignant CLL.
What was learned from this study?
Levels of GRASP and AC016745.3 mRNAs were progressively lower and C11orf80, ROR1, METTL8, and LEF1 mRNAs progressively higher in B lymphocytes from F-MBL and F-CLL cases compared to F-Controls. PARP3 was increased in F-MBL compared to F-Controls but decreased in F-CLL compared to F-MBL.
The findings for F-CLL were the same in S-CLL, except for PARP3, which was higher in S-CLL.Multiple CLL case families, though limited by small numbers of patients, can be studied to identify differentially abundant mRNAs in normal B lymphocytes, MBL and CLL cells and provide new molecular signatures for targeted therapies.

INTRODUCTION

Chronic lymphocytic leukemia (CLL) accounts for > 25% of all leukemia cases in Western countries [1], and a family history is found in up to 10% of cases [2]. Familial clustering of CLL (F-CLL) has been consistently demonstrated in epidemiological studies [3], and a higher frequency of monoclonal B-cell lymphocytosis (MBL), a precursor to CLL, is found in CLL families [4, 5].

To detect patterns of multiple low-risk loci, genome-wide association studies (GWAS) have analyzed large numbers of F-CLL and sporadic CLL (S-CLL) cases and controls using dense-coverage single nucleotide polymorphism (SNP) arrays [6–8]. Over 40 risk mutations have been identified to have a role in the etiology of CLL [9], and 30 of these account for ~ 19% of CLL heritability [8], suggesting that a significant proportion of genetic susceptibility has not been detected. Some of this “missing heritability” may be associated with non-DNA sequence-based inheritance factors that affect gene expression, including epigenetic variations, which have been found in several familial cancers [10, 11]. The simultaneous presence of F-CLL and familial MBL (F-MBL) in families provides an opportunity to study changes in mRNA levels associated with progression to CLL against similar genetic backgrounds.

DNA microarray studies have identified differential mRNA expression among normal B lymphocytes, MBL and CLL cells from unrelated individuals [12]. However, this is the first study of gene expression from normal B lymphocytes, F-MBL and F-CLL from within one family. We previously performed a genome-wide linkage scan of the family using high-density SNP markers; however, there was no significant evidence for a single gene model of disease susceptibility, suggesting that susceptibility to CLL has a more complex basis [13]. Although individual family studies are limited by low subject numbers, background genetic variation is reduced, increasing the detection of epigenetic and environmental modifiers associated with variation in gene expression and phenotype [14].

To identify differential mRNAs associated with B lymphocytes, F-MBL and F-CLL, blood samples were collected from members of one of the largest multiple-case CLL kindreds reported in the literature [13]. In the present study, DNA microarrays were used to compare mRNAs in enriched B lymphocytes to determine whether mRNA abundances of genes differed among B lymphocytes from control subjects, F-MBL, F-CLL and S-CLL cases.

METHODS

Patients and Samples

The experimental protocol was approved by the Nepean and Blue Mountains Local Health District Human Research Ethics Committee (01/70). Peripheral blood samples (40 ml) were collected from six patients (2 with F-CLL and 4 with F-MBL) and three unaffected members from a family with multiple cases of CLL (Fig. 1) [13]. In addition, samples were collected from six unrelated S-CLL cases and three NK-Controls. All CLL subjects were treatment naïve. The diagnosis of CLL was based on the presence of a clonal B lymphocyte count $\geq 5 \times 10^9$/l for ≥ 3 months, co-expression of CD19, CD5 and CD23, and weak or no expression of CD20, CD79b, CD22 and surface IgM [15]. The diagnosis of F-MBL was based on the same immunophenotype but clonal B cells were $< 5 \times 10^9$/l. For comparisons of mRNA levels in F-MBL, F-CLL and S-CLL, we sought to reduce the effect of genetic relatedness by combining NK-Controls with F-Controls (Combined Controls).

B lymphocytes were enriched using a RosetteSep™ B Cell isolation cocktail (StemCell Technologies Inc., Vancouver, BC, Canada) to provide > 95% B lymphocyte purity confirmed by flow cytometry [16].

IgVH Usage and Mutation Analysis

Genomic DNA was extracted using the Wizard® Genomic DNA Purification kit (Promega, Madison, WI, USA) and quantified using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Amplification by polymerase chain reaction (PCR) and sequence analysis of *IgVH* rearrangements were conducted according to BIOMED-2 protocols [17, 18], using *IgVH* gene clonality master mixes

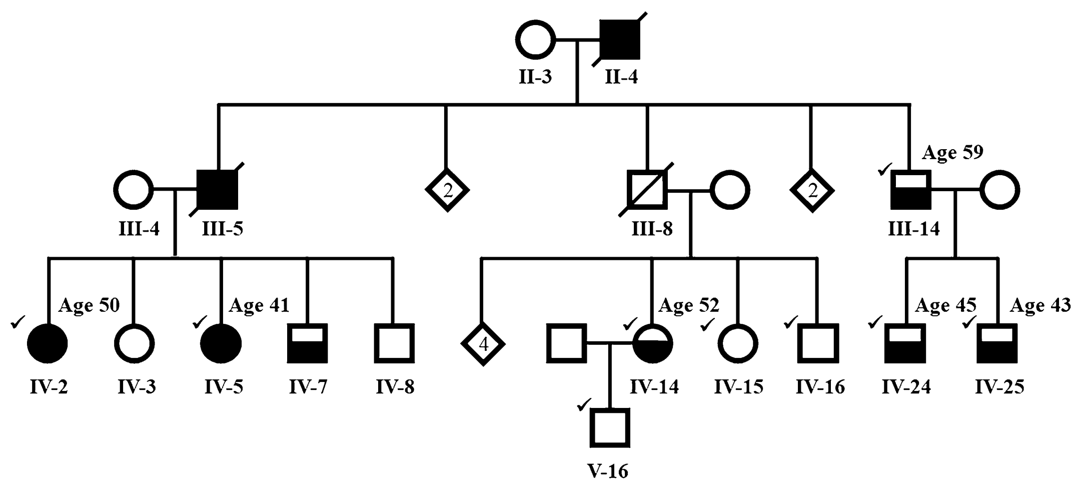

Fig. 1 Pedigree of the family. The pedigree in abbreviated form showing segregation of CLL. Blackened symbols denote family members affected with F-CLL; half-shaded symbols F-MBL; clear symbols unaffected; ticked symbols individuals studied from whom mRNA were collected; diamonds represent grouped siblings. The pedigree numbering system corresponds to the original report of this family [13], where each generation is identified by a Roman numeral and each child and cousin in the same generation is identified by an Arabic numeral. Ages (in years) at diagnosis of F-MBL or F-CLL are shown

(InVivoScribe Technologies, San Diego, CA, USA). Purified PCR products were sequenced at the Australian Genome Research Facility, Brisbane, Australia. Ig blast GenBank and the IMGT/V-QUEST portal for immunoglobulin and T cell receptor sequences (International ImMunoGeneTics Information System) were used to analyze and align *IgVH* sequences [19]. Sequences with germline homology ≥ 98% were considered as unmutated and those < 98% as mutated [19].

