Cancer Causes Control (2015) 26:1845–1855
DOI 10.1007/s10552-015-0677-2

ORIGINAL PAPER

Determinants of the t(14;18) translocation and their role in t(14;18)-positive follicular lymphoma

Rachel S. Kelly[1,2] · Sandrine Roulland[3,4,5] · Ester Morgado[3,4,5] · Stéphanie Sungalee[3,4,5] · Nathalie Jouve[3,4,5] · Rosario Tumino[6] · Vittorio Krogh[7] · Salvatore Panico[8] · Silvia Polidoro[9] · Giovanna Masala[10] · María-José Sánchez[11,12] · Maria-Dolores Chirlaque[12,13] · Núria Sala[14] · Aurelio Barricarte Gurrea[12,15] · Miren Dorronsoro[16] · Ruth C. Travis[17] · Elio Riboli[1] · Marc Gunter[1] · Neil Murphy[1] · Roel Vermeulen[18] · H. B. Bueno-de-Mesquita[19,20] · Petra H. Peeters[21] · Antonia Trichopoulou[22,23] · Dimitrios Trichopoulos[2,22,23] · Pagona Lagiou[2,23,24] · Alexandra Nieters[25] · Federico Canzian[26] · Rudolf Kaaks[26] · Heiner Boeing[27] · Elisabete Weiderpass[28,29,30,31] · Tanja Stocks[32,33] · Beatrice Melin[34] · Kim Overvad[35] · Anne Tjønneland[36] · Anja Olsen[36] · Paul Brennan[37] · Mattias Johansson[37] · Bertrand Nadel[3,4,5] · Paolo Vineis[1,9]

Received: 12 February 2015 / Accepted: 22 September 2015 / Published online: 30 September 2015

Abstract

Purpose The strong association between t(14;18) translocation and follicular lymphoma (FL) is well known. However, the determinants of this chromosomal aberration and their role in t(14;18) associated FL remain to be established.

Methods t(14;18) frequency within the B cell lymphoma 2 major breakpoint region was determined for 135 incident FL cases and 251 healthy controls as part of a nested case–control study within the European Prospective Investigation into Cancer cohort. Quantitative real-time PCR was performed in DNA extracted from blood samples taken at recruitment. The relationship between prevalence and frequency of the translocation with baseline anthropometric, lifestyle, and dietary factors in cases and controls was determined. Unconditional logistic regression was used to explore whether the risk of FL associated with these factors differed in $t(14;18)^{+}$ as compared to $t(14;18)^{-}$ cases.

Rachel S. Kelly, Sandrine Roulland, Paolo Vineis and Bertrand Nadel have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10552-015-0677-2) contains supplementary material, which is available to authorized users.

✉ Paolo Vineis
p.vineis@imperial.ac.uk

[1] MRC/PHE Centre for Environment and Health, School of Public Health, Imperial College London, St Mary's Campus Norfolk Place, London W2 1PG, UK

[2] Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

[3] Center of Immunology of Marseille-Luminy (CIML), Université de la Méditerranée, Marseille, France

[4] INSERM U631, Marseille, France

[5] CNRS UMR6102, Marseille, France

[6] Cancer Registry and Histopathology Unit, "Civic - M.P.Arezzo" Hospital, ASP Ragusa, Ragusa, Italy

[7] Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[8] Department of Clinical and Experimental Medicine, Federico II University, Naples, Italy

[9] HuGeF - Human Genetics Foundation, Turin, Italy

[10] Molecular and Nutritional Epidemiology Unit, Cancer Research and Prevention Institute – ISPO, Florence, Italy

[11] Escuela Andaluza de Salud Pública, Instituto de Investigación Biosanitaria de Granada (Granada.ibs), Granada, Spain

[12] CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

[13] Department of Epidemiology, Murcia Regional Health Authority, Murcia, Spain

[14] Unit of Nutrition and Cancer, Cancer Epidemiology Research Program and Translational Research Laboratory, Catalan Institute of Oncology (IDIBELL), Barcelona, Spain

Results Among incident FL cases, educational level (χ^2 $p = 0.021$) and height (χ^2 $p = 0.025$) were positively associated with t(14;18) prevalence, and cases with high frequencies [$t(14;18)^{HF}$] were significantly taller (*t* test *p* value = 0.006). These findings were not replicated in the control population, although there were a number of significant associations with dietary variables. Further analyses revealed that height was a significant risk factor for $t(14;18)^{+}$ FL [OR 6.31 (95 % CI 2.11, 18.9) in the tallest versus the shortest quartile], but not $t(14;18)^{-}$ cases.

Conclusions These findings suggest a potential role for lifestyle factors in the prevalence and frequency of the t(14;18) translocation. The observation that the etiology of FL may differ by t(14;18) status, particularly with regard to height, supports the subdivision of FL by translocation status.

