VOLUME 32 · NUMBER 13 · MAY 1 2014

JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

t(14;18) Translocation: A Predictive Blood Biomarker for Follicular Lymphoma

Sandrine Roulland, Rachel S. Kelly, Ester Morgado, Stéphanie Sungalee, Philippe Solal-Celigny, Philippe Colombat, Nathalie Jouve, Domenico Palli, Valeria Pala, Rosario Tumino, Salvatore Panico, Carlotta Sacerdote, José R. Quirós, Carlos A. Gonzáles, Maria-José Sánchez, Miren Dorronsoro, Carmen Navarro, Aurelio Barricarte, Anne Tjønneland, Anja Olsen, Kim Overvad, Federico Canzian, Rudolf Kaaks, Heiner Boeing, Dagmar Drogan, Alexandra Nieters, Françoise Clavel-Chapelon, Antonia Trichopoulou, Dimitrios Trichopoulos, Pagona Lagiou, H. Bas Bueno-de-Mesquita, Petra H.M. Peeters, Roel Vermeulen, Göran Hallmans, Beatrice Melin, Signe Borgquist, Joyce Carlson, Eiliv Lund, Elisabete Weiderpass, Kay-Tee Khaw, Nick Wareham, Timothy J. Key, Ruth C. Travis, Pietro Ferrari, Isabelle Romieu, Elio Riboli, Gilles Salles, Paolo Vineis, and Bertrand Nadel

Author affiliations appear at the end of this article.

Published online ahead of print at www.jco.org on March 31, 2014.

Support information appears at the end of this article.

P.V. and B.N. contributed equally to this work.

Presented orally at the 12th International Conference on Malignant Lymphoma, Lugano, Switzerland, June 19-22, 2013; by invitation at the 13th International Lymphoma Epidemiology Consortium, Dijon, France, June 24-26, 2013; and orally at the 55th Annual Meeting of the American Society of Hematology, New Orleans, LA, December 7-10, 2013.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Sandrine Roulland, PharmD, PhD, Centre d'Immunologie de Marseille-Luminy, Aix-Marseille University, 13288 Marseille Cedex 9, France; e-mail: roulland@ciml.univ-mrs.fr.

0732-183X/14/3213w-1347w/$20.00

DOI: 10.1200/JCO.2013.52.8190

ABSTRACT

Purpose
The (14;18) translocation constitutes both a genetic hallmark and critical early event in the natural history of follicular lymphoma (FL). However, t(14;18) is also detectable in the blood of otherwise healthy persons, and its relationship with progression to disease remains unclear. Here we sought to determine whether t(14;18)-positive cells in healthy individuals represent tumor precursors and whether their detection could be used as an early predictor for FL.

Participants and Methods
Among 520,000 healthy participants enrolled onto the EPIC (European Prospective Investigation Into Cancer and Nutrition) cohort, we identified 100 who developed FL 2 to 161 months after enrollment. Prediagnostic blood from these and 218 controls were screened for t(14;18) using sensitive polymerase chain reaction–based assays. Results were subsequently validated in an independent cohort (65 case participants; 128 controls). Clonal relationships between t(14;18) cells and FL were also assessed by molecular backtracking of paired prediagnostic blood and tumor samples.

Results
Clonal analysis of t(14;18) junctions in paired prediagnostic blood versus tumor samples demonstrated that progression to FL occurred from t(14;18)-positive committed precursors. Furthermore, healthy participants at enrollment who developed FL up to 15 years later showed a markedly higher t(14;18) prevalence and frequency than controls ($P < .001$). Altogether, we estimated a 23-fold higher risk of subsequent FL in blood samples associated with a frequency $> 10^{-4}$ (odds ratio, 23.17; 95% CI, 9.98 to 67.31; $P < .001$). Remarkably, risk estimates remained high and significant up to 15 years before diagnosis.

Conclusion
High t(14;18) frequency in blood from healthy individuals defines the first predictive biomarker for FL, effective years before diagnosis.

J Clin Oncol 32:1347-1355. © 2014 by American Society of Clinical Oncology

INTRODUCTION

Follicular lymphoma (FL), the second most common adult B-cell lymphoma in Western countries, is usually characterized by an indolent clinical course, evolving asymptomatically over many years. Consequently, diagnosis is frequently delayed, and treatment performed on a largely disseminated tumor. Although the advent of highly effective therapies and the availability of anti-CD20 antibodies have significantly improved clinical outcome, FL remains incurable, and patients continue to die as a result of the disease after development of resistance to treatments or disease transformation into aggressive diffuse large B-cell lymphoma.[1,2] FL represents a particularly attractive model to study the early phases of cancer development, because the acquisition of the genetic hallmark t(14;18) translocation—the earliest recurrent event giving rise to a *BCL2-IGH* fusion—originates from bone marrow (BM) pre–B cell; conversely, malignant FL clones derive from the transformation of germinal center B cells

in secondary lymphoid organs.[3,4] Accordingly, a large proportion of healthy individuals (50% to 70%) harbor low levels of circulating t(14;18)-positive cells[5-7] but will never develop FL, indicating that BCL2 ectopic expression is necessary but not sufficient for tumor progression.[8] However, the relationship between t(14;18) and progression to disease remains unclear. In rare cases, apparently tumor-free individuals carry an unusually high t(14;18) frequency (10 to 100× greater than average population),[7] and sporadic cases of progression to FL have been reported.[9-11] We and others previously demonstrated that such high frequency of t(14;18)-positive cells constitutes an expanding clonal population of atypical B cells, issued from the germinal center and sharing illegitimate genotypic and phenotypic features exclusively seen in FL.[12,13] Nevertheless, key gaps in our understanding are whether such t(14;18)-positive cells with FL-like features in healthy individuals constitute committed FL precursors, and whether high t(14;18) frequency in blood represents a suitable predictive biomarker of FL development.

PARTICIPANTS AND METHODS

Study Population

Discovery cohort. We conducted a nested case-control study on FL risk within the EPIC (European Prospective Investigation Into Cancer and Nutrition) cohort. Approximately 520,000 healthy individuals (age 35 to 70 years) were recruited between 1992 and 2000 from 23 centers in 10 European countries (Appendix, online only). At recruitment, participants were asked to complete a self-administered questionnaire to collect demographic, medical, and lifestyle information. Eligible participants were those diagnosed with FL during follow-up, according to the International Classification of Diseases for Oncology, third revision (ICD-O-3), who had an archived cryopreserved buffy-coat sample collected at enrollment. For each participant case, two controls were selected from among cohort members alive and cancer free at the time participant cases were diagnosed and matched by study center, age (± 1 year), sex, and blood collection date (± 45 days). Twenty-one participant cases were excluded because of insufficient DNA quality. Participant cases with ICD-O-3 codes that could not be unequivocally assigned to FL were also excluded (n = 15). Approval was obtained from the International Agency for Research on Cancer Ethics Board and local review boards of participating institutions. Informed consent was obtained from all participants.

Validation Cohort. The validation cohort included samples from the prospective population-based NSHDS (Northern Sweden Health and Disease Study). According to the eligibility criteria used for EPIC, 68 incident FL participant cases and 128 matched controls were identified through the Swedish Cancer Registries.

French Clinical Series of Patients With FL. Blood (n = 25) and BM (n = 31) DNA samples were obtained from patients with FL enrolled onto a clinical study and collected at diagnosis before therapy.[14] All patients carried a low tumor burden according to Groupe d'Etude des Lymphomes Folliculaires criteria.[15]

Screening Assay

A quantitative polymerase chain reaction (PCR) assay with primers and probes for the *BCL2-IGH* and reference *GAPDH* genes was used to quantify t(14;18) frequency in buffy coat–extracted DNA samples (Appendix Table A1, online only; Appendix Fig A1, online only). DNA was encoded by EPIC/NSHDS investigators and sent for blind screening to the Marseille laboratory. DNA of each individual was tested in triplicate using 0.5- to 1-μg DNA for t(14;18) and two replicates for *GAPDH*. Standard curves were generated from cloned *BCL2-IGH* and *GAPDH* PCR products, and absolute frequency of t(14;18)-positive cells was calculated using the standard curve method, normalized to the amount of total cells in the samples (Appendix, online only).

Statistical Analysis

Differences in t(14;18) prevalence and frequency were analyzed using χ^2 and Mann-Whitney U tests, respectively. Correlations between frequency and time to diagnosis were performed using Spearman's rank test. A receiver operating characteristic curve analysis was used to test the overall ability of t(14;18) frequency to identify incident FL cases.[16] Frequency was then dichotomized based on several predefined thresholds, and the risk of FL associated with a frequency greater than each of these thresholds was calculated. Unconditional logistic regression was applied to determine the association between the frequency thresholds and FL, adjusting for age, sex, time since blood draw, and country of residence. An unconditional model was used to maximize power, because after exclusion of poor-quality DNA samples, there was no longer a one-to-two ratio for all participant cases and controls, and therefore, using matched analysis would have resulted in additional exclusions. The optimal threshold was determined based on sensitivity, specificity, and association with risk. This threshold was applied to the validation cohort to confirm its ability to predict FL risk in an independent cohort. Meta-analysis of both cohorts was used to assess heterogeneity and provide a summary risk estimate. Stratified analyses of time to diagnosis were performed on the pooled data set, and the Kaplan-Meier method was used to construct curves according to t(14;18) status (Appendix, online only).