Interphase Fluorescence In Situ Hybridization (FISH)

FISH analyses for common abnormalities associated with CLL were performed in affected individuals using the following probes: DLEU/LAMP at 13q14, chromosome 12 centromere, ATM at 11q22 and TP53 at 17p13. Interphase FISH studies were performed based on techniques adapted from the Cytogenetics and the Molecular Genetics Laboratory, the Children's Hospital at Westmead, Sydney, NSW, Australia. Two hundred images of interphase nuclei were captured for every probe set according to the manufacturer's instructions. Results were abnormal when the percentage of cells with any given abnormality exceeded 5% in 200 interphase nuclei for trisomy 12 and 8% for deletions of 13q, 11q and 17p.

RNA Extraction

RNA was extracted using the Isolate II RNA mini kit (Bioline, Taunton, MA, USA). Samples were quantified and purity determined using a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). RNA purity was assessed by measuring absorbances at 260 and 280 nm (A260 and A280, respectively). Samples with concentrations between 50 and 100 ng/μl and with A260/A280 > 1.8 were analyzed using Affymetrix gene expression microarrays (Affymetrix Inc, Santa Clara, CA, USA). An additional RNA quality assessment was performed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to determine the ratio of two ribosomal RNAs (rRNA; 28S/18S) and the RNA integrity number (RIN). Only RNA preparations with a 28S/18S rRNA ratio > 2 and RIN > 7 were used for microarray analyses.

Transcriptome Profiling

RNA was prepared as described for the GeneChip® WT Pico Reagent Kit (Affymetrix Inc, Santa Clara, CA, USA) and analyzed using Affymetrix GeneChip® Human Transcriptome 2.0 Arrays. Affymetrix transcriptome analysis console (TAC 3.0) software was used to perform statistical analyses. Gene expression intensity was calculated for each sample using Tukey's Bi weight average for all eligible exons' intensities in that gene and represented as a bi-weight average shown in a log2 scale. The quality of each Affymetrix Human Transcriptome Array was determined using Affymetrix spike-in controls, perfect match expression and relative log expression during data summarization and normalization in the Affymetrix expression console software, version 1.4.1. The Affymetrix transcriptome analysis console (TAC 3.0) software was used to perform statistical analysis and generate a list of differentially expressed mRNAs. The following formula was used to compare fold change in expression between CLL and controls: $\log_2(\text{CLL/control}) = \log_2(\text{CLL}) - \log_2(\text{control})$ and converted to a linear scale fold-change value using the formula $[2^{\log2(\text{CLL/control})}]$. Quantitative reverse transcription PCR (qRT-PCR) was used to confirm *GRASP* mRNA levels. mRNA was converted to cDNA using a Tetro cDNA synthesis kit (Bioline, Taunton, MA, USA), and qRT-PCR was performed using a Rotor-Gene 2000 cycler (Corbett Life Science; Qiagen, Hilden, Germany) with validated primer pairs (Supplementary Material, Table S1) [20]. Gene expression of *GRASP* relative to *GAPDH* was calculated using the delta cycle threshold (delta Ct) method [21].

ELECTRONIC-DATABASE INFORMATION

URLs for programs and data presented herein are as follows: US National Library of Medicine, National Center for Biotechnology Information

(accessed 31 March 2017) available from https://www.ncbi.nlm.nih.gov/igblast; the International ImMunoGeneTics Information System (accessed 31 March 2017) available from http://www.imgt.org/IMGT_vquest/vquest; The R Project for Statistical Computing (accessed 15 October 2017) available at http://www.R-project.org; National Genetics Reference Laboratory, Manchester, UK (accessed 1 December 2017) available from http://www.ngrl.org.uk/Manchester/projects/snpcheck.html.

Statistical and Bioinformatic Analyses

Identification of differentially abundant mRNAs was performed using one-way analysis of variance (ANOVA) tests, and to correct for multiple comparisons, false discovery rate (FDR) *P*-values were calculated [22–24]. Hierarchical clustering was performed using Affymetrix transcriptome analysis console version 3.0 software (Affymetrix Inc, Santa Clara, CA, USA). Distances between clusters were computed using the complete linkage method (maximum distance between a pair of objects in the two clusters), and results are displayed in a heat map and dendrogram. To determine mRNAs with FDR < 0.05 that differed among Combined controls, F-MBL, F-CLL and S-CLL, one-way ANOVA with Tukey's post hoc tests were performed using GraphPad Prism version 7.00 for Windows (GraphPad Software, La Jolla, CA, USA).

RESULTS

Clinical and Laboratory Attributes of Patients

The attributes of two F-CLL, four F-MBL and six S-CLL, including Binet stage [25], are shown in Table 1, and an abbreviated family pedigree is shown in Fig. 1. There were no significant differences in mean ages among F-Controls ($n = 3$, mean 48 years; SD 6 years), F-MBL (n = 4, mean 62 years; SD 10 years) and F-CLL ($n = 2$, mean 54 years; SD 0 years), and no difference in mean ages among F-MBL, F-CLL and S-CLL ($n = 6$, mean 73, SD 12, one-way ANOVA with Tukey's post hoc test); however, there was a significant difference in age between combined F-Controls and NK-Controls and S-CLL (mean age 49 vs 73; $P = 0.005$). To reduce the effect of genetic relatedness, NK-Controls were added to F-Controls for analyses of non-kindred S-CLL, F-MBL and F-CLL.