Keywords Follicular lymphoma · Translocation · t(14;18) · Height

Introduction

Follicular lymphoma (FL) constitutes the second most common non-Hodgkin lymphoma (NHL) subtype, accounting for roughly 20 % of NHL cases in Europe [1]. FL is a lymphoma of follicle center B cells characterized by the t(14;18)(q32;q21) translocation [2]. t(14;18) juxtaposes the *B cell leukemia/lymphoma 2* (*BCL2*) proto-oncogene on chromosome 18q21 near to the immunoglobulin heavy chain (IGH) locus on chromosome 14q32. This rearrangement subjects *BCL2* to the control of the IGH locus enhancers leading to overexpression of the BCL2 protein, which inhibits apoptosis, thereby increasing cell survival and uncontrolled cell proliferation in the germinal centers [3]. t(14;18) probably constitutes the initiating lesion of FL and is believed to trigger the multistep oncogenic cascade leading to malignant disease [3]. The association between the t(14;18) chromosomal aberration and FL has long been known, with 90 % of FL cases positive for the translocation [commonly defined as a frequency >1 cell bearing the translocation per 1,000,000 circulating lymphocytes ($\geq 1 \times 10^{-6}$)]. However, the translocation is also present at very low frequencies in the blood of ~50 % of "healthy" individuals [4–6], with some studies reporting prevalence as high as 70 % in certain groups [7]. While the presence of t(14;18) clones alone thus cannot constitute a biomarker of FL risk, we have recently demonstrated that t(14;18) frequency in the blood is associated with the risk of FL development [8]. This previous study, nested within the European Prospective Investigation into Cancer (EPIC), was the first to define a high-frequency threshold [this is denoted by $t(14;18)^{HF}$ and refers to a circulating frequency $>1 \times 10^{-4}$], above which the risk of FL was increased by more than 23-fold. We were able to demonstrate that this biomarker was effective in risk prediction up to 15 years before diagnosis and to replicate these findings in an independent cohort.

However, the causes of the initiation and proliferation of these premalignant $t(14;18)^{+}$ cells remain largely unknown. The varying prevalence and frequency of $t(14;18)^{+}$ lymphocytes in healthy individuals [9] suggests a role for environmental factors [10]. Age, smoking, UV radiation, and exposure to certain pesticides and dioxins have all been implicated, but the data remain inconclusive [6, 7, 9, 11–13]. In this study, we exploited the detailed lifestyle and dietary data available from EPIC to search for

[15] Navarre Public Health Institute, Pamplona, Spain

[16] Public Health Direction and Ciberesp-Biodonostia Basque Regional Health Department, Vitoria, Spain

[17] Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

[18] Division of Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

[19] Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

[20] Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, The Netherlands

[21] Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands

[22] Hellenic Health Foundation, 13 Kaisareias Street, 115 27 Athens, Greece

[23] Bureau of Epidemiologic Research, Academy of Athens, 23 Alexandroupoleos Street, 115 27 Athens, Greece

[24] Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, 75 M. Asias Street, Goudi, 115 27 Athens, Greece

[25] Center for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Germany

[26] Genomic Epidemiology Group, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

[27] Department of Epidemiology, German Institute of Human Nutrition (DIfE), Potsdam-Rehbrüucke, Nuthetal, Germany

[28] Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[29] Cancer Registry of Norway, Oslo, Norway

[30] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[31] Department of Genetic Epidemiology, Folkhälsan Research Center, Helsinki, Finland

[32] Department of Clinical Sciences in Malmö, Lund University, Lund, Sweden

variables associated with the prevalence and frequency of the t(14;18) translocation and to investigate whether they differed in subjects who remained healthy compared to those who went on to develop FL.

Methods

Study population

A nested case–control study was conducted in a population drawn from the EPIC cohort [14]. This multicenter prospective cohort, following over half a million healthy participants from middle age onwards, was initiated in 1990 and now includes ten European countries: Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the UK. Full details can be found in [15]. In brief, a self-completed questionnaire was used to collect information on lifestyle, occupation, and environmental factors. Anthropometric variables were measured using standardized methods and quality control checks for between- and within-observer variability. Height was measured to the nearest 0.1, 0.5, or 1.0 cm dependent on study center, and weight to the nearest 0.1 kg, without shoes [16]. At baseline, diet over the previous 12 months was assessed using validated country-specific dietary questionnaires based on a common core-protocol. Questionnaires were self-administered in all countries with the exception of Greece, Spain, and the Italian center of Ragusa. In France, Greece, Germany, Italy, the Netherlands, and Spain, individual average portions were assessed quantitatively, whereas semiquantitative food-frequency questionnaires with the same standard portions were assigned to all subjects in Denmark, the UK, Norway, and Umea (Sweden). In Malmo (Sweden), a combination of these methods was employed. A standardized nutrient database was used to estimate energy, alcohol, and nutrient intake. In addition to the baseline questionnaires, standardized computer-based 24-h dietary recall measurements were collected from a random sample of participants (8 %) in order to calibrate measurements across the centers [17–19].

Ascertainment of cancer incidence is ongoing and obtained through population cancer registries, health insurance programs, or through periodic personal contact, dependent on the center. Participants for this study were selected from the EPIC cohort among those who had an archived cryopreserved buffy coat sample collected at enrollment. Eligible cases were those with incident FL (International Classification of Diseases for Oncology—3rd Edition: 9690/3, 9691/3, 9695/3, 9698/3), and controls were matched 1:2 by study center (and therefore country), age (±1 year), sex, and blood collection date (±45 days). Eligible controls were those who fulfilled the matching criteria and who were alive and cancer free at the time of diagnosis in the corresponding case. FL cases ($n = 4$) and controls ($n = 17$) with poor quality or non-amplifiable DNA were excluded.