RESULTS

t(14,18) Prevalence and Frequency in Blood From Healthy Individuals Before FL Diagnosis

We investigated FL onset in the EPIC cohort, which includes more than 520,000 healthy participants with a follow-up for cancer incidence of more than 15 years.[17] We identified 100 participants who developed FL during follow-up (here referred to as incident FL) and had good-quality archived prediagnostic blood collected at enrollment (Table 1). Time between blood draw and FL diagnosis ranged from 2 months to 13 years (mean, 6.4 years). The control population included 218 participants who did not develop lymphoma during follow-up (here referred to as controls), matched for age, sex, study center, and time since blood draw. This constituted the discovery cohort. A blinded screening quantitative PCR–based assay was performed to evaluate t(14;18) prevalence and quantify t(14;18) frequency. The assay enabled high experimental sensitivity (< one translocation per 500,000 cells) and covered approximately 70% of *BCL2* breakpoints.

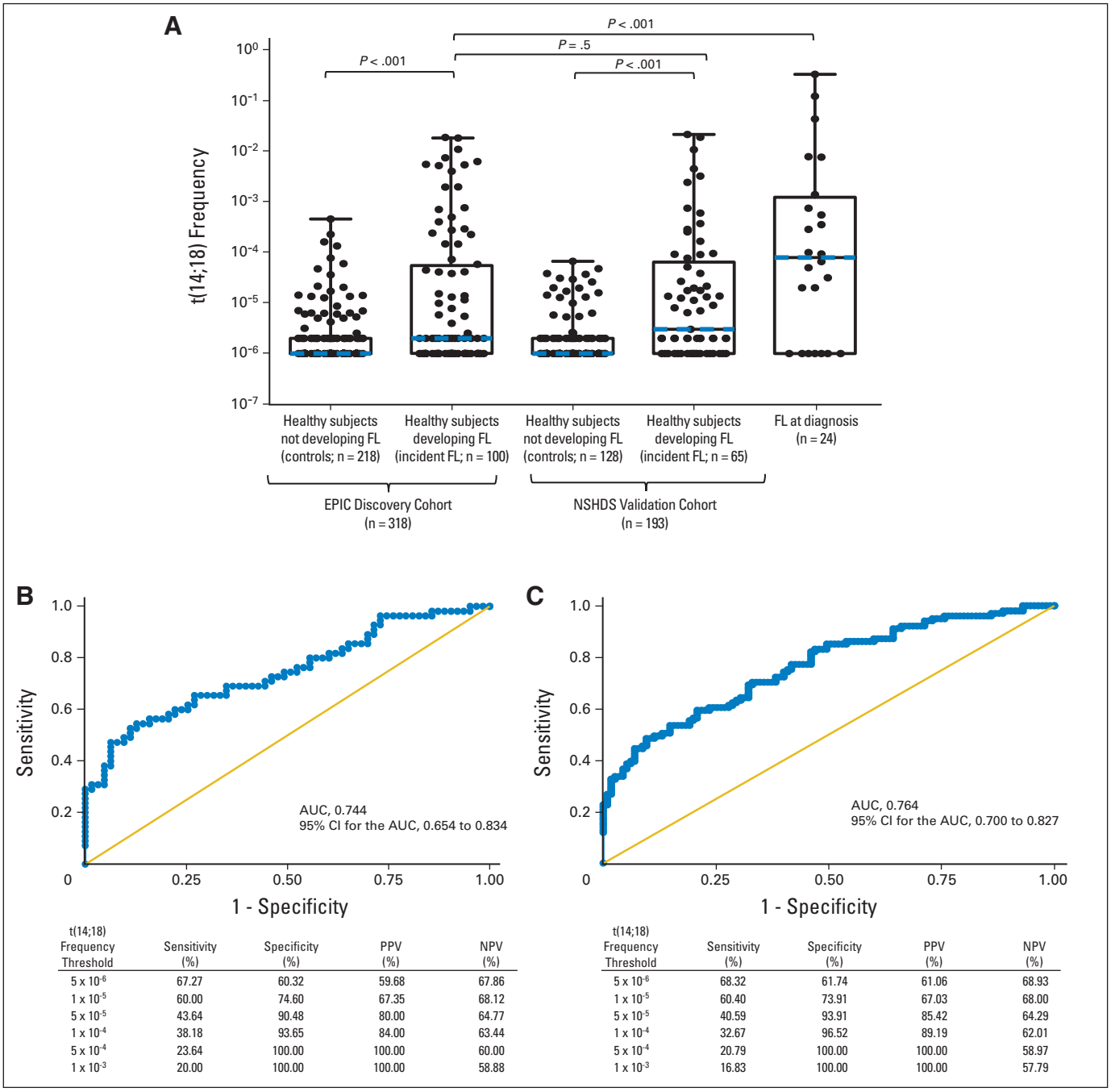
Results showed that prevalence of t(14;18) was significantly higher in incident FL participant cases compared with controls (56% *v* 28.9%; $P < .001$; Table 1). Similarly, the mean and median frequencies were significantly higher, both when considering the total population and t(14;18)-positive participants only (Table 1; Fig 1A). Although t(14;18) frequency varied over a 4-log range (from undetectable to one every 50 cells), the highest frequency quartile in prediagnostic FL reached levels unseen in the control group and never observed previously in large screenings of healthy participants, to our knowledge.[7] Because blood samples at time of FL clinical diagnosis from the incident participant cases with FL were not available, we selected blood samples from an independent series of 24 untreated patients with FL with low tumor burden.[14] Although the frequency distribution was distinct ($P < .001$), it spread over 5 logs, and the range of higher frequencies in the FL series largely overlapped with prediagnostic FL participant cases (Figs 1A to 1C).

from 131.211.105.239

Table 1. Participant Demographic and Clinical Characteristics and t(14;18) Analysis

Characteristic	EPIC Discovery Cohort					NSHDS Validation Cohort					Total Cohort				
	Did Not Develop FL (controls)		Developed FL (incident participant cases)			Did Not Develop FL (controls)		Developed FL (incident FL)			Did Not Develop FL (controls)		Developed FL (incident FL)		
	No.	%	No.	%	P*	No.	%	No.	%	P*	No.	%	No.	%	P*
No. of participants	218		100			128		65			346		165		
Age at blood draw, years										.03					
Mean	54.8		54.1			55.8		52.9			55.2		51.9		
Range	27.3-77.7		27.2-77.2			30.0-73.70		30.3-72.0			27.2-77.5		27.2-77.5		
Age at FL diagnosis, years															
Mean	NA		60.3			NA		63.1					61.6		
Range			37.7-79.2					38.4-81.7					37.7-81.7		
Sex															
Male	66	30.3	32	32.0		65	50.8	32	49.2		131	37.9	64	38.8	
Female	152	69.7	68	68.0		63	49.2	33	50.8		215	62.1	101	61.2	
Calendar years of enrollment	1993-2002		1993-2002			1990-2010		1990-2010			1990-2010		1993-2010		
Calendar years of FL diagnosis	NA		1995-2007			NA		1994-2013			NA		1994-2013		
Time from blood collection to diagnosis, months															
Mean	NA		76.9			NA		124.0			NA		91		
Range			2.7-161.3					5.9-244.5					2-241		
t(14;18) prevalence	63 of 218	28.9	56 of 100	56.0	< .001†	52 of 128	40.6	46 of 65	70.8	< .001†	115 of 346	33.0	102 of 165	61.8	< .001†
t(14;18) frequency, $\times 10^{-5}$					< .001‡					.003‡					< .001‡
Mean	0.8		97			0.4		98.3			0.6		94.1		
Range	0.1-45.6		0.1-1,860			0.1-6.7		0.1-2,129.5							
Frequency in t(14;18)–positive samples, $\times 10^{-5}$					< .01‡					.026‡					< .001‡
Mean	2.3		162.0			0.90		139			1.71		153		
Range						0.1-6.6		0.1-2130			0.1-45.6		0.1-2130		
Median t(14;18) frequency, $\times 10^{-5}$	0.1		0.2		< .001‡	0.1		0.3		< .001‡	0.1		0.2		< .001‡
Median frequency in t(14;18)-positive samples, $\times 10^{-5}$	0.2		3.8		< .001‡	0.2		1.6		< .001‡	0.2		1.8		< .001‡
Prevalence of t(14;18)-positive samples, $\times 10^{-4}$	4 of 218	1.8	21 of 100	12.0	< .001†	1 of 128	0.78	12 of 65	18.5	< .001†	5 of 346	1.4	33 of 165	20	< .001†

Abbreviations: EPIC, European Prospective Investigation Into Cancer and Nutrition; FL, follicular lymphoma; NA, not applicable; NSHDS, Northern Sweden Health and Disease Study.
*Only significant values are indicated.
†χ^2 test.
‡Mann-Whitney *U* test.



t(14;18) Frequency Threshold	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
5×10^{-6}	67.27	60.32	59.68	67.86
1×10^{-5}	60.00	74.60	67.35	68.12
5×10^{-5}	43.64	90.48	80.00	64.77
1×10^{-4}	38.18	93.65	84.00	63.44
5×10^{-4}	23.64	100.00	100.00	60.00
1×10^{-3}	20.00	100.00	100.00	58.88

t(14;18) Frequency Threshold	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
5×10^{-6}	68.32	61.74	61.06	68.93
1×10^{-5}	60.40	73.91	67.03	68.00
5×10^{-5}	40.59	93.91	85.42	64.29
1×10^{-4}	32.67	96.52	89.19	62.01
5×10^{-4}	20.79	100.00	100.00	58.97
1×10^{-3}	16.83	100.00	100.00	57.79