Comparison of mRNA Levels in F-Controls, F-MBL and F-CLL

RNA extracts of enriched B lymphocytes were prepared and analyzed to identify differences in abundance of mRNAs. Using flow cytometry, there were no differences among the purity of F-MBL CD20 +, CD5 + cases and CLL (mean purity 83% versus 94% respectively; $P > 0.05$, Student *t*-test) or between CD5 mRNA expression in F-MBL and F-CLL cases (mean log2 bi-weight avg signal 8.7 versus 9.5, respectively, $P > 0.05$, Student's *t*-test). The levels of 2095 mRNAs (1794 coding, 301 non-coding) differed among F-Controls, F-MBL and F-CLL (ANOVA $P < 0.01$) (Fig. 2 and Supplementary Material, Table S2). After correcting for multiple comparisons (FDR *P*-value < 0.05), seven mRNAs were identified that segregated F-Controls from F-MBL and F-CLL (Table 2; Fig. 3). Compared to F-Control B-lymphocytes, levels of *GRASP* mRNA and the novel transcript *ACO16745.3* were decreased in F-MBL and further decreased in F-CLL (Fig. 4a, b). *C11orf80* and *METTL8* levels were higher in F-MBL and further increased in F-CLL (Fig. 4c, d). The mean mRNA level for *PARP3* was increased in F-MBL compared to F-Controls, however less increased in F-CLL (Fig. 4e). Compared to F-Controls, *ROR1* and *LEF1* mRNA levels were increased in both F-MBL and F-CLL (Fig. 4f, g); however, there were no differences between F-MBL and F-CLL.

qRT-PCR was used to measure mRNA levels for *GRASP* in F-Controls ($n = 3$), F-MBL ($n = 4$) and F-CLL ($n = 1$), and changes in expression were determined relative to *GAPDH* using the delta Ct method. The delta Ct for *GRASP* was highest (mRNA less abundant) in F-CLL, intermediate in F-MBL and lowest in normal

Table 1 Clinical and B-cell phenotype of subjects with F-MBL, F-CLL and S-CLL

Sample identity	Age at diagnosis	Age when studied	Status	Gender	Lymphocyte count (× 10^9/l)	Binet stage[a] at time of study	Course predicted by lymphocyte doubling time[b]	IgVH mutational status	Karyotype by FISH
III-14	59	78	F-MBL	M	1	A	Indolent	Mutated	Normal
IV-14	52	61	F-MBL	F	1	A	Indolent	Mutated	Normal
IV-24	45	54	F-MBL	M	2	A	Indolent	Mutated	Normal
IV-25	43	52	F-MBL	M	2	A	Indolent	Mutated	Normal
IV-5	41	54	F-CLL	F	10	B	Indolent	Mutated	del 13q14
IV-2	50	54	F-CLL	F	17	A	Progressive	Un-Mutated	Normal
S-CLL 1	62	66	S-CLL	M	9	A	Indolent	Mutated	ND
S-CLL 2	82	93	S-CLL	F	106	C	Progressive	Un-Mutated	ND
S-CLL 3	69	80	S-CLL	F	28	A	Indolent	Mutated	ND
S-CLL 4	63	64	S-CLL	M	9	A	Indolent	Mutated	Normal
S-CLL 5	53	58	S-CLL	M	17	B	Progressive	Un-Mutated	del TP53 del 13q14
S-CLL 6	54	73	S-CLL	F	108	C	Indolent	Mutated	del TP53

ND Not done

[a] Binet stage A: no anemia (hemoglobin ≥ 100 g/l), no thrombocytopenia (platelets ≥ 100 × 10^9/l) and < 3 areas of lymphoid tissue enlargement; Binet stage B: as for stage A but ≥ 3 areas of lymphoid tissue enlargement; Binet stage C: hemoglobin ≥ 100 g/l and/or platelets ≥ 100 × 10^9/l and any number of areas of lymphoid tissue enlargement

[b] Indolent lymphocyte doubling time ≥ 12 months; progressive lymphocyte doubling < 12 months

F-Controls (Fig. 4h), consistent with the microarray data.

Comparison of mRNA Levels in Related and NK-Controls, F-MBL, F-CLL and S-CLL

S-CLL cases were analyzed to determine whether mRNA abundances of genes differed to familial cases and whether the same mRNAs that showed changes in abundance among F-Controls, F-MBL and F-CLL were also differentially abundant in combined familial and NK-Controls (Combined Controls), F-MBL, F-CLL and S-CLL (Fig. 5). For six of the genes, there were no differences in mRNA levels between S-CLL cases and F-CLL (Supplementary Material, Table S3). However, there was a difference in *PARP3* levels between S-CLL cases and F-CLL. These results were the same when F-Controls