Assessment of t(14;18) prevalence and frequency

Blood samples were extracted in the centers at recruitment, and buffy coat was stored in 0.5-ml plastic straws in liquid nitrogen containers at −196 °C. Measurement of t(14;18) frequency within the *BCL2* major breakpoint region (MBR) was reported previously [8]. Briefly, quantitative real-time PCR (Q-PCR) was performed in triplicate on DNA extracted from buffy coats. The absolute frequency was calculated using the standard curve method and normalized to the total number of cells in the samples. Prevalence [$t(14:18)^{+}$] was defined as more than one cell bearing the translocation per 1,000,000 circulating lymphocytes ($\geq 1 \times 10^{-6}$). Frequency was defined as the number of circulating t(14;18) copies per 1,000,000 circulating lymphocytes.

Statistical analysis

Baseline differences between the control and incident FL cases were assessed using the Student's *t* test and the Chi-squared test for continuous and categorical variables, respectively. The Wilcoxon rank-sum test was used to assess differences in median translocation frequency.

To identify variables that may be involved in the etiology of the translocation, three outcomes were considered: $t(14;18)^{+}$ as a binary variable, t(14;18) frequency as a continuous variable, and $t(14;18)^{HF}$, the previously defined biomarker threshold indicating a greatly increased risk of FL. This was a binary variable categorized as a frequency greater than 1×10^{-4} and was associated with a 23-fold higher risk of FL [8]. Individuals below this frequency were considered to have a low frequency [denoted as $t(14;18)^{LF}$]. The following variables available from the EPIC cohort were considered as possible determinants of these outcomes: age (years), sex, measured height (cm),

[33] Department of Perioperative and Surgical Sciences, Umeå University, Umeå, Sweden

[34] Department of Radiation Sciences, Head Regional Cancer Center North, Umeå University, Umeå, Sweden

[35] Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark

[36] Danish Cancer Society Research Center, Copenhagen, Denmark

[37] International Agency for Research on Cancer (IARC-WHO), 69372 Lyon, France

Table 1 Demographic and clinical characteristics of EPIC incident FL cases and controls

Demographic and clinical characteristics	Controls	Cases	*p* value
N	251	135	
Age at blood draw in years, mean (range)	54.4 (26.8–77.7)	53.3 (26.7–77.2)	0.237
Sex—no (%)			
Male	88 (35.1)	52 (38.5)	0.500
Female	163 (64.9)	83 (61.5)	
Calendar years of enrollment	1992–1998	1992–1998	
Country			
Germany	13 (5.4)	17 (12.0)	0.154
Greece	9 (3.7)	4 (2.8)	
Italy	54 (22.2)	23 (16.2)	
Norway	13 (5.4)	5 (3.5)	
Spain	27 (11.1)	15 (10.6)	
Sweden[a]	33 (13.2)	35 (24.7)	
The Netherlands	46 (18.9)	19 (13.4)	
UK	49 (20.2)	24 (16.9)	
t(14;18) characteristics			
Prevalence (%)	73 (29.2)	80 (59.3)	<0.001*
Range ($\times 10^{-6}$)	1–456	1–21, 295.4	
Mean frequency ($\times 10^{-6}$)	7	993.1	<0.001*
Mean frequency in t(14;18)$^{+}$ samples ($\times 10^{-6}$)	21.6	1,658.8	0.0011*
Median frequency ($\times 10^{-6}$)	1	2	<0.001*
Median frequency in t(14;18)$^{+}$ samples ($\times 10^{-6}$)	2	38	<0.001*
Prevalence of samples above 10^{-4} (%)	4 (1.6)	29 (22.0)	<0.001*

* Significant difference at the 95 % CI

[a] In the study [8] which forms the basis of this current work, these Swedish cases were included in the validation dataset, as they are additionally participants in the Northern Sweden Health and Disease study

measured weight (kg), computed body mass index (BMI) (kg/m^2), country of residence, smoking status (never, former, current), Cambridge physical activity index (inactive, moderately inactive, moderately active, active), highest educational level achieved (as a proxy for socioeconomic status; none, primary, technical/professional, secondary, university), and 48 dietary variables measured by computed mean daily intake. Age, height, weight, BMI, and the dietary variables were considered as both continuous and categorical variables. The continuous variables were categorized into ordered quartiles based on the distribution of exposure in the total population. The only exception was BMI, which was based on the WHO defined categories [normal (18.5–24.9 kg/m^2), overweight (25–29.9 kg/m^2), obese (>30 kg/m^2)]. Three participants who had a BMI between 18.3 and 18.5 were grouped together with participants of normal weight for the purposes of these analyses.