Fig 1. Analysis of t(14;18) frequency. (A) t(14;18) frequency in blood from healthy individuals who subsequently developed follicular lymphoma (FL; incident FL) as compared with healthy individuals who did not (controls) in EPIC (European Prospective Investigation Into Cancer and Nutrition) discovery and NSHDS (Northern Sweden Health and Disease Study) validation cohorts. Blood samples from clinical series of patients with established FL with low tumor burden are also depicted for comparison (dashed blue line indicates median). Negative threshold value defined as half detection limit of our quantitative polymerase chain reaction assay (eg, 1×10^{-6}). Receiver operating characteristic (ROC) curves for incident FL versus controls in (B) discovery and (C) combined discovery and validation cohorts. ROC curve applied to t(14;18)-positive samples only. Gold diagonal line indicates performance of random predictor. Area under curve (AUC), sensitivity, and specificity are shown at each threshold value. NPV, negative predictive value; PPV, positive predictive value.

t(14;18) Frequency in Blood Is a Predictive Marker of FL Progression

To evaluate the potential of t(14;18) frequency to discriminate individuals who will develop FL, we next performed a receiver operating curve analysis. Area under curve was computed to be 0.744 (95% CI, 0.654 to 0.834) when carried out on t(14;18)-positive samples, indicating that t(14;18) frequency constitutes an excellent predictive marker among individuals who will develop FL (Fig 1B). We estimated that an optimal cutoff frequency value (ie, one providing optimal tradeoff between sensitivity and specificity) for predicting FL development was 10^{-4}. Using unconditional logistic regression, we evaluated an approximately 15-fold higher risk of subsequent FL

from 131.211.105.239

associated with a frequency greater than 10^{-4} (odds ratio [OR], 15.52; 95% CI, 5.12 to 47.11; $P < .001$), when adjusting for age, sex, time since blood draw, and country of residence.

An independent validation cohort (NHSDS) was selected using the same criteria as for the discovery data set and consisted of 65 incident FL participant cases and 128 matched controls (Appendix Fig A1, online only). Findings from the discovery cohort were fully validated (Table 1; Fig 1A; Appendix Figs A2A and A2B, online only). When both discovery and validation cohorts were combined, the area under the curve was 0.764 (95% CI, 0.700 to 0.827; Fig 1C). Meta-analysis of the discovery and validation phases found no heterogeneity in the effect estimates between the two cohorts at the thresholds where an OR could be computed in both cohorts: 5×10^{-6}, 1×10^{-5}, and 5×10^{-5} ($P > .17$; Appendix Figs A2A and A2B, online only). Overall, the combined analysis reported a 23-fold increase in risk associated with this 10^{-4} threshold (OR, 23.17; 95% CI, 9.98 to 67.31; $P < .001$), confirming the cutoff of 10^{-4} as a highly reliable discriminator of FL development.

Alternative BCL2 *Translocations*

Our results show that in 20% of individuals who will develop FL later, t(14;18) frequency in blood greater than 10^{-4} precedes disease development. However, a notable proportion of EPIC participants who developed FL during follow-up carried lower and/or undetectable t(14;18) frequencies (Fig 1A). To first delineate the possibility of false negatives, we refined our screen. Although 70% of FLs harbor translocations in *BCL2* major breakpoint region (MBR), 15% to 20% have alternative *BCL2* breakpoints or lack t(14;18) altogether.[18,19] We thus designed a nested PCR assay to screen the discovery cohort for three additional breakpoint clusters (minor cluster region [mcr], 3′MBR, intermediate cluster region; Appendix Fig A3, online only), each involved in fewer than 10% of cases of FL. Among 30 screened classical t(14;18)-negative EPIC participants who subsequently developed FL, four were identified with alternative *BCL2-IGH* translocations (mcr/J_H, n = 2; 3′MBR/J_H, n = 2). Translocation frequencies were distributed similarly among those with MBR/J_H junctions (from 2×10^{-6} to 2×10^{-4}), indicating that samples with alternative breakpoints did not constitute a significant contingent of false-negative participant cases.

t(14;18) Frequency and Time to Diagnosis

We next asked whether increased t(14;18) frequency might be preferentially observed in prediagnostic samples collected close to diagnosis. Elapsed time between blood draw and diagnosis in incident FL ranged from 2 months to 20 years (average, 7.6 years). Overall, median time to diagnosis did not differ according to t(14;18) positivity ($P = .538$; Appendix Fig A4A, online only). Among t(14;18)-positive participant cases, no significant correlation was observed between t(14;18) level and time to diagnosis ($P = .56$; $r_s = -0.06$; Appendix Fig A4B, online only). When conducting a stratified analysis in 5-year intervals (adjusted for matching variables), we found a significantly increased risk of FL associated with a threshold of more than 1×10^{-4} in all intervals up to 15 years (0 to 5, 5 to 10, and 10 to 15 years; Table 2). Interestingly, when time to diagnosis was stratified into 1-year intervals (up to 5 years), the t(14;18) distribution was significantly different between prediagnostic samples from incident participant cases with FL and patients with FL at diagnosis, except for samples collected close to diagnosis (< 12 months; $P = .108$; Table 3). These constituted a small contingent of incident participant cases with FL (five of 168), who might represent undiagnosed patients and not precursor state.

Table 2. Stratified Analysis of t(14;18) Frequency According to Time to Diagnosis

Time to Diagnosis (years)	No. of Participant Cases*	No. of Controls*	OR	95% CI	*P*
< 5	30	115	17.17	4.48 to 65.81	< .001
5-10	36	115	22.86	6.69 to 78.08	< .001
10-15	25	115	8.40	1.92 to 36.73	< .01
> 15	11	115	5.85	0.42 to 80.94	.187

NOTE. *P* for heterogeneity = .665.
Abbreviation: OR, odds ratio.
*t(14;18)-positive samples.

Altogether, these data indicate that high t(14;18) frequency discriminates commitment to FL development up to 15 years before diagnosis. This latency is in keeping with previous anecdotal reports suggesting that t(14;18)-positive precursors can be committed to FL development up to 9 years before diagnosis.[9,11] This also suggests that two forms of FL preclinical courses may exist: one involving a massive release of t(14;18)-positive precursors in blood, and one where trafficking is more limited, potentially constrained to homing niches in lymphoid organs. In line with this possibility, a fraction of patients with established FL with low tumor burden also displayed low t(14;18) frequency in blood (Fig 1A).

Early t(14;18)-Positive Clones Precede the Disease

To confirm that the t(14;18)-positive clones present in prediagnostic samples constituted bona fide precursors, which progressed to overt FL, three good-quality archival formalin-fixed paraffin-embedded FL biopsies from EPIC participants were recovered, and their t(14;18) breakpoints were determined and compared with those found in the prediagnostic samples. All FL and pre-FL samples were t(14;18) positive, including one mcr variant. In all cases, the same *BCL2-IGH* breakpoint was found before and after FL outcome (Fig 2), demonstrating that t(14;18)-positive clones in blood from these healthy individuals constituted bona fide precursors, which progressed to FL. Importantly, this clonal filiation was apparent not only in individuals with high frequencies; one precursor clone was detected at a low frequency (one per 500,000 cells). Time to diagnosis was variable (range, 4.9 to 10.6 years), further adding to the evidence that committed t(14;18)-positive precursors can be present and circulate in the blood several years before diagnosis.

DISCUSSION

The presence at low frequency of t(14;18)-bearing clones in the blood from healthy individuals is well documented.[5,7,20] Nevertheless, whether some (and/or which) positive individuals are prone to develop clinical FL and the timeframe of progression to disease remain to date largely circumstantial.[9-11] Considering that FL is diagnosed in approximately one in 25,000 individuals annually[18,19] and likely develops over several decades, large prospective cohort studies provide a

From 31.211.105.239

Table 3. t(14;18) Prevalence and Frequency Among Prediagnostic FL samples Diagnosed Between 1 and 5 Years After Blood Draw Compared With Blood at Diagnosis From Patients With Established FL

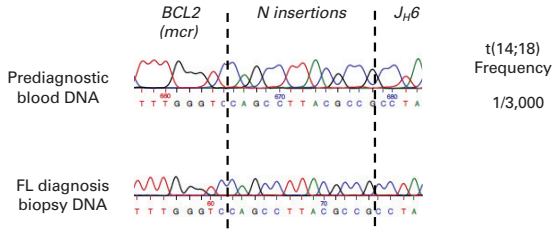
Time to Diagnosis (years)	Prevalence		P*	Median t(14;18) Frequency ($\times 10^{-5}$)	P†
	No.	%			
Series of French patients with FL	17 of 25	68		6.5	
EPIC and NSHDS cohorts					
< 1	5 of 8	63	.774	4.5	.108
1-2	13 of 16	81	.35	0.2	.010
2-3	18 of 27	67	.918	0.2	.002
3-4	23 of 39	59	.467	0.2	.003
4-5	30 of 49	61	.567	3.1	.007

Abbreviations: EPIC, European Prospective Investigation Into Cancer and Nutrition; FL, follicular lymphoma; NSHDS, Northern Sweden Health and Disease Study.
*χ^2 test.
†Mann-Whitney *U* test.