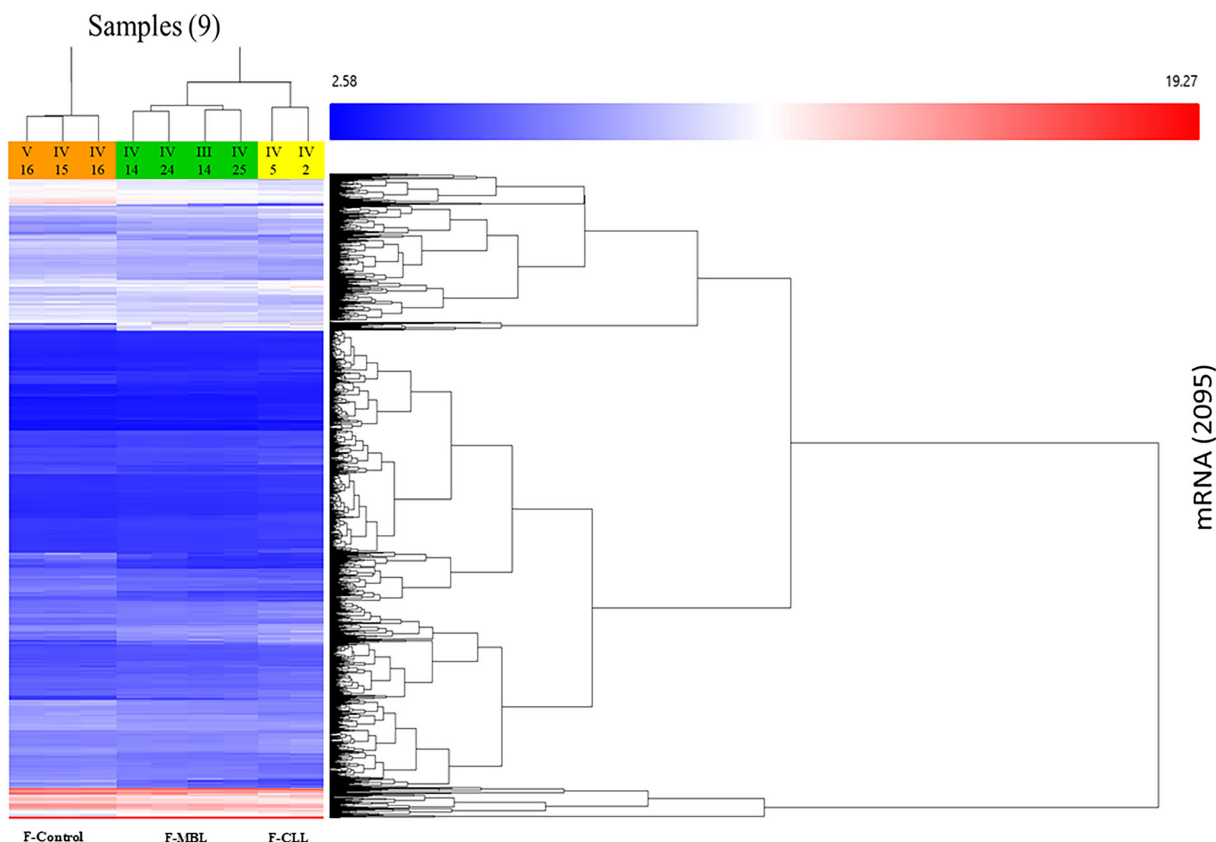

Fig. 2 Hierarchical clustering of B lymphocyte mRNA abundance in F-Controls, F-MBL and F-CLL. Data are displayed as a heat map where rows represent mRNAs and columns represent samples from patients. Colored pixels indicate the magnitude of the response for each gene, where shades of red and blue represent induction and repression, respectively, relative to the median for all genes. Differential expression analysis identified 2095 differentially expressed mRNAs (1794 coding, 301 non-coding) among F-Controls, F-MBL and F-CLL (ANOVA $P < 0.01$; not corrected for multiple comparisons). The range of differential expression (log2) was 1.58 suppression to 19.27 increased expression. The cluster dendrograms at the right segregate F-Controls, F-MBL and F-CLL

were removed from the Combined Controls group (Supplementary Material, Fig. S1).

The abundances of mRNAs for *LEF1*, *GRASP*, *ROR1* and *METTL8* were different between S-CLL and F-MBL; however, there were no differences among *C11orf80*, *PARP3* and *AC016745.3*.

DISCUSSION

In this article we report that mRNA levels of *GRASP* and *AC016745.3* were lower and of *C11orf80*, *PARP3*, *ROR1*, *METTL8* and *LEF1* were higher in enriched B lymphocytes from F-MBL and F-CLL cases compared to F-Control subjects. Furthermore, there were no differences in mRNA levels of *GRASP*, *AC016745.3*, *C11orf80*, *ROR1*, *METTL8* and *LEF1* between F-CLL and S-CLL. *PARP3* was differentially abundant but increased in F-CLL and S-CLL compared to F-Controls and combined F- and NK-Controls. Previous studies have found changes in mRNA levels in both sporadic MBL and early-stage S-CLL cases compared to normal B lymphocytes [26], including a prognostic seven-gene signature (*FMOD*, *PIK3C2B*, *LEF1*, *CKAP4*, *PFTK1*, *BCL-2* and *GPM6a*) [12]. Furthermore, mRNA levels of genes involved in MAPKinase, protein kinase A and proliferation pathways have been

Table 2 mRNAs differentially abundant in F-Controls, F-MBL and F-CLL

Gene	F-Control mean log2 transformed expression[a]	F-MBL mean log2 transformed expression	F-CLL mean log2 transformed expression	ANOVA *P*-value	FDR *P*-value[b]	F-CLL vs F-Control log2 transformed fold change (linear scale fold change)[c]
GRASP	11.1	9.7	6.7	4.09E-07	0.03	− 4.4 (− 21.1)
LEF1	5.0	10.4	11.2	0.000001	0.03	+ 6.2 (+ 73.5)
C11orf80	9.9	11.4	11.8	0.000002	0.03	+ 1.9 (+ 3.7)
ROR1	5.8	10.4	10.6	0.000002	0.03	+ 4.8 (+ 27.9)
AC016745.3	7.1	6.0	5.7	0.000004	0.04	− 1.4 (− 2.6)
METTL8	8.0	10.5	11.2	0.000004	0.04	+ 3.2 (+ 9.2)
PARP3	6.7	7.1	6.9	0.000004	0.04	+ 0.2 (+ 1.1)

[a] Bi-weight average signal (log2) gene intensity for each sample was calculated using the Tukey's Bi weight average for all eligible exons' intensities in that gene

[b] False discovery rate (FDR) *P*-values [22–24]

[c] CLL vs control log2 transformed fold change [$\log_2(\text{CLL/control}) = \log_2(\text{CLL}) - \log_2(\text{control})$]. Linear scale fold-change = $2^{\log2(\text{CLL/control})}$

found to differentiate normal B lymphocytes from sporadic MBL and S-CLL cases [27].

More than 40 mutations have been associated with an inherited risk of CLL [9, 28]. Significantly, susceptibility alleles and haplotypes are enriched in regulatory elements including B-cell transcription factor binding sites, and it is likely that a proportion of the genetic susceptibility to CLL results from mutations that affect gene regulation [28]. Furthermore, non-DNA sequence-based inheritance factors, including epigenetic variations, that regulate gene expression have been described for hereditary cancers [10, 11]. In the present study, the simultaneous presence of F-MBL and F-CLL in a single family provided an opportunity to study changes in mRNA associated with progression to F-CLL against similar genetic backgrounds. In this family, the mRNA levels of *GRASP* and *AC016745.3* were decreased in F-MBL (2.6- and 2.1-fold, respectively) compared to F-Controls and further decreased in F-CLL (21.1- and 2.6-fold, respectively), whereas *C11orf80* and *METTL8* mRNA levels were increased in F-MBL (2.8- and 5.7-fold, respectively) and further increased in F-CLL (3.7- and 9.2-fold, respectively). The mRNA levels of *ROR1* and *LEF1* were also higher in F-CLL compared to F-Controls (27.9- and 73.5-fold, respectively); however, there were no differences between F-CLL and F-MBL, and for *PARP3*, levels were higher in F-MBL (1.3-fold) but less so for F-CLL (1.1-fold) compared to F-Controls.

The incidence of CLL increases with age; however, familial cases are more likely to be younger (≤ 55 years) than sporadic cases [29], and consequently age-matching of cases and F-Controls was difficult for this single family-based study. The younger F-Controls may develop F-CLL in the future, which would be expected to reduce differences in expression among the seven mRNAs that were differentially abundant between F-Controls and F-MBL or F-CLL cases.