For the binary outcomes [t(14;18)$^{+}$ and t(14;18)^{HF}], the Chi-squared and Student's *t* tests were used to explore categorical and continuous explanatory variables, respectively. Analyses of translocation frequency as a continuous variable were conducted in t(14;18)$^{+}$ individuals only. Kruskal–Wallis nonparametric ANOVA and Spearman's rank correlation were used to assess correlation with frequency for categorical and continuous explanatory variables, respectively. Stepwise analyses including all variables in the same model were then run to determine which variables retained significance when all other variables were taken into account. Analyses were performed in incident FL cases [those participants who went on to develop FL ($n = 135$)] and in controls (those who remained healthy until the end of follow-up, $n = 251$), and the findings were compared.

Variables which retained statistical significance in the stepwise models were then explored further. Unconditional logistic regression adjusting for the matching factors (age, sex, and country) and using all controls as reference was used to determine whether the risk associated with exposure differed between the t(14;18)$^{+}$, t(14;18)$^{-}$, t(14;18)^{HF}, and t(14;18)^{LF} cases. Unconditional analyses were used to maximize the included population and therefore the power, as after the exclusion of participants with insufficient DNA quantity or quality matching was lost for a number of cases. Tests for trend were performed using the Wald test statistic based on the continuous quartiles. A log-linear

Table 2 t(14;18) frequency and prevalence by lifestyle and anthropometric variables

Variable	Controls					Incident FL cases				
	n	Prevalence $t(14;18)^+$ *n* (%)	Median frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Mean frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Prevalence $t(14;18)^{HF}$ *n* (%)	*n*	Prevalence $t(14;18)^+$ *n* (%)	Median frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Mean frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Prevalence $t(14;18)^{HF}$ *n* (%)
Total	251	73 (29.2)	2.0	21.6	4 (1.6)	135	80 (59.3)	38.0	1,658.8	29 (22.0)
Age (years)										
25–45	32	5 (15.6)	6.0	34.1	1 (3.1)	21	13 (61.9)	17.6	1,733.7	4 (3.0)
45–55	103	31 (30.1)	5.3	19.4	1 (1.0)	59	35 (59.3)	41.0	2,300.0	14 (10.4)
55–65	87	27 (31.4)	2.0	28.2	2 (2.3)	44	27 (61.4)	21.4	986.1	8 (5.9)
65+	29	10 (34.5)	2.0	4.5	0 (0)	11	5 (45.5)	338.0	344.5	3 (2.2)
Sex										
Males	88	23 (26.4)	2.0	13.6	1 (1.5)	52	33 (63.5)	41.0	1,899.2	15 (28.9)
Females	163	50 (30.7)	2.6	25.3	3 (1.8)	83	36 (56.6)	14.5	1,486.3	14 (17.5)
BMI										
Underweight	2	0 (0)	–	–	–	1	0 (0)	–	–	–
Normal	113	31 (27.4)	5.0	30.6	3 (2.7)	71	42 (59.2)	14.0	1,020.7	14 (19.7)
Overweight	93	29 (31.2)	2.0	10.2	0 (0)	48	27 (56.3)	76.6	2,435.4	12 (25.0)
Obese	43	13 (30.2)	6.0	25.6	1 (2.3)	14	10 (71.4)	13.1	1,896.1	2 (14.3)
Unknown	0	–	–	–	–	1	1 (100)	4,475.9	4,475.9	1 (100.0)
Height										
144–159.9 cm	70	15 (21.4)	2.0	8.4	0 (0)	25	9 (36.0)*	14.0	677.6	1 (4.6)
160–165 cm	63	23 (36.5)	4.2	45.0	4 (4.8)	33	18 (54.6)*	35.0	1,821.0	6 (18.2)
165.5–172.9 cm	56	21 (38.2)	2.0	8.4	0 (0)	32	20 (62.5)*	31.4	1,951.6	7 (21.9)
173–189.3 cm	62	14 (22.6)	2.3	17.3	1 (1.6)	44	32 (72.7)*	41.0	1,541.8	14 (31.8)
Unknown	0	–	–	–	–	1	1 (100.0)	4,475.9	4,475.9	1 (100.0)
Weight										
41.2–62.1 kg	67	19 (28.4)	4.2	16.4	1 (1.5)	29	12 (41.4)	94.8	153.8	5 (19.2)
62.4–70.1 kg	60	16 (26.7)	2	42.6	2 (3.3)	36	23 (63.9)	12	1,626.9	5 (13.9)
70.2–79.5 kg	61	16 (26.3)	9	13.71	0 (0)	35	21 (60.0)	38.5	1,357.3	8 (22.9)
79.7–125.4 kg	63	22 (35.5)	2	16.6	1 (1.6)	33	22 (66.7)	54.7	1,711.7	9 (27.3)
Unknown	0	–	–	–	–	2	2 (100)	12,885.7	12,885.7	2 (100)
Smoking status										
Never	127	40 (31.5)	3.1	18.5	1 (0.8)	64	34 (53.1)	14	454.6	10 (15.6)
Former	71	19 (26.8)	2	19.1	1 (1.4)	40	27 (67.5)	41	2,354.6	11 (27.5)
Current	49	14 (28.6)	5.6	34	2 (4.1)	30	19 (6.3)	73	2,761.3	8 (26.7)
Unknown	4	0 (0.0)	–	–	0 (0.0)	1	0 (0.0)	–	–	0 (0.0)
Educational level										
None	11	1 (9.1)	13	13	0 (0.0)	7	1 (14.3)*	58	58	0 (0.0)
Primary	78	22 (28.2)	2.3	26.6	1 (1.3)	41	21 (51.2)*	38	1,939.1	9 (22.0)
Tech/Prof	54	17 (31.5)	2	5.1	0 (0.0)	33	20 (60.6)*	17.6	1,629.2	7 (21.9)
Secondary	43	12 (27.9)	4.7	27.6	1 (2.3)	20	15 (75.0)*	17.7	1,802.4	4 (22.2)
University	48	17 (35.4)	6.3	28.1	2 (4.2)	27	20 (74.1)*	39.7	1,609.8	9 (33.3)
Not specified	17	4 (23.5)	3.5	21.5	0 (0.0)	7	3 (42.9)	2.5	25.8	0 (0.0)
Activity level										
Inactive	55	16 (29.1)	6.5	40	1 (1.8)	32	17 (53.1)	17.7	1,132.3	4 (12.5)
Mod. inactive	67	21 (31.3)	3.2	14.5	1 (1.5)	41	22 (53.7)	28	2,731.5	6 (14.6)
Mod. active	69	16 (23.2)	2	9.5	0 (0.0)	35	24 (68.6)	110	1,094.1	12 (34.3)