unique opportunity to screen prediagnostic blood samples and investigate the association of a predictive molecular marker with disease onset.[21,22] Here we quantified t(14;18) frequency in archived blood samples from 511 healthy adults observed in the EPIC and NSHDS cohorts, 165 of whom subsequently developed FL up to 20 years later. Our data unambiguously demonstrate that in approximately 20% of healthy individuals who went on to develop FL, commitment to FL progression is significantly associated with elevated t(14;18) frequency. We determined that individuals with t(14;18) frequency reaching one in every 10,000 blood cells had a 23-fold greater risk of progression to FL. Importantly, the predictive value of high t(14;18) in blood is effective both close to and long before disease onset (1 to 15 years), the risk estimate remaining strongly significant regardless of time to disease. Although we cannot exclude the possibility that some of the incident participant cases with FL (especially those from whom samples were collected < 12 months before diagnosis) might have been asymptomatic undiagnosed patients at the time of blood draw, it seems unlikely that they represent an important proportion of our healthy participants. To our knowledge, this constitutes the first predictive biomarker of FL. Together with chronic lymphocytic leukemia/monoclonal B-cell lymphocytosis (MBL) and multiple myeloma (MM)/monoclonal gammopathy of undetermined significance (MGUS), FL/high t(14;18) frequency before FL diagnosis represents the third example of a B-cell premalignant state that will progress into overt malignancy.[21-23]

It is well acknowledged that the use of a biomarker to predict the occurrence of a low-incidence disease in the general population is difficult to achieve because of a demanding tradeoff between sensitivity and specificity.[24,25] This is the case for MBL/MGUS, the systematic screening of which is currently not recommended outside a research study. Similarly, although the high-frequency t(14;18) biomarker confers a strong risk for the development of clinically diagnosed FL, alone it does not provide a basis for early prediction at the individual level, because at the 10^{-4} cutoff, which achieves reasonable sensitivity, specificity is only 96.5%.

Recognition of the MBL/MGUS precursor entities has nevertheless been of great value, notably in allowing the identification of high-risk patients[26,27] who might benefit from active surveillance and/or early therapeutic intervention.[28,29] Fine MBL characterization has allowed stratification into low-count MBL and clinical MBL,[30] the latter seeming to overlap with chronic lymphocytic leukemia at both genomic and clinical levels.[31] A score combining elevated serum-free light chain, M spike, and gene expression profile identified a subset of patients with high risk of progression to clinical MM requiring therapy.[32] In this context, recent studies have begun to explore early initiation of therapy in asymptomatic MM.[33,34] For FL, one of the



Case	Origin	*BCL2* Break*	N Insertions	J_H Segment*	t(14;18) Frequency
#1945/01	Prediagnostic blood	GGGGC (3056)	CTGCCCGGCGGTCCTGGA	(88758) ACTACTT - J_H4	1/500,000
	Diagnostic biopsy	GGGGC (3056)	CTGCCCGGCGGTCCTGGA	(88758) ACTACTT - J_H4	
#99-I-02520	Prediagnostic blood	CCTCCCCGG (3164)	TGCCCTC	(89762) TACTACTACTACTAC - J_H6	1/5,000
	Diagnostic biopsy	CCTCCCCGG (3164)	TGCCCTC	(89762) TACTACTACTACTAC - J_H6	

Fig 2. Clonal identity between *BCL2*/J_H junctions in paired prediagnostic blood samples and follicular lymphoma (FL) biopsies at diagnosis (case no. 03-12089). Blood t(14;18) frequency in prediagnostic samples indicated. (*) Breakpoint positions numbered according to Genbank M14745 (*BCL2*) and X97051 (*IGH*).

from 131.211.105.239

major challenges in the future will be to develop similar individualized risk profiles. However, this requires the identification of a complementing set of molecular predictors in pre-FL cells. The present identification of committed high t(14;18) cells allows thorough biologic characterization of such precursors and constitutes an important starting tool to address this central issue. Longitudinal studies will be instrumental in estimating the risk of clinical progression of such precursors.

Non-Hodgkin B-cell lymphomas, including FL, have shown considerable incidence increases during the past 30 years.[35] This progression is to date largely unexplained, and the identification of the underlying risk factors remains a major public health issue. Considering the strength of the association in our study (OR $>$ 20, predictive years before diagnosis, and validation in independent cohort), the high-frequency t(14;18) blood-based biomarker could be of great value in identifying groups of at-risk individuals and associated risk factors in prospective large-scale epidemiologic studies.[36] In FL, trends include immune dysfunction and exposure to environmental factors, such as pesticides.[37,38] Recently, common genetic variants at specific loci (including HLA) identified from genome-wide association studies emerged as additional risk factors for FL development, which might be associated with early premalignant stages.[39] It is expected that the combination of risk factors issued from genomic, environmental, and clinical investigations will be required to further define healthy subpopulations at risk for FL development.

One unexpected observation from our study was that a fraction of both patients with FL and incident FL did not display elevated t(14;18) levels in blood. The basis of the difference in blood tumor cell release remains to be clarified, but a more limited trafficking might be linked to increased expression of adhesion receptors and/or microenvironment-promoted chemotaxis.[40,41] We investigated t(14;18) frequency in BM in our series of established FL cases with low tumor burden, and it closely reflected frequency in blood ($P < .001$; Appendix Fig A5, online only), further arguing for a link between increased frequency in blood and the propensity to disseminate. Such concordance between BM and blood frequency, together with evidence of bidirectional trafficking, was also found in a series of otherwise healthy individuals undergoing thoracic surgery (Sungalee et al, submitted for publication), suggesting that distinctive trafficking properties and ensuing dissemination might be established early on during pre-FL development. The microenvironment has an important prognostic significance in FL.[42,43] Identifying whether and how such distinctive trafficking properties of t(14;18)-positive clones are linked to differential interactions with the microenvironment might be of particular interest for tailored therapy, especially with the appearance of new and promising drugs targeting tumor microenvironment components.[44,45]

Early committed precursors likely represent the FL cancer progenitor cells from which relapses reemerge,[46,47] and the identification of candidate genes or oncogenes involved in early steps of FL development is currently an area of intense genomic research.[48-52] Recently, the report of synchronous FL development occurring from the same t(14;18)-positive precursor 9 years after BM transplantation in both the donor and recipient[11] demonstrated that commitment to malignant transformation can occur nearly a decade before disease onset. In line with this, we show here that progression to FL systematically occurred from clonally related t(14;18)-positive bona fide lymphoma precursors up to 10 years before diagnosis. The identification of committed t(14;18) cells at high frequency in participants developing FL provides a unique opportunity to unravel key genomic alterations paving the progression from pre-FL to FL through genome-wide approaches.[53,54] This might contribute to the design of noncytotoxic targeted therapy in the near future, which may be beneficial in well-defined, less-advanced/refractory forms of the disease.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Sandrine Roulland, Paolo Vineis, Bertrand Nadel
Provision of study materials or patients: Philippe Solal-Celigny, Philippe Colombat, Carlotta Sacerdote, Gilles Salles
Collection and assembly of data: Sandrine Roulland, Ester Morgado, Stephanie Sungalee, Philippe Solal-Celigny, Philippe Colombat, Nathalie Jouve, Domenico Palli, Valeria Pala, Rosario Tumino, Salvatore Panico, Carlotta Sacerdote, Jose Ramon Quiros, Carlos A. González, Maria-Jose Sanchez, Mirren Dorronsoro, Carmen Navarro, Aurelio Barricarte, Anne Tjonneland, Anja Olsen, Kim Overvad, Federico Canzian, Rudolf Kaaks, Heiner Boeing, Dagmar Drogan, Alexandra Nieters, Francoise Clavel-Chapelon, Antonia Trichopoulou, Dimitrios Trichopoulos, Pagona Lagiou, H. Bas Bueno-de-Mesquita, Petra H.M. Peeters, Roel Vermeulen, Göran Hallmans, Beatrice Melin, Signe Borgquist, joyce Carlson, Eiliv Lund, Elisabete Weiderpass, Kay-Tee Khaw, Nick Wareham, Timothy J. Key, Ruth C. Travis, Pietro Ferrari, Isabelle Romieu, Elio Riboli
Data analysis and interpretation: Sandrine Roulland, Rachel S. Kelly, Ester Morgado, Gilles Salles, Paolo Vineis, Bertrand Nadel
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Montoto S, Fitzgibbon J: Transformation of indolent B-cell lymphomas. J Clin Oncol 29:1827-1834, 2011
2. Relander T, Johnson NA, Farinha P, et al: Prognostic factors in follicular lymphoma. J Clin Oncol 28:2902-2913, 2010
3. Shaffer AL 3rd, Young RM, Staudt LM: Pathogenesis of human B cell lymphomas. Annu Rev Immunol 30:565-610, 2012
4. Kridel R, Sehn LH, Gascoyne RD: Pathogenesis of follicular lymphoma. J Clin Invest 122: 3424-3431, 2012
5. Limpens J, Stad R, Vos C, et al: Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. Blood 85:2528-2536, 1995
6. Dölken G, Illerhaus G, Hirt C, et al: BCL-2/JH rearrangements in circulating B cells of healthy blood donors and patients with nonmalignant diseases. J Clin Oncol 14:1333-1344, 1996
7. Schüler F, Dölken L, Hirt C, et al: Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals. Int J Cancer 124:958-963, 2009
8. Roulland S, Faroudi M, Mamessier E, et al: Early steps of follicular lymphoma pathogenesis. Adv Immunol 111:1-46, 2011
9. Bretherick KL, Bu R, Gascoyne RD, et al: Elevated circulating t(14;18) translocation levels prior to diagnosis of follicular lymphoma. Blood 116:6146-6147, 2010
10. Hart J, Turner AR, Larratt L, et al: Transmission of a follicular lymphoma by allogeneic bone marrow transplantation: Evidence to support the existence of lymphoma progenitor cells. Br J Haematol 136:166-167, 2007
11. Weigert O, Kopp N, Lane AA, et al: Molecular ontogeny of donor-derived follicular lymphomas

from 131.211.105.239

occurring after hematopoietic cell transplantation. Cancer Discov 2:47-55, 2012

12. Roulland S, Navarro JM, Grenot P, et al: Follicular lymphoma-like B cells in healthy individuals: A novel intermediate step in early lymphomagenesis. J Exp Med 203:2425-2431, 2006