Of the seven differential genes identified, *LEF1* and *ROR1* have previously been associated with either the development of CLL or progression of MBL to CLL [30–36]. The transcription factor, *LEF1*, is involved in the development of B lymphocytes and is highly expressed in mouse pro-B and pre-B lymphocytes but downregulated in mature B cells [37]. LEF1 functions in the Wnt/β-catenin signaling pathway, recruiting β-catenin to activate

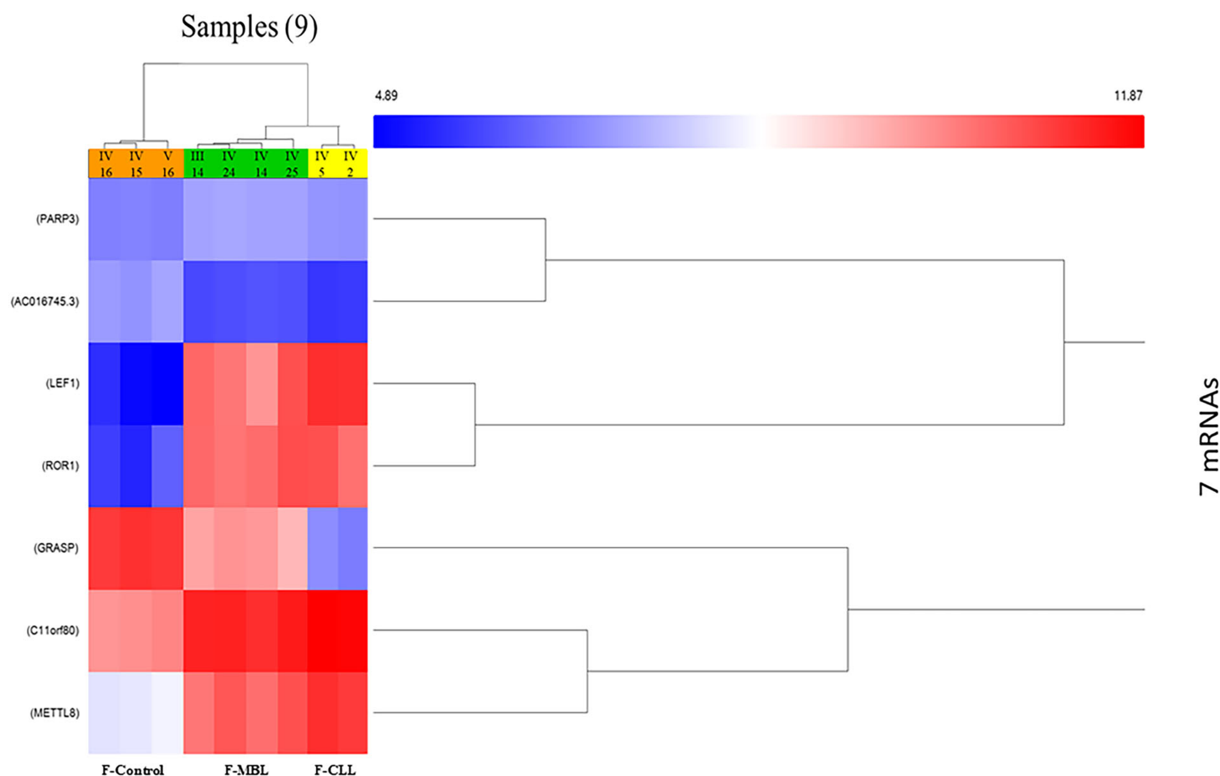

Fig. 3 Hierarchical clustering of B lymphocyte mRNA levels in F-Controls, F-MBL and F-CLL cases. Array elements that significantly varied between groups (FDR < 0.05) were included (7 mRNAs). The range of differential expression (log2) was 4.89 suppression (LEF1) to 11.87 increased expression (C11orf80). The cluster dendrograms at right segregate F-Controls, F-MBL and F-CLL

transcription of several target genes in response to constitutive Wnt pathway activation, which regulates B lymphocyte proliferation and survival [31]. CLL cells aberrantly express *LEF1* compared to normal B lymphocytes and *LEF1* knockdown or LEF1 inhibition by small molecule decreases CLL B-cell survival [31, 38].

ROR1 signaling is involved in cell proliferation and differentiation, and over-expression of ROR1 on the surface of B-CLL has been documented in several studies [33, 39]. ROR1 acts as a receptor for Wnt5 signaling, which increases CLL cell survival, proliferation and migration [40]. These effects are blocked by cirmtuzumab, a humanized anti-ROR1 monoclonal antibody [40]. siRNA silencing of *ROR1* in CLL cells induces apoptosis of B-CLL cells but not control B cells [41]. Consequently, ROR1 has been considered as a target for new CLL therapies [42].

This study identified five novel associations, of which *PARP3, GRASP*, *METTL8* and *C11orf80* may be plausible candidate genes associated with the neoplastic transformation of B lymphocytes. *PARP3* facilitates the formation and maintenance of the mitotic spindle and genome integrity [43] and is a potential target for cancer therapy [44]. The *GRASP* gene encodes the general receptor for phosphoinositide 1-associated scaffold protein, which promotes ADP ribosylation factors to Rac signaling networks and cell migration [45, 46]. Consistent with our findings, *GRASP* has been found to be down-regulated in CLL compared to control B lymphocytes [47]. *METTL8* encodes a methyltransferase associated with CLL [48] and may be responsible for epigenetic effects in CLL. *C11orf80* encodes a component of a topoisomerase 6 complex specifically required for meiotic recombination and may be a potential