Table 2 continued

Variable	Controls					Incident FL cases				
	n	Prevalence $t(14;18)^+$ *n* (%)	Median frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Mean frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Prevalence $t(14;18)^{HF}$ *n* (%)	*n*	Prevalence $t(14;18)^+$ *n* (%)	Median frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Mean frequency within all $t(14;18)^+$ individuals ($\times 10^{-6}$)	Prevalence $t(14;18)^{HF}$ *n* (%)
Active	43	14 (32.6)	5.2	33	2 (4.7)	21	13 (61.9)	92	607.9	6 (28.6)
Missing	17	6 (35.3)	2	3.1	0 (0.0)	6	4 (66.7)	23.5	4,537.3	1 (1.7)

HF high frequency of translocations ($>10^{-4}$)

* Significant difference across strata at the 95 % CI. The Chi-squared test was used to assess difference in the proportion of individuals who were positive for the translocation and the percentage who were above the high-frequency threshold. Kruskal–Wallis nonparametric ANOVA was used to consider difference in translocation frequency between the groups. Where the variable was categorized as unknown, these individuals were excluded from analysis

model was used to search for interactions between these variables and the translocation in the risk of FL.

Results

A total of 386 participants were eligible for inclusion in this study, including 135 incident cases of FL and 251 controls who were free of cancer at the end of follow-up. Incident FL diagnosis occurred between 2 months and 20 years (mean 4.6 years) after blood draw. Due to the exclusion of participants with insufficient DNA, the 1:2 case–control matching was lost for a number of cases; however, there were no significant differences in age, gender, years of enrollment, or country of residence between the included cases and controls (Table 1). As reported previously [8], the participants who went on to develop FL were significantly more likely to be $t(14;18)^+$ and displayed significantly higher translocation frequencies in the blood samples taken at recruitment. In total, 22 % of incident FL cases displayed frequencies $>1 \times 10^{-4}$ [$t(14;18)^{HF}$] compared to only 2 % of controls ($p < 0.0001$).

Tables 2 and 3 describe the relationship between $t(14;18)^+$, translocation frequency and $t(14;18)^{HF}$ with lifestyle, anthropometric, and dietary variables that were considered potential determinants of these outcomes. In Table 3, only those 10 dietary factors (of the 48 investigated) with a significant *p* value in either cases ($n = 1$) or controls ($n = 9$) are shown.

Among the incident FL cases, height (χ^2 $p = 0.025$) and educational level (χ^2 $p = 0.021$) were both associated with t(14;18) prevalence. $t(14;18)^{HF}$ cases reported a taller mean height than non-$t(14;18)^{HF}$ cases (*t* test *p* value = 0.006), although there was no significant difference in the prevalence of high-frequency cases across the categories (χ^2 *p* value = 0.08). Consumption of fruits (grams per day) was associated with a lower frequency of the translocation (Spearman's $p = 0.017$).

In the control only population $t(14;18)^+$ individuals reported lower daily consumption of calories (*t* test *p* value = 0.0321), carbohydrates ($p = 0.0142$), cereals ($p = 0.0272$), and starch ($p = 0.0126$) than $t(14;18)^-$ individuals. However, starch consumption was associated with an increased frequency among those who were positive for the translocation (Spearman's $p = 0.027$), as was consumption of cake products ($p = 0.021$). High-frequency controls [$t(14;18)^{HF}$] reported significantly lower daily consumption of dietary fiber ($p = 0.041$), vitamin C ($p = 0.025$), thiamine ($p = 0.027$), and folate ($p = 0.040$).

There was a high degree of correlation between the lifestyle (online resource 1) and dietary (online resource 2) variables. To account this correlation, stepwise regression models were utilized to determine which variables retained statistical significance. Under these analyses, additional variables showed an association with translocation frequency or prevalence (Table 4). These 19 variables were therefore taken forward to determine whether they were differentially associated with the risk of NHL according to translocation status in the prediagnostic blood samples.