13. Hirt C, Dölken G, Janz S, et al: Distribution of t(14;18)-positive, putative lymphoma precursor cells among B-cell subsets in healthy individuals. Br J Haematol 138:349-353, 2007

14. Colombat P, Salles G, Brousse N, et al: Rituximab (anti-CD20 monoclonal antibody) as single first-line therapy for patients with follicular lymphoma with a low tumor burden: Clinical and molecular evaluation. Blood 97:101-106, 2001

15. Brice P, Bastion Y, Lepage E, et al: Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: A randomized study from the Groupe d'Etude des Lymphomes Folliculaires—Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol 15:1110-1117, 1997

16. Søreide K: Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. J Clin Pathol 62:1-5, 2009

17. Riboli E, Hunt KJ, Slimani N, et al: European Prospective Investigation Into Cancer and Nutrition (EPIC): Study populations and data collection. Public Health Nutr 5:1113-1124, 2002

18. Buchonnet G, Jardin F, Jean N, et al: Distribution of BCL2 breakpoints in follicular lymphoma and correlation with clinical features: Specific subtypes or same disease? Leukemia 16:1852-1856, 2002

19. Leich E, Salaverria I, Bea S, et al: Follicular lymphomas with and without translocation t(14;18) differ in gene expression profiles and genetic alterations. Blood 114:826-834, 2009

20. Roulland S, Lebailly P, Lecluse Y, et al: Long-term clonal persistence and evolution of t(14;18)-bearing B cells in healthy individuals. Leukemia 20:158-162, 2006

21. Landgren O, Albitar M, Ma W, et al: B-cell clones as early markers for chronic lymphocytic leukemia. N Engl J Med 360:659-667, 2009

22. Weiss BM, Abadie J, Verma P, et al: A monoclonal gammopathy precedes multiple myeloma in most patients. Blood 113:5418-5422, 2009

23. Landgren O, Kyle RA, Pfeiffer RM, et al: Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: A prospective study. Blood 113:5412-5417, 2009

24. Ware JH: The limitations of risk factors as prognostic tools. N Engl J Med 355:2615-2617, 2006

25. Elmore JG, Fletcher SW: The risk of cancer risk prediction: "What is my risk of getting breast cancer"? J Natl Cancer Inst 98:1673-1675, 2006

26. Rawstron AC, Bennett FL, O'Connor SJ, et al: Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. N Engl J Med 359:575-583, 2008

27. Shanafelt TD, Kay NE, Rabe KG, et al: Brief report: Natural history of individuals with clinically recognized monoclonal B-cell lymphocytosis compared with patients with Rai 0 chronic lymphocytic leukemia. J Clin Oncol 27:3959-3963, 2009

28. Shanafelt TD, Ghia P, Lanasa MC, et al: Monoclonal B-cell lymphocytosis (MBL): Biology, natural history and clinical management. Leukemia 24:512-520, 2010

29. Korde N, Kristinsson SY, Landgren O: Monoclonal gammopathy of undetermined significance (MGUS) and smoldering multiple myeloma (SMM): Novel biological insights and development of early treatment strategies. Blood 117:5573-5581, 2011

30. Fazi C, Scarfò L, Pecciarini L, et al: General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. Blood 118:6618-6625, 2011

31. Vardi A, Dagklis A, Scarfò L, et al: Immunogenetics shows that not all MBL are equal: The larger the clone, the more similar to CLL. Blood 121:4521-4528, 2013

32. Dhodapkar MV, Sexton R, Waheed S, et al: Clinical, genomic and imaging predictors of myeloma progression from asymptomatic monoclonal gammopathies (SWOG S0120). Blood 123:78-85, 2014

33. Mateos MV, Hernández MT, Giraldo P, et al: Lenalidomide plus dexamethasone for high-risk smoldering multiple myeloma. N Engl J Med 369: 438-447, 2013

34. Barlogie B, van Rhee F, Shaughnessy JD Jr, et al: Seven-year median time to progression with thalidomide for smoldering myeloma: Partial response identifies subset requiring earlier salvage therapy for symptomatic disease. Blood 112:3122-3125, 2008

35. Fisher SG, Fisher RI: The epidemiology of non-Hodgkin's lymphoma. Oncogene 23:6524-6534, 2004

36. Landgren O, Goedert JJ, Rabkin CS, et al: Circulating serum free light chains as predictive markers of AIDS-related lymphoma. J Clin Oncol 28:773-779, 2010

37. Chiu BC, Dave BJ, Blair A, et al: Agricultural pesticide use and risk of t(14;18)-defined subtypes of non-Hodgkin lymphoma. Blood 108:1363-1369, 2006

38. Giannelli F, Moscarella S, Giannini C, et al: Effect of antiviral treatment in patients with chronic HCV infection and t(14;18) translocation. Blood 102: 1196-1201, 2003

39. Conde L, Halperin E, Akers NK, et al: Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet 42:661-664, 2010

40. Mandigers CM, Meijerink JP, Mensink EJ, et al: Lack of correlation between numbers of circulating t(14;18)-positive cells and response to first-line treatment in follicular lymphoma. Blood 98:940-944, 2001

41. Drillenburg P, Pals ST: Cell adhesion receptors in lymphoma dissemination. Blood 95:1900-1910, 2000

42. Kiaii S, Clear AJ, Ramsay AG, et al: Follicular lymphoma cells induce changes in T-cell gene expression and function: Potential impact on survival and risk of transformation. J Clin Oncol 31:2654-2661, 2013

43. Dave SS, Wright G, Tan B, et al: Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med 351:2159-2169, 2004

44. Schuster SJ, Neelapu SS, Santos CF, et al: Idiotype vaccination as consolidation therapy: Time for integration into standard of care for follicular lymphoma? J Clin Oncol 29:4845-4846, 2011

45. Fowler N, Neelapu S, Hagemeister F, et al: Lenalidomide and rituximab for untreated indolent lymphoma: Final results of a phase II study. Blood, 2012 (abstr 901)

46. Carlotti E, Wrench D, Matthews J, et al: Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. Blood 113:3553-3557, 2009

47. Ruminy P, Jardin F, Picquenot JM, et al: S(mu) mutation patterns suggest different progression pathways in follicular lymphoma: Early direct or late from FL progenitor cells. Blood 112:1951-1959, 2008

48. Green MR, Gentles AJ, Nair RV, et al: Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. Blood 121:1604-1611, 2013

49. Pasqualucci L, Dominguez-Sola D, Chiarenza A, et al: Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 471:189-195, 2011

50. Morin RD, Johnson NA, Severson TM, et al: Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 42:181-185, 2010

51. Morin RD, Mendez-Lago M, Mungall AJ, et al: Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 476:298-303, 2011

52. Bödör C, Grossmann V, Popov N, et al: EZH2 mutations are frequent and represent an early event in follicular lymphoma. Blood 122:3165-3168, 2013

53. Landau DA, Carter SL, Stojanov P, et al: Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 152:714-726, 2013

54. Harismendy O, Schwab RB, Bao L, et al: Detection of low prevalence somatic mutations in solid tumors with ultra-deep targeted sequencing. Genome Biol 12:R124, 2011

Affiliations

Sandrine Roulland, Ester Morgado, Stéphanie Sungalee, Nathalie Jouve, and Bertrand Nadel, Aix-Marseille Université, Institut National de la Santé et de la Recherche Médicale (INSERM) U1104, and Centre National de la Recherche Scientifique (CNRS) Unités Mixtes de Recherche (UMR) 7280, Marseille; Philippe Solal-Celigny, Jean Bernard Center, Le Mans; Philippe Colombat, Bretonneau University Hospital, Tours; Françoise Clavel-Chapelon, INSERM U1018 Centre de Recherche en Epidémiologie et Santé des Populations, Villejuif; Pietro Ferrari and Isabelle Romieu, International Agency for Research on Cancer, Lyon; Gilles Salles, Hospices Civils de Lyon, Université de Lyon, UMR CNRS 5239, Pierre Bénite, France; Rachel S. Kelly, Petra H.M. Peeters, Roel Vermeulen, Elio Riboli, and Paolo Vineis, School of Public Health, Imperial College London, London; Kay-Tee Khaw, University of Cambridge; Nick Wareham, Institute of Metabolic Science, Cambridge; Timothy J. Key and Ruth C. Travis, University of Oxford, Oxford, United Kingdom; Domenico Palli, Istituto per lo Studio e la Prevenzione Oncologica, Florence; Valeria