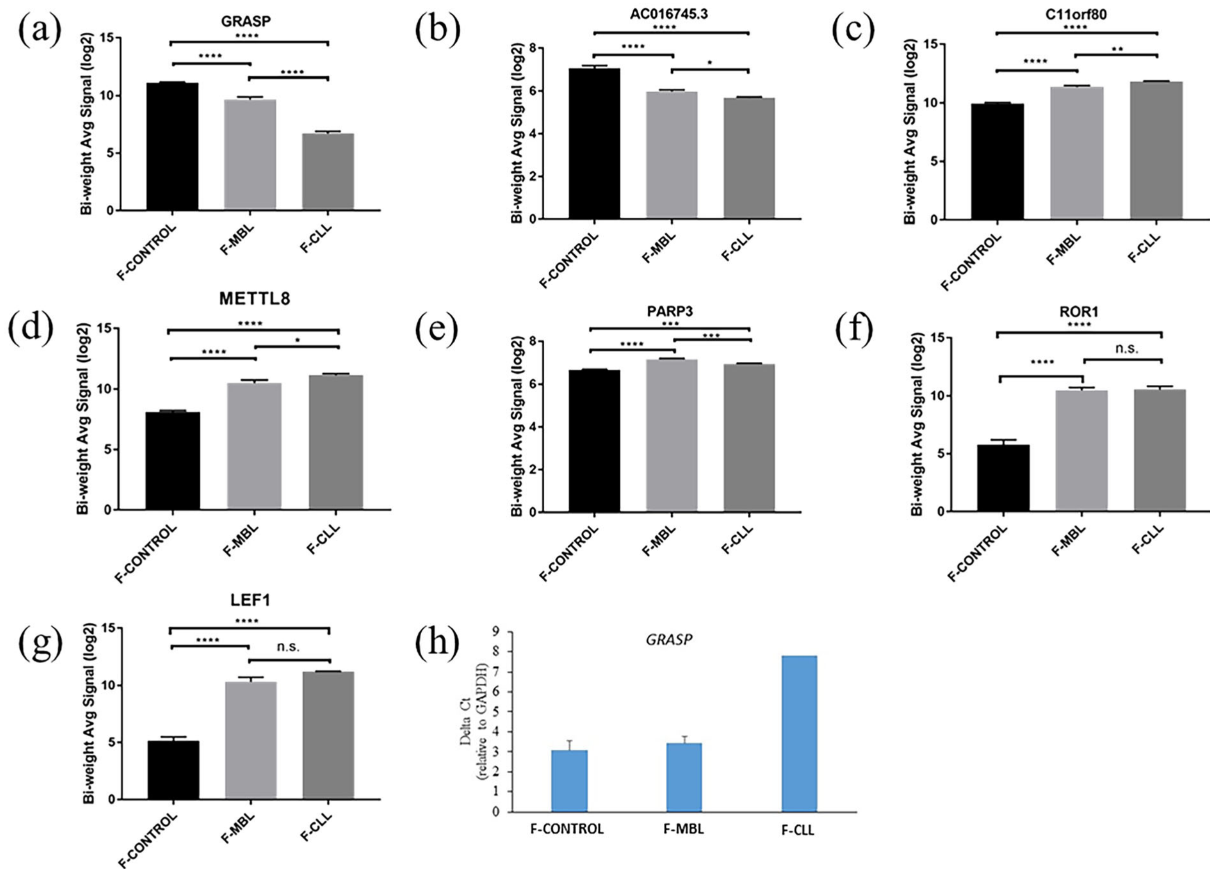

Fig. 4 Comparison of seven mRNAs differentially abundant among F-Controls, F-MBL and F-CLL cases. **a–g** Bi-weight average signal (log2) intensity for each of the seven mRNAs found to be differentially abundant among F-Controls, F-MBL and F-CLL. Significances were determined using one-way ANOVA with Tukey's post hoc test. $P \leq 0.05$ values are summarized with 1 asterisk, $P \leq 0.01$ with 2 asterisks, $P \leq 0.001$ with 3 asterisks and $P \leq 0.0001$ with 4 asterisks. **h** Real-time reverse transcription-PCR (qRT-PCR) validation of microarray results for GRASP. qRT-PCR was performed in F-Controls ($n = 3$), F-MBL ($n = 4$) and F-CLL ($n = 1$). Changes in expression were determined relative to GAPDH (delta Ct)

target for treatment if overexpressed in CLL cells.

Although CLL and F-MBL samples were not 100% pure and contained contaminating CD20 +, CD5- B lymphocytes, for comparisons of mRNA expression there were no differences among the purity of F-MBL CD20 +, CD5 + cases and F-CLL using flow cytometry (mean purity 83% versus 94%, respectively; ns, Student's *t*-test) and CD5 mRNA expression in F-MBL and F-CLL cases (mean log2 bi-weight avg signal 8.7 versus 9.5, respectively, ns, Student's *t*-test). Furthermore, the possibility of activating downstream pathways was reduced by using a negative selection method to purify CLL and F-MBL cells rather than positively sorting CD5 + cells, which induces protein kinase C signaling [49].

CONCLUSIONS

In conclusion, although studies of single families are limited by small numbers, identification of differentially abundant mRNAs in normal B lymphocytes, F-MBL and CLL cells has provided new molecular signatures for targeted therapies. Significantly, the similarities between F-CLL and S-CLL in this study and previous studies