Association with disease status

Only three of these nineteen variables—height, educational status, and cake consumption—showed a significant association with either $t(14;18)^-$, $t(14;18)^+$, $t(14;18)^{LF}$, or $t(14;18)^{HF}$ FL. These results are presented in Table 5. Additionally when all nineteen variables were entered into the model simultaneously, these associations for height, educational level, and cake consumption remained consistent.

Table 3 Dietary variables significantly ($p < 0.05$) associated with t(14;18) prevalence and frequency

Dietary nutrient	Population	t(14;18) prevalence			t(14;18) frequency		High translocation frequency		
		Mean daily intake		p value	Spearman's ρ	p value	Mean daily intake		p value
		$t(14;18)^-$	$t(14;18)^+$				non-$t(14;18)^{HF}$	$t(14;18)^{HF}$	
Total calories (kcal)	Cases	2,091	2,068.5	0.857	−0.0221	0.856	2,091.4	2,070.1	0.889
	Controls	2103.6	1907.7	0.032*	0.141	0.236	2049.8	1,833.7	0.513
Potatoes (g)	Cases	86.8	97.9	0.462	0.2038	0.091	97.3	87	0.577
	Controls	100.7	86.9	0.402	0.184	0.122	97.5	47.5	0.398
Vegetables (g)	Cases	195.2	150.3	0.063	−0.0814	0.503	173.6	153.9	0.506
	Controls	190.2	197.7	0.699	−0.185	0.12	194.1	93.7	0.147
Fruit, nuts, seeds (g)	Cases	242.8	230.2	0.699	−0.2837	0.017*	247.2	192.1	0.164
	Controls	237.6	261.7	0.313	−0.108	0.366	246.4	137.8	0.205
Cakes (g)	Cases	59	44.4	0.139	0.0682	0.575	51.3	50.6	0.954
	Controls	46.4	42.8	0.584	0.272	0.021*	44.9	72.3	0.25
Carbohydrates (g)	Cases	235.7	238.2	0.877	−0.0429	0.725	237.4	241.6	0.825
	Controls	235.6	210.3	0.014*	0.197	0.097	228.8	189.5	0.293
Starch (g)	Cases	127.8	134.7	0.573	0.0079	0.948	130.5	141.2	0.473
	Controls	127.7	108.8	0.013*	0.261	0.027*	122.5	104.7	0.517
Cereal (g)	Cases	220.1	245.9	0.348	−0.0092	0.939	227.9	266.1	0.252
	Controls	228	190.1	0.027*	0.15	0.208	217.7	165.9	0.404
Dietary fiber (g)	Cases	21.8	21.8	0.999	−0.0814	0.503	22	21.6	0.788
	Controls	22.1	21.1	0.329	0.001	0.992	22	14.3	0.041*
Vitamin C (mg)	Cases	117.7	113.1	0.696	−0.1092	0.368	117.6	107.4	0.481
	Controls	120.8	131.5	0.21	−0.123	0.305	125.1	56.4	0.025*
Thiamine (B1) (mg)	Cases	1.29	1.2	0.262	−0.0455	0.708	1.28	1.16	0.248
	Controls	1.29	1.21	0.358	0.029	0.809	1.27	0.66	0.027*
Folate (B9) (μg)	Cases	292	275.2	0.328	−0.0613	0.614	286.5	271.2	0.461
	Controls	296.4	285.7	0.491	−0.01	0.932	295.1	181.5	0.040*

* Significant difference at the 95 % CI

Among incident FL cases who were positive for the translocation, there was a more than sixfold increase in risk of malignancy associated with being in the tallest quartile [OR 6.31 (95 % CI 2.11, 18.9)] with evidence of a significant trend (p trend = 0.001). This was also evident for $t(14;18)^{HF}$ cases, with an increased risk observed in the tallest compared to the shortest participants [OR 6.74 (95 % CI 1.69, 27.0, p trend = 0.007]. Conversely, incident $t(14;18)^-$ cases showed no significant association with height, and the risk estimates were decreased in the two tallest quartiles. An increased level of formal education was associated with a decreased risk of incident $t(14;18)^-$ FL (p trend 0.029), but an increased (although nonsignificant) risk of $t(14;18)^+$ FL. There was little evidence that the association between risk and cake consumption differed according to translocation frequency. According to the log-linear model, there was evidence of an interaction between the translocation with height ($p = 0.008$) and with educational level ($p = 0.030$) in the risk of FL, but there was no interaction with cake consumption ($p = 0.534$). Similarly, there was an evidence of interaction between a high frequency with height ($p = 0.003$), while for educational level it was borderline significant ($p = 0.061$) and for cake consumption nonsignificant ($p = 0.163$).

In all analyses, a significance level of 5 % was utilized; however, it should be noted that if multiple testing were accounted for, none of the reported associations would retain significance.