Pala, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico, Istituto Nazionale Tumori, Milan; Rosario Tumino, "Civile–M.P. Arezzo" Hospital, Ragusa; Salvatore Panico, Federico II University, Naples; Carlotta Sacerdote, Centro di Riferimento per l'Epidemiologia e la Prevenzione Oncologica–Piemonte, Torino, Italy; José R. Quirós, Public Health and Health Planning Directorate, Asturias; Carlos A. Gonzáles, Catalan Institute of Oncology, Barcelona; Maria-José Sánchez, Andalusian School of Public Health and Biomedical Research Centre Network for Epidemiology and Public Health (CIBERESP), Granada; Miren Dorronsoro, Basque Regional Health Department and CIBERESP Biodonostia, San Sebastian; Carmen Navarro, Murcia Regional Health Council, Universidad de Murcia, and CIBERESP, Murcia; Aurelio Barricarte, Navarre Public Health Institute and CIBERESP, Pamplona, Spain; Anne Tjønneland and Anja Olsen, Danish Cancer Society Research Center, Copenhagen; Kim Overvad, Aarhus University, Aarhus, Denmark; Federico Canzian and Rudolf Kaaks, German Cancer Research Centre, Heidelberg; Heiner Boeing and Dagmar Drogan, German Institute of Human Nutrition, Potsdam-Rehbrücke; Alexandra Nieters, University Medical Center Freiburg, Freiburg, Germany; Antonia Trichopoulou, Dimitrios Trichopoulos, and Pagona Lagiou, University of Athens Medical School; Antonia Trichopoulou and Dimitrios Trichopoulos, Hellenic Health Foundation; Pagona Lagiou, Academy of Athens, Athens, Greece; Dimitrios Trichopoulos and Pagona Lagiou, Harvard School of Public Health, Boston, MA; H. Bas Bueno-de-Mesquita, National Institute for Public Health and the Environment, Bilthoven; Bas Bueno-de-Mesquita, Petra H.M. Peeters, and Roel Vermeulen, University Medical Centre Utrecht, Utrecht, the Netherlands; Göran Hallmans and Beatrice Melin, Umeå University, Umeå; Signe Borgquist and Joyce Carlson, Lund University, Lund; Elisabete Weiderpass, Karolinska Institutet, Stockholm, Sweden; Eiliv Lund and Elisabete Weiderpass, Institute of Community Medicine, University of Tromsø, Tromsø; Elisabete Weiderpass, Cancer Registry of Norway, Oslo, Norway; and Elisabete Weiderpass, Samfundet Folkhälsan, Helsinki, Finland.

Support

B.N. supported by grants from Medimmune, Fondation de France (Grant No. 2010-15034), Association pour la Recherche sur le Cancer (Grant No. ARC#20101201415), and France Lymphome Espoir Association and by institutional grants from Institut National de la Santé et de la Recherche Médicale (INSERM) and Centre National de la Recherche Scientifique. S.S. supported by fellowship from French Ministry of Research and Fondation pour la Recherche Medicale. EPIC study supported by Europe Against Cancer Program of European Commission; Deutsche Krebshilfe, Deutsches Krebsforschungszentrum, and German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund of Spanish Ministry of Health; Spanish regional governments of Andalusia, Asturias, Basque, Murcia, and Navarra; Catalan Institute of Oncology; Instituto de Salud Carlos III of Spanish Ministry of Health (Red Temática de Investigación Cooperativa en Cáncer Grant No. DR06/0020); Cancer Research United Kingdom; Medical Research Council United Kingdom; Hellenic Health Foundation; Italian Association for Research on Cancer; Italian National Research Council, Fondazione Istituto Banco Napoli; Compagnia di San Paolo; Dutch Ministry of Public Health, Welfare and Sports, Netherlands Cancer Registry, LK Research Fund, Dutch Prevention Funds, Dutch Zorg Onderzoek Nederland, World Cancer Research Fund, and Statistics Netherlands; Swedish Cancer Society; Swedish Scientific Council; Regional Government of Västerbotten, Sweden; Norwegian Cancer Society; Research Council of Norway; French League Against Cancer; INSERM; Mutuelle Générale de l'Education Nationale; 3M (France); Gustave Roussy Institute; and General Councils of France.

From 131.211.105.239

Acknowledgment

We thank EPIC (European Prospective Investigation Into Cancer and Nutrition) and NSHDS (Northern Sweden Health and Disease Study) participants for their involvement in this study.

Appendix

EPIC Cohort

EPIC (European Prospective Investigation Into Cancer and Nutrition) is a multicenter prospective cohort established to investigate the role of biologic, dietary, lifestyle, and environmental factors in the etiology of cancer and other chronic diseases. Approximately 520,000 healthy men and women, age 35 to 70 years, were recruited between 1992 and 2000 in 23 centers located in 10 European countries, including Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom. At enrollment, participants were asked to complete a self-administered questionnaire to collect demographic, medical, and lifestyle information, including a self-administered food frequency questionnaire. Participants agreed to have their health status monitored for the rest of their lives, and standardized procedures were used to identify incident cancers based on automated linkages to cancer and mortality registries, municipal population offices, and hospital discharge systems. Diagnosis, tumor site classification, and morphology of each participant case were based on International Classification of Diseases for Oncology, second revision (ICD-O-2), which was reclassified according to the recently published WHO Classification of Tumors of the Hematopoietic and Lymphoid Tissues (ICD-O-3). Inclusion was strictly limited to patients diagnosed with follicular lymphoma (FL); those with transformed FL or de novo diffuse large B-cell lymphoma with or without t(14;18) were excluded. For the current analysis, participants were observed from study recruitment until the date of last known contact, date of diagnosis, or date of death (report complete until December 2008). None of the EPIC controls had been diagnosed with clinically recognized FL or any other lymphoid malignancies at the last follow-up.

NSHDS Cohort

The NSHDS (Northern Sweden Health and Disease Study) is a large prospective population-based study that comprises three cohorts: the Västerbotten Intervention Project, the Northern Sweden WHO Monitoring of Trends and Cardiovascular Disease Study, and the Local Mammography Screening Project (MSP). A total of 95,000 healthy individuals age 40 to 60 years were invited for inclusion in the project between 1990 and 2006. At initial recruitment, participants were asked to complete a self-administered questionnaire to collect demographic, medical, and lifestyle information and a separate self-administered food frequency questionnaire. Informed consent was obtained from all participants, and a medical examination was conducted, during which a blood sample was taken. Patients with FL (ICD-O-3 classification) diagnosed between 1994 and 2013 were identified by linking the Northern Sweden Health and Disease Cohort with the National Cancer Registry. For each participant case, two controls were selected from among cohort members alive and matched by study center, age ($\pm$ 1 year), sex, and blood collection date ($\pm$ 45 days). Controls were alive and free from cancer at the time of FL diagnosis of the matched participant case. According to the eligibility criteria used for the EPIC cohort, a total of 195 samples were included (68 incident FL participant cases and 128 matched controls), of which three were removed for nonamplifiable DNA (one incident FL participant case and two controls). The study protocol was approved by the Umeå University Research Ethics Committee (Umeå, Sweden). Written informed consent was obtained from all NSHDS participants at enrollment.

DNA Isolation

DNA samples from EPIC participants were extracted from buffy coats using the QIAsymphony DNA Midi Kit or the QIAamp DNA formalin-fixed paraffin-embedded (FFPE) tissue kit for FFPE tumor biopsies (Qiagen, Crawley, United Kingdom). DNA quantity and quality were evaluated using two methods: NanoDrop spectrophotometer (ultraviolet absorbance; Wilmington, DE) and the lab-on-a-chip technique (Agilent 2100 Bioanalyzer; Santa Clara, CA). Quality of DNA extracted from FFPE biopsies was further evaluated using a gel-based multiplex polymerase chain reaction (PCR) assay allowing the coamplification of four PCR fragments ranging from 100 to 400 bp and determining the maximum size of amplifiable DNA fragments. Samples giving rise to four amplicons had enough high-quality DNA for molecular analysis, whereas samples with low-size amplicons or no amplification were not analyzed.

Quantitative Real-Time PCR Amplification of BCL2 Major Breakpoint Region/J_H Translocations

Quantitative real-time PCR (Q-PCR) was conducted using an ABIPRISM 7500 system and Sequence Detection software (Applied Biosystems, Carlsbad, CA). Q-PCR for the *GAPDH* endogenous reference gene was systematically performed as a test for amplifiable DNA and to determine the number of cells analyzed in each single assay. Each PCR reaction mixture contained the forward and reverse primers at a concentration of 300 nmol/L, a fluorescent probe at a concentration of 200 nmol/L, the standard TaqMan Universal PCR Master Mix (Applied Biosystems), and 500-ng DNA in each PCR replicate. After a 2-minute incubation at 50°C to allow for cleavage by uracil-DNA glycosylase and initial step at 95°C for 10 minutes, the PCR program was followed by 50 cycles of 95°C for 15 seconds and 60°C for 1 minute of combined annealing and extension. Plasmidic standard curves with initial copy number going from 10^6 to 10^1 were established for both the t(14;18) major breakpoint region (MBR) –specific PCR and the *GAPDH*-specific PCR and run in parallel and in duplicate together with the samples. Genomic DNA from buffy-coat samples was tested in triplicate for the target gene and with two replicates for the reference gene. The absolute t(14;18) frequency for each individual was extrapolated from the standard curve and normalized on total

number of cells in the assay. The frequency of t(14;18)-negative samples was arbitrary, defined as half the detection threshold of our Q-PCR assay (eg, 1×10^{-6}), as previously performed. Both primers and probes are listed in Appendix Table A1 (online only).

t(14;18) Fluctuation PCR and DNA Sequencing

A sensitive two-step fluctuation PCR (F-PCR) was performed, as previously reported,[13] in Q-PCR–negative prediagnostic FL samples to determine the presence of alternative *BCL2-IGH* translocations within the intermediate cluster region, minor cluster region, and 3′MBR breakpoint clusters where heterogeneous amplicon PCR lengths (> 250 bp) are not compatible with Q-PCR assay (Appendix Fig A3, online only). Briefly, for each individual, the F-PCR assay consisted of 10 to 15 PCR replicates performed in parallel using 0.1 to 0.2 μg of DNA per reaction. In the fluctuation range, at most, one target molecule is present per PCR replicate, and if so, it will give rise to a detectable amplicon. The frequency of the event can then be calculated from the number of positive *BCL2*/J_H amplicons using Poisson's assumption, as previously described.[13] Moreover, an additional F-PCR was applied to all previously identified Q-PCR–positive pre-FL samples to generate MBR/J_H amplicons compatible with sequencing. All amplicons were purified from each independent PCR replicate and were cloned and sequenced.