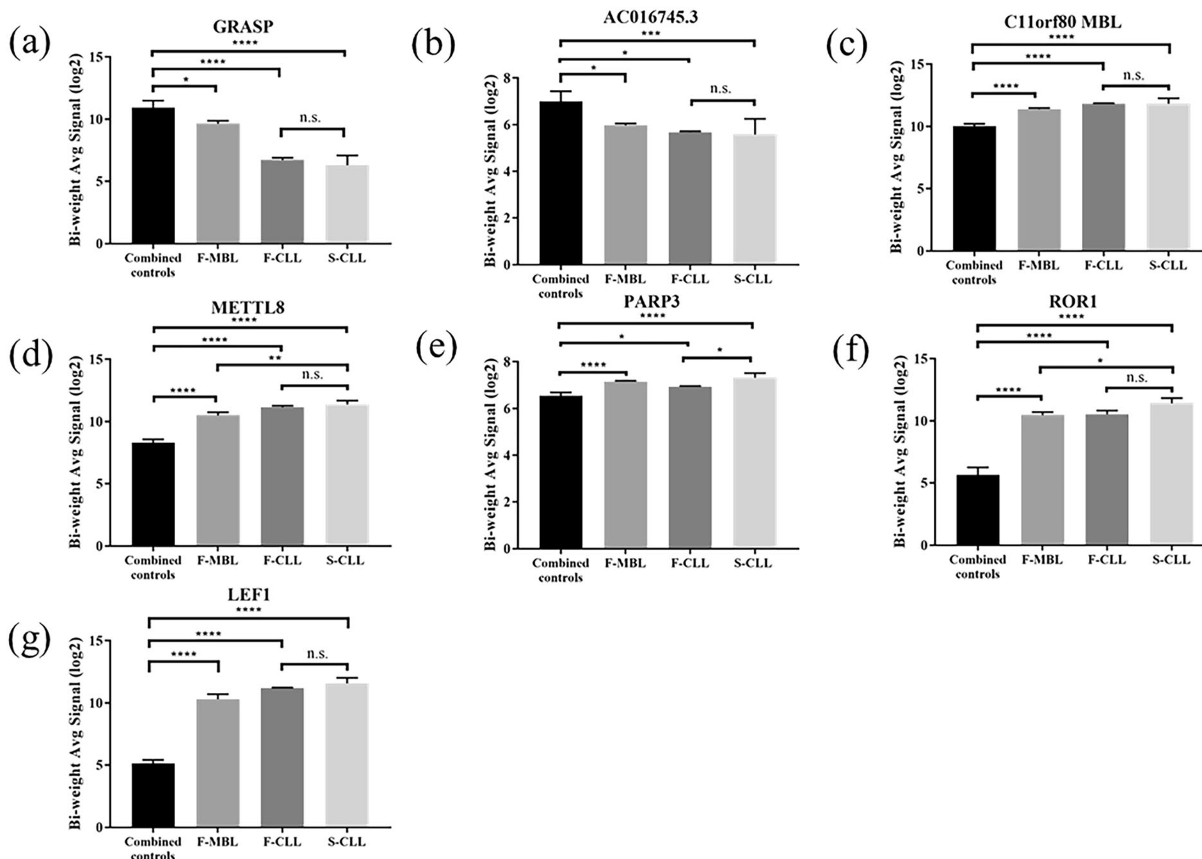

Fig. 5 Comparison of seven mRNAs among combined F-Controls and NK-Controls, F-MBL, F-CLL and S-CLL cases. **a-g** Bi-weight average signal (log2) intensity for each of the seven mRNAs among combined familial and NK-Controls (Combined Controls), F-MBL and F-CLL compared to S-CLL. Significances were determined using one-way ANOVA with Tukey's post hoc test. $P \leq 0.05$ values are summarized with 1 asterisk, $P \leq 0.01$ with 2 asterisks, $P \leq 0.001$ with 3 asterisks, and $P \leq 0.0001$ with 4 asterisks. ns: $P > 0.05$

[50] indicate that findings from familial studies may translate to sporadic cases.

This study was limited by the small sample sizes, especially for F-CLL, and inability to standardize the collection times of samples after diagnosis. In addition, all F-MBL samples were *IgVH* mutated, which may alter expression of downstream effectors.

ACKNOWLEDGEMENTS

We are grateful to all subjects who provided blood samples.

Funding. The Cultural Mission of Saudi Arabia provided principal funding for the study. We acknowledge generous support from Gary and Marion Krelle and the Nepean Medical Research Fund. The Rapid Service Fee was funded by the authors.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. All authors contributed to the study conception and design.

Material preparation, data collection and analysis were performed by Abdullah Alshahrani, Kristen K. Skarratt, Maryam Hassanvand, Benjamin Tang and Stephen J. Fuller. Statistical analyses were performed by Kristy P. Robledo, Abdullah Alshahrani and Stephen J. Fuller. The first draft of the manuscript was written by Stephen J. Fuller, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Disclosures. Abdullah Alshahrani declares that he has no conflict of interest. Kristen K. Skarratt declares that she has no conflict of interest. Kristy P. Robledo declares that she has no conflict of interest. Maryam Hassanvand declares that she has no conflict of interest. Benjamin Tang declares that he has no conflict of interest. Stephen J. Fuller declares that he has no conflict of interest.

Compliance with Ethics Guidelines. The experimental protocol was approved by the Nepean and Blue Mountains Local Health District Human Research Ethics Committee (01/70). The study was performed in accordance with the Helsinki Declaration of 1964 and its later amendments. All subjects provided informed consent to participate in the study.

Data Availability. All data generated or analyzed during this study are included in this published article/as supplementary information files.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018;68(1):7–30.
2. Bevan S, Catovsky D, Matutes E, Antunovic P, Auger MJ, Ben-Bassat I, et al. Linkage analysis for major histocompatibility complex-related genetic susceptibility in familial chronic lymphocytic leukemia. Blood. 2000;96(12):3982–4.
3. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. Int J Cancer. 2002;99(2):260–6.
4. Rawstron A, Hillmen P, Houlston R. Clonal lymphocytes in persons without known chronic lymphocytic leukemia (CLL): implications of recent findings in family members of CLL patients. Semin Hematol. 2004;41(3):192–200.
5. Rawstron AC, Yuille MR, Fuller J, Cullen M, Kennedy B, Richards SJ, et al. Inherited predisposition to CLL is detectable as subclinical monoclonal B-lymphocyte expansion. Blood. 2002;100(7):2289–90.
6. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2008;40(10):1204–10.
7. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, Dobbins SE, Torres M, Mansouri M, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet. 2010;42(2):132–6.
8. Speedy HE, Di Bernardo MC, Sava GP, Dyer MJ, Holroyd A, Wang Y, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2014;46(1):56–60.
9. Kleinstern G, Camp NJ, Goldin LR, Vachon CM, Vajdic CM, de Sanjose S, et al. Association of polygenic risk score with the risk of chronic

lymphocytic leukemia and monoclonal B-cell lymphocytosis. Blood. 2018;131(23):2541–51.

10. Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Res. 2002;62(14):3925–8.

11. Esteller M, Fraga MF, Guo M, Garcia-Foncillas J, Hedenfalk I, Godwin AK, et al. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet. 2001;10(26): 3001–7.

12. McCarthy BA, Yancopoulos S, Tipping M, Yan XJ, Wang XP, Bennett F, et al. A seven-gene expression panel distinguishing clonal expansions of pre-leukemic and chronic lymphocytic leukemia B cells from normal B lymphocytes. Immunol Res. 2015;63(1–3):90–100.

13. Fuller SJ, Papaemmanuil E, McKinnon L, Webb E, Sellick GS, Dao-Ung LP, et al. Analysis of a large multi-generational family provides insight into the genetics of chronic lymphocytic leukemia. Br J Haematol. 2008;142(2):238–45.

14. Borecki IB, Province MA. Genetic and genomic discovery using family studies. Circulation. 2008;118(10):1057–63.

15. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood. 2008;111(12):5446–56.

16. Essakali S, Carney D, Westerman D, Gambell P, Seymour JF, Dobrovic A. Negative selection of chronic lymphocytic leukaemia cells using a bifunctional rosette-based antibody cocktail. BMC Biotechnol. 2008;8:6.

17. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12): 2257–317.

18. van Krieken JH, Langerak AW, Macintyre EA, Kneba M, Hodges E, Sanz RG, et al. Improved reliability of lymphoma diagnostics via PCR-based clonality testing: report of the BIOMED-2 Concerted Action BHM4-CT98-3936. Leukemia. 2007;21(2):201–6.

19. Ghia P, Stamatopoulos K, Belessi C, Moreno C, Stilgenbauer S, Stevenson F, et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia. 2007;21(1):1–3.

20. Skarratt KK, Fuller SJ. Quantitative real-time PCR eliminates false-positives in colony screening PCR. J Microbiol Methods. 2014;96:99–100.

21. Yuan JS, Reed A, Chen F, Stewart CN. Statistical analysis of real-time PCR data. BMC Bioinform. 2006;7(1):85.

22. Pounds SB. Estimation and control of multiple testing error rates for microarray studies. Brief Bioinform. 2006;7(1):25–36.

23. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics. 2003;19(3):368–75.

24. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological). 1995;57(1):289–300.

25. Binet JL, Auquier A, Dighiero G, Chastang C, Piguet H, Goasguen J, et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer. 1981;48(1): 198–206.

26. Morabito F, Mosca L, Cutrona G, Agnelli L, Tuana G, Ferracin M, et al. Clinical monoclonal B lymphocytosis versus Rai 0 chronic lymphocytic leukemia: a comparison of cellular, cytogenetic, molecular, and clinical features. Clin Cancer Res. 2013;19(21):5890–900.

27. Lanasa MC, Allgood SD, Slager SL, Dave SS, Love C, Marti GE, et al. Immunophenotypic and gene expression analysis of monoclonal B-cell lymphocytosis shows biologic characteristics associated with good prognosis CLL. Leukemia. 2011;25(9): 1459–66.

28. Law PJ, Sud A, Mitchell JS, Henrion M, Orlando G, Lenive O, et al. Genome-wide association analysis of chronic lymphocytic leukaemia, Hodgkin lymphoma and multiple myeloma identifies pleiotropic risk loci. Sci Rep. 2017;7:41071.

29. Parikh SA, Rabe KG, Kay NE, Call TG, Ding W, Schwager SM, et al. Chronic lymphocytic leukemia in young (</= 55 years) patients: a comprehensive analysis of prognostic factors and outcomes. Haematologica. 2014;99(1):140–7.

30. Wu W, Zhu H, Fu Y, Shen W, Miao K, Hong M, et al. High LEF1 expression predicts adverse prognosis in

chronic lymphocytic leukemia and may be targeted by ethacrynic acid. Oncotarget. 2016;7(16): 21631–43.

31. Gutierrez A Jr, Tschumper RC, Wu X, Shanafelt TD, Eckel-Passow J, Huddleston PM 3rd, et al. LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. Blood. 2010;116(16): 2975–83.

32. van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L, et al. Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. Genes Dev. 1994;8(22):2691–703.

33. Baskar S, Kwong KY, Hofer T, Levy JM, Kennedy MG, Lee E, et al. Unique cell surface expression of receptor tyrosine kinase ROR1 in human B-cell chronic lymphocytic leukemia. Clin Cancer Res. 2008;14(2):396–404.

34. Palermo G, Maisel D, Barrett M, Smith H, Duchateau-Nguyen G, Nguyen T, et al. Gene expression of INPP5F as an independent prognostic marker in fludarabine-based therapy of chronic lymphocytic leukemia. Blood Cancer J. 2015;5:e353.

35. Douglas RS, Capocasale RJ, Lamb RJ, Nowell PC, Moore JS. Chronic lymphocytic leukemia B cells are resistant to the apoptotic effects of transforming growth factor-β. Blood. 1997;89(3):941–7.

36. Matveeva A, Kovalevska L, Kholodnyuk I, Ivanivskaya T, Kashuba E. The TGF-beta - SMAD pathway is inactivated in cronic lymphocytic leukemia cells. Exp Oncol. 2017;39(4):286–90.

37. Reya T, O'Riordan M, Okamura R, Devaney E, Willert K, Nusse R, et al. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. Immunity. 2000;13(1):15–24.

38. Gandhirajan RK, Staib PA, Minke K, Gehrke I, Plickert G, Schlosser A, et al. Small molecule inhibitors of Wnt/beta-catenin/lef-1 signaling induces apoptosis in chronic lymphocytic leukemia cells in vitro and in vivo. Neoplasia. 2010;12(4):326–35.

39. Daneshmanesh AH, Mikaelsson E, Jeddi-Tehrani M, Bayat AA, Ghods R, Ostadkarampour M, et al. Ror1, a cell surface receptor tyrosine kinase is expressed in chronic lymphocytic leukemia and may serve as a putative target for therapy. Int J Cancer. 2008;123(5):1190–5.

40. Yu J, Chen L, Cui B, Widhopf GF 2nd, Shen Z, Wu R, et al. Wnt5a induces ROR1/ROR2 heterooligomerization to enhance leukemia chemotaxis and proliferation. J Clin Invest. 2016;126(2):585–98.

41. Choudhury A, Derkow K, Daneshmanesh AH, Mikaelsson E, Kiaii S, Kokhaei P, et al. Silencing of ROR1 and FMOD with siRNA results in apoptosis of CLL cells. Br J Haematol. 2010;151(4):327–35.

42. Aghebati-Maleki L, Shabani M, Baradaran B, Motallebnezhad M, Majidi J, Yousefi M. Receptor tyrosine kinase-like orphan receptor 1 (ROR-1): An emerging target for diagnosis and therapy of chronic lymphocytic leukemia. Biomed Pharmacother. 2017;88:814–22.

43. Boehler C, Gauthier LR, Mortusewicz O, Biard DS, Saliou JM, Bresson A, et al. Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. Proc Natl Acad Sci U S A. 2011;108(7):2783–8.

44. Oplustil O'Connor L, Rulten SL, Cranston AN, Odedra R, Brown H, Jaspers JE, et al. The PARP inhibitor AZD2461 provides insights into the role of PARP3 inhibition for both synthetic lethality and tolerability with chemotherapy in preclinical models. Cancer Res. 2016;76(20):6084–94.

45. White DT, McShea KM, Attar MA, Santy LC. GRASP and IPCEF promote ARF-to-Rac signaling and cell migration by coordinating the association of ARNO/cytohesin 2 with Dock180. Mol Biol Cell. 2010;21(4):562–71.

46. Attar MA, Santy LC. The scaffolding protein GRASP/Tamalin directly binds to Dock180 as well as to cytohesins facilitating GTPase crosstalk in epithelial cell migration. BMC Cell Biol. 2013;14(1):9.

47. Liao W, Jordaan G, Nham P, Phan RT, Pelegrini M, Sharma S. Gene expression and splicing alterations analyzed by high throughput RNA sequencing of chronic lymphocytic leukemia specimens. BMC Cancer. 2015;15(1):1.

48. Chadeau-Hyam M, Vermeulen RC, Hebels DG, Castagne R, Campanella G, Portengen L, et al. Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. Ann Oncol. 2014;25(5):1065–72.

49. Perez-Chacon G, Vargas JA, Jorda J, Morado M, Rosado S, Martin-Donaire T, et al. CD5 provides viability signals to B cells from a subset of B-CLL patients by a mechanism that involves PKC. Leuk Res. 2007;31(2):183–93.

50. Brown JR. Inherited susceptibility to chronic lymphocytic leukemia: evidence and prospects for the future. Ther Adv Hematol. 2013;4(4):298–308.