Discussion

The identification of a committed precursor to FL development linked to t(14;18) translocation frequency may afford the possibility of translation into a clinically useful predictive biomarker [8]. Yet the causes of this translocation and the factors which promote its clonal expansion remain unknown. In this study, we explored lifestyle and

Table 4 Variables that retained significance in the stepwise regression models for t(14;18) prevalence and frequency

Outcome	Population	Variable[a]	OR	95 % CI	*p* value[b]
t(14;18) prevalence	Cases	Height (cm)	1.06	(1.01, 1.11)	0.025
		Educational status	1.62	(1.16, 2.25)	0.004
	Controls	Veal (g)	1.04	(1.01, 1.08)	0.009
		Vitamin B6 (mg)	0.11	(0.03, 0.42)	0.001
		Vitamin C (mg)	1.01	(1.01, 1.02)	0.001
		Beef (g)	1.02	(1.01, 1.04)	0.003
		Pork (g)	0.97	(0.94, 1.00)	0.023
		Thiamine (B1) (mg)	4.71	(1.07, 20.77)	0.041
		Olive oil (g)	0.97	(0.95, 1.00)	0.027

Outcome	Population	Variable[a]	Beta coef.	95 % CI	*p* value[b]
t(14;18) frequency	Cases	Height (cm)	−509.50	(−848.49, −170.51)	0.004
		Weight (kg)	659.06	(272.39, 1,045.73)	0.001
		BMI	−1,777.78	(−2,869.94, −685.61)	0.002
		Sugar (g)	−296.758	(−452.46, −141.06)	<0.001
		Meat (g)	21.06227	(7.78, 34.35)	0.002
		Iron (mg)	−288.8968	(−490.67, −87.12)	0.005
		Carbohydrate (g)	295.7241	(140.99, 450.46)	<0.001
		Starch (g)	−287.2917	(−439.21, −135.37)	<0.001
	Controls	Cakes (g)	0.13	(0.03, 0.23)	0.009
		Rabbit (g)	1.85	(0.97, 2.74)	<0.001
		Olive oil (g)	−0.34	(−0.65, −0.04)	0.027
		Polyunsaturated fats (g)	−0.90	(−1.73, −0.07)	0.033

[a] Dietary variables are measured by computed mean daily intake

[b] All displayed variables are significant at the 95 % CI

dietary factors as possible initiators of the t(14;18) translocation or its clonal expansion. Furthermore, we considered whether the association between these factors and disease risk differed according to the t(14;18) status of the cases, in order to provide further inferences on this novel predictive biomarker.

The prevalence of $t(14;18)^{+}$ individuals among the incident FL cases was 59 %, and the prevalence of $t(14;18)^{HF}$ was 22 %, while among controls it was only 29 and 2 %, respectively. A taller height and a higher level of formal education [as a proxy for socioeconomic status (SES)] were significantly associated with a greater prevalence of the translocation in the cases, but not in controls. Similarly, an inverse association with fruit consumption was evident only among the cases. Taller cases also displayed significantly increased frequencies of circulating translocated cells and were more likely to be $t(14;18)^{HF}$. Additionally, a large number of dietary variables showed significant associations for both prevalence and frequency among the controls. No associations were common between the two populations, and this remained true when a stepwise model was run to account for potential correlation between the variables.

Among those variables associated with translocation prevalence in controls, thiamine and folate both play a role in the one-carbon metabolism and therefore impact epigenetic pathways. Given the importance of epigenetic processes in the regulation of genomic integrity and stability, these associations might be worth further consideration. The lack of association of these factors with prevalence among cases could be explained by impaired epigenetic signalling in the pathogenesis of some hematopoietic cancers, including FL [20–22].

A potential interpretation of the association between height and translocations is based on the existing evidence that both height and a higher SES have been associated with an increased risk of NHL [23, 24]. In fact, in this study the increased risk of NHL associated with height, in agreement with a previous study of FL [25], was restricted to the $t(14;18)^{+}$-positive FL cases and was particularly marked in the $t(14;18)^{HF}$ cases. In a smaller subset of this population for whom the information was available, this association was also found to be apparent for sitting height (online resource 1). Sitting height has previously been shown to be positively associated with indicators of both higher childhood SES and delayed infection [26], the two variables being strongly correlated. The increased risk of FL may therefore be related to the altered th1/th2 response, which is thought to be mediated by delayed exposure to

Table 5 Association between selected risk factors and FL stratified by t(14;18) status of the cases