Statistical Methods

Prevalence was defined as the proportion of individuals who had a detectable t(14;18)-positive blood sample (ie, those with a frequency of more than two copies per 10^6 WBCs). Frequency was defined as the number of circulating t(14;18) copies per 10^6 circulating lymphocytes. Frequency analyses were conducted for both the total population and the t(14;18)-positive subgroup.

Receiver operating characteristic curves and areas under curve were generated by plotting sensitivity (true positives) against 1-specificity (false positives). To obtain an optimal cutoff predicting FL development, we calculated sensitivity, specificity, positive predictive value, and negative predictive value at different t(14;18) thresholds.

For the stratified analysis, participant cases were categorized into four groups based on time between blood draw and diagnosis: less than 5, 5 to 10, 10 to 15, and more than 15 years. For each stratum, unconditional logistic regression was run adjusting for the matching factors using all controls as baseline. Heterogeneity between the strata was assessed using the metan meta-analysis command in STATA (STATA, College Station, TX). Significance was set at the 5% confidence level, and statistical analysis was performed using STATA software (version 11.2).

Table A1. Primers and Probes for PCR and Quantitative Real-Time PCR

Name	Target	Sequence (5′ to 3′)	Purpose
PCR and sequencing			
MBR20A*	*BCL2*	CCCTGGGCAATTCCGCATTTAATTCATGG	
MBR21A†	*BCL2*	GGCAAATGACCAGCAGATTCAAATCTATGG	
mcr-1A*	*BCL2*	TAGAGCAAGCGCCCAATAAATA	
mcr-1Aint†	*BCL2*	GCAGCTATTGTTACTCTTGCAG	
J_HCoB*	*IGH*	ACCTGAGGAGACGGTGACC	
J_HCoint B†	*IGH*	CAGGGTCCCTTGGCCCCAG	
Quantitative real-time PCR			
Q-MBR35A	*BCL2*	TGACCTTTAGAGAGTTGCTTTACGT	
Q-JHCo-B	*IGH*	ACCTGAGGAGACGGTGACC	
P-MBR2A	*BCL2*	FAM-TGTTTCAACACAGACCCA-MGB	Probe
Q-GAPDH-1A	*GAPDH*	CTGAGAACGGGAAGCTTGTCAT	
Q-GAPDH-1B	*GAPDH*	ATCCTAGTTGCCTCCCCAAAG	
P-GAPDH-1A	*GAPDH*	VIC-TCACCATCTTCCAGGAGT-MGB	Probe

Abbreviation: PCR, polymerase chain reaction.
*Primary PCR.
†Nested secondary PCR.

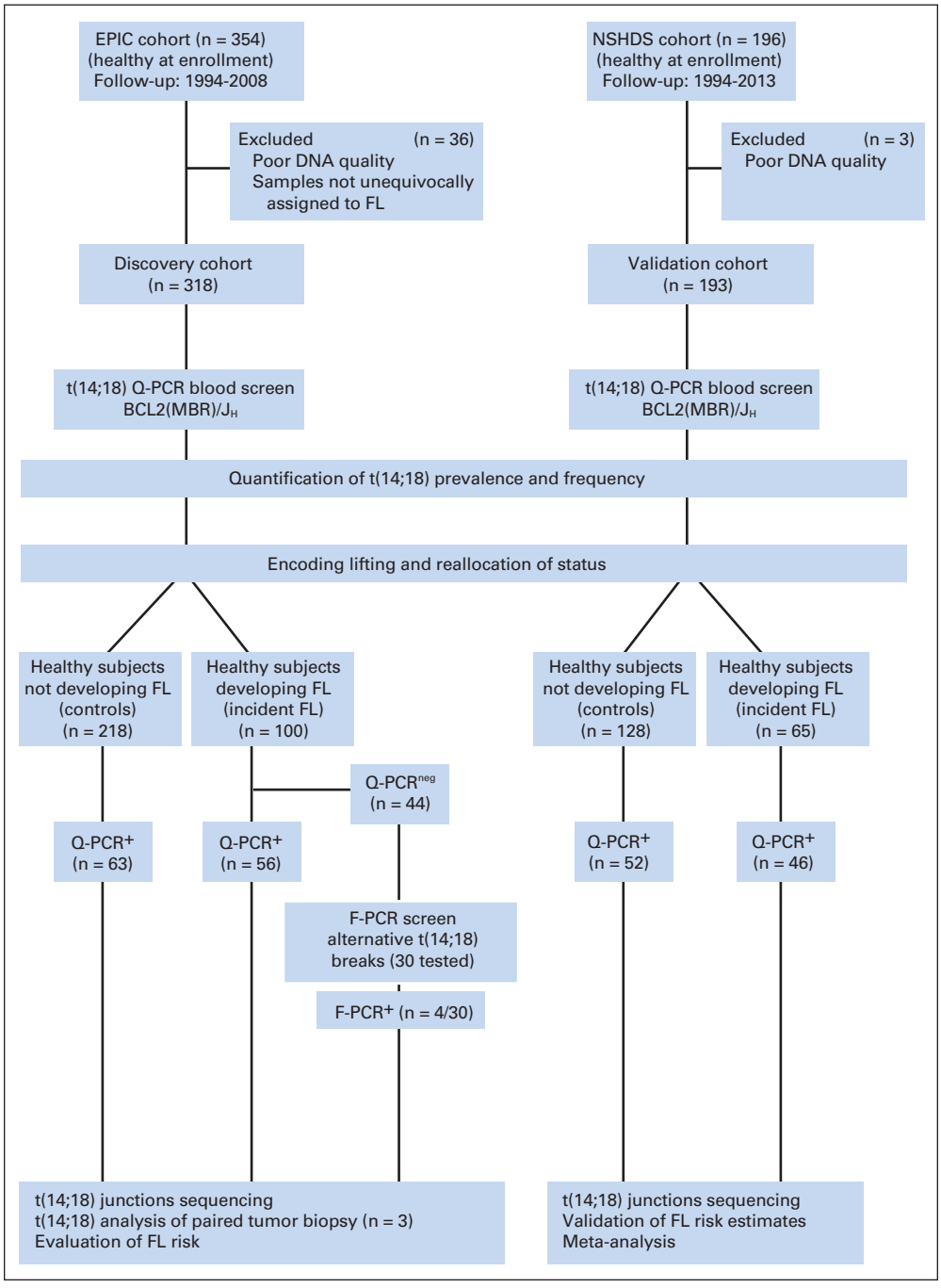
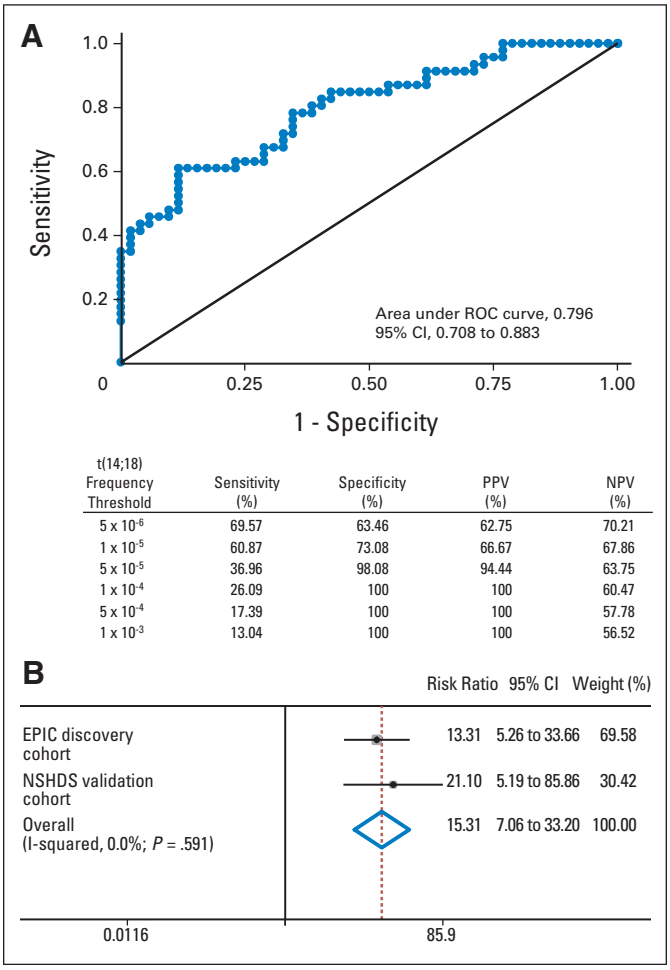


Fig A1. t(14;18) blood screening workflow in discovery (EPIC [European Prospective Investigation Into Cancer and Nutrition]) and validation (NSHDS [Northern Sweden Health and Disease Study]) cohorts. FL, follicular lymphoma; F-PCR, fluctuation polymerase chain reaction; Q-PCR, quantitative polymerase chain reaction.



t(14;18) Frequency Threshold	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
5×10^{-6}	69.57	63.46	62.75	70.21
1×10^{-5}	60.87	73.08	66.67	67.86
5×10^{-5}	36.96	98.08	94.44	63.75
1×10^{-4}	26.09	100	100	60.47
5×10^{-4}	17.39	100	100	57.78
1×10^{-3}	13.04	100	100	56.52

Fig A2. (A) Receiver operating characteristic (ROC) curve for all t(14;18)-positive incident follicular lymphoma (FL) participant cases (n = 46) versus controls (n = 52) in validation cohort (NSHDS [Northern Sweden Health and Disease Study]) and effectiveness of t(14;18) frequency for predicting FL development at different cutoff values. Cutoff value of 5×10^{-5} had highest accuracy for FL detection in validation cohort (sensitivity, 36.9%; specificity, 98.08%). (B) Forest plot showing meta-analysis of test and validation phases at highest threshold where odds ratio (OR) could be computed in both phases: 5×10^{-5} (no controls had frequency $> 1 \times 10^{-4}$ in validation cohort). ORs and 95% CIs shown for validation and discovery (EPIC [European Prospective Investigation Into Cancer and Nutrition]) cohorts. NPV, negative predictive value; PPV, positive predictive value.

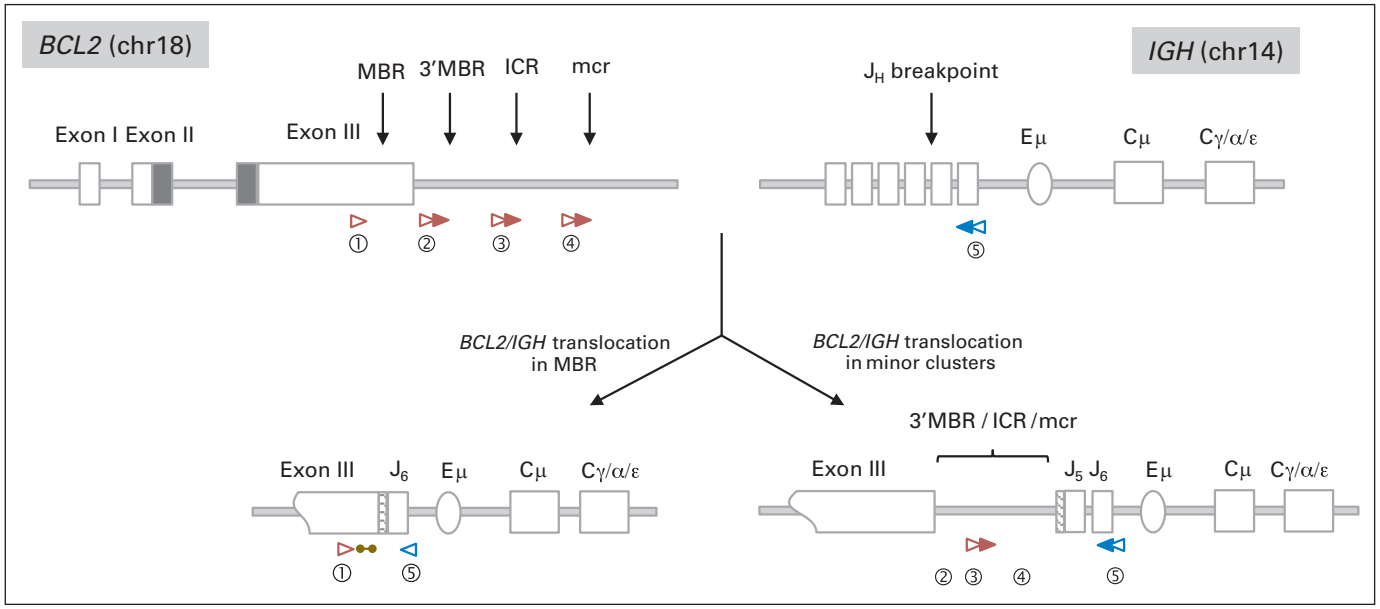


Fig A3. Design of t(14;18) polymerase chain reaction (PCR) –based assays. Schematic diagram of *BCL2* and *IGH* breakpoint regions involved in t(14;18) translocations. *IGH* breakpoint is located at one of six joining regions. *BCL2* breakpoints are indicated by vertical arrows: major breakpoint region (MBR), located within 3′ noncoding region of exon 3; minor cluster region (mcr), located 20 to 30 kb 3′ from MBR; and additional clusters 3′MBR and intermediate cluster region (ICR) in between. Configuration of *BCL2-IGH* rearrangement at MBR (bottom left), and *BCL2-IGH* rearrangement at minor breakpoint regions (bottom right). Positions of (1) MBR, (2) 3′MBR, (3) ICR, and (4) mcr primers for PCR are indicated by red arrows. (5) Consensus J_H primers for quantitative PCR (Q-PCR) and nested fluctuation PCR are indicated by blue arrows. Open arrows indicate first-round PCR; solid arrows, second-round nested PCR. Solid circles indicate *BCL2* MBR probe used for Q-PCR. Dashed boxes indicate N-nucleotide additions.

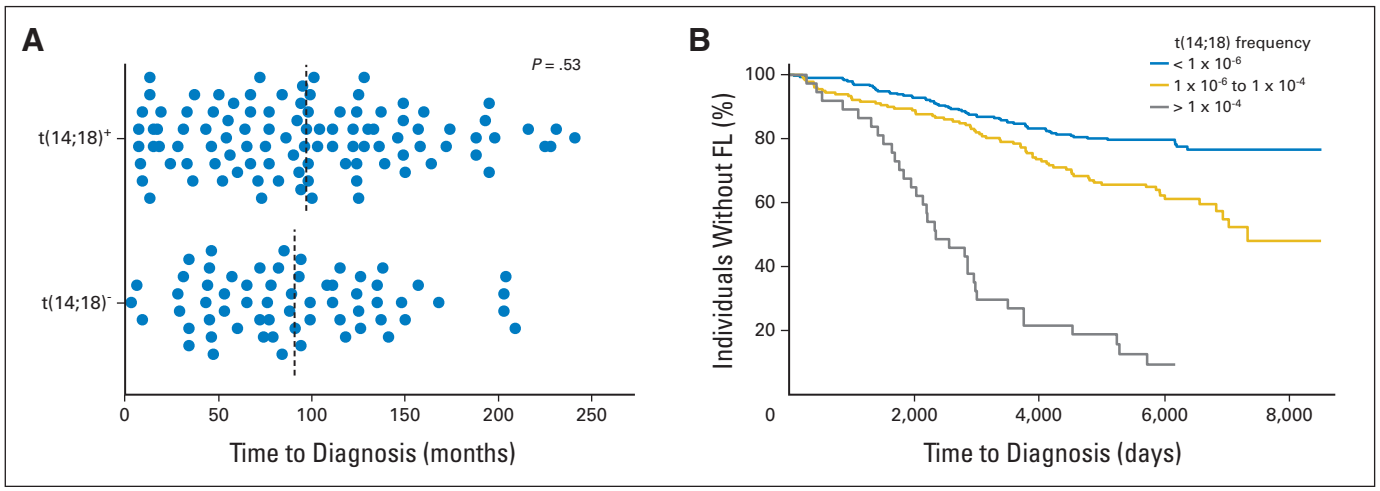


Fig A4. t(14;18) translocation and time to diagnosis. (A) Absence of difference in time to diagnosis among incident follicular lymphoma (FL) participant cases by t(14;18) prevalence status (85 *v* 94.5 months; $P = .538$) for pooled cohort (discovery plus validation cohorts; Mann-Whitney $P < .001$). (B) Kaplan-Meier curve showing estimated time to FL diagnosis according to level of circulating t(14;18)-positive cells: t(14;18) high, $> 1 \times 10^{-4}$; positive, 1×10^{-6} to 1×10^{-4}; negative, $< 1 \times 10^{-6}$. Differences between groups were assessed using log-rank test. Risk of diagnosis was significantly greater in t(14;18)-positive individuals compared with t(14;18)-negative individuals ($P < .001$) and in individuals with high t(14;18) frequency compared with t(14;18)-positive individuals ($P < .001$).

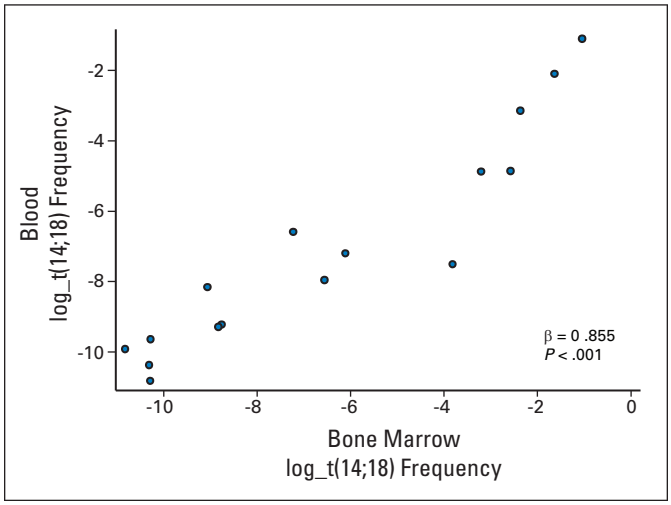


Fig A5. Scatter plot of log-transformed t(14;18) frequency in 16 patients with established follicular lymphoma with paired bone marrow/blood t(14;18)-positive samples.