Variable	n. co	$t(14;18)^-$ cases			$t(14;18)^+$ cases			p for interaction[b]	$t(14;18)^{LF}$ cases			$t(14;18)^{HF}$ cases			p for interaction[c]
		n. ca	OR (95 % CI)[a]	p trend	n. ca	OR (95 % CI)[a]	p trend		n. ca	OR (95 % CI)[a]	p trend	n. ca	OR (95 % CI)[a]	p trend	
Height															
144–159.9 cm	70	16	1		9	1			21	1			–[a]		
160–165 cm	63	15	1.00 (0.45, 2.23)		18	2.14 (0.88, 5.18)			27	1.33 (0.67, 2.62)		7	1		
165.5–172.9 cm	56	12	0.81 (0.31, 2.09)		20	3.07 (1.23, 7.69)			25	1.46 (0.69, 3.08)		7	2.78 (0.87, 8.81)		
173–189.3 cm	62	12	0.64 (0.19, 2.14)	0.485	32	6.31 (2.11, 18.9)	0.001*	0.008*	30	1.65 (0.65, 4.20)	0.271	14	6.74 (1.69, 27.0)	0.007*	0.003*
Educational status															
None/primary	89	26	1		22	1			39	1		9	1		
Tech/Prof	54	13	0.79 (0.37, 1.69)		20	1.40 (0.69, 2.83)			25	0.94 (0.51, 1.75)		7	1.23 (0.43, 3.56)		
Secondary	43	5	0.36 (0.13, 1.02)		15	1.44 (0.66, 3.14)			14	0.68 (0.33, 1.41)		4	0.99 (0.28, 3.53)		
University	48	7	0.44 (0.17, 1.12)	0.029*	20	1.53 (0.74, 3.17)	0.251	0.030*	18	0.74 (0.37, 1.48)	0.276	9	1.61 (0.57, 4.49)	0.434	0.061
Cake intake															
0–15.1 (g)	66	8	1		18	1			21	1		4	1		
15.7–34.2 (g)	62	14	1.89 (0.74, 4.84)		16	0.91 (0.43, 1.95)			19	0.93 (0.46, 1.91)		10	2.64 (0.78, 8.91)		
34.5–59.5 (g)	63	7	0.93 (0.32, 2.74)		22	1.2 (0.59, 2.48)			22	1.03 (0.51, 2.08)		6	1.54 (0.41, 5.76)		
>60 (g)	56	21	3.19 (1.31, 7.80)	0.029*	15	1.00 (0.46, 2.18)	0.794	0.534	30	1.72 (0.88, 3.34)	0.103	6	1.65 (0.44, 6.21)	0.751	0.163

* Significant difference at the 95 % CI

[a] Only one case had a height between 144 and 159.9 cm, and so this category was combined with the 160–165 cm category ($n = 7$ cases and 133 controls) p for interaction between the variable and [b] translocation prevalence/[c] high translocation frequency in the risk of FL computed using a log-linear model

infectious agents [27, 28]. Education status, as a proxy for SES, also showed an inverse association with $t(14;18)^-$ FL, but a positive association with $t(14;18)^+$ FL, that was robust to adjustment for height. We report evidence of a potential interaction for both height and education level with the t(14;18) translocation in the risk of FL.

Previous studies have observed that cases of NHL positive for the t(14;18) translocation might constitute a more homogenous group of malignancies, including in etiological terms, than NHL as a whole [29]. Taken together our reported findings may add further support to this hypothesis. However, in this study we did not replicate the previously reported differing associations between $t(14;18)^+$ and $t(14;18)^-$ FL related to milk and coffee consumption or cigarette smoking [29–31]. The risk associated with pesticide and PCB exposure, lifetime number of surgeries, and a family history of hematopoietic cancer [25, 30, 32] has also been observed to differ by t(14;18) status, but this information was only available for a small proportion of our cohort and so has not been included in these analyses.

Our study has a number of strengths. The study population is from EPIC, a large well-characterized cohort with stringent inbuilt quality control, calibration, and validation procedures. The biomarker was measured in a blinded fashion using a recognized and proven screening assay in the same laboratory under the same conditions for all samples. A proportion of participants were excluded on the basis of unavailable or poor DNA quality DNA samples, but this is unlikely to be related to translocation or disease status. Finally, the disease analyses were adjusted for potential confounding factors which may be related to t(14;18) frequency.

A potential limitation is that all analyses rely on a single spot measurement from the same blood sample provided at recruitment together with the questionnaire data. The questionnaire data therefore reflect the exposure at the time of translocation assessment and may not be representative of lifetime experience, making the determination of a temporal pathway difficult. Further, we are reliant on self-reported data. The nesting within EPIC raises issues of the generalizability and external validity of the findings due to the European Caucasian population and the related geographical preferences and regulations. There was a high correlation between the identified dietary variables such as fruit, vegetables and vitamin C, and potatoes and starch, but we attempted to account for this using stepwise analyses. Translocation status was determined based on the prediagnostic blood samples, rather than tissue biopsies taken at diagnosis. However, we believe this approach is justified as we have previously demonstrated that t(14;18)-positive clones present in prediagnostic samples constituted bona fide precursors which progressed to overt FL and that those cases of FL associated with a high translocation frequency in blood represent a distinct subset of cases [8].

In conclusion, this hypothesis generating study suggests a potential role for lifestyle factors in the prevalence and frequency of translocations and provides a number of avenues that may be worthy of further exploration. Despite the limitations discussed, the weight of these findings additionally lends some support to the previously proposed hypothesis that the etiology of FL may differ by t(14;18) status.

Acknowledgments We thank the investigators and participants of the EPIC (European Prospective Investigation into Cancer and Nutrition).

Funding 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, DutchZorgOnderzoek 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 del'Education Nationale;3 M(France); Gustave Roussy Institute; and General Councils of France.

Compliance with ethical standards

Conflicts of interest R. Kelly, S. Roulland, P.Vineis, and B. Nadal hold a patent for the $t(14;18)^{HF}$ biomarker (Nadal B, Roulland S, Vineis P, Kelly RS. Methods for Predicting Whether a Subject is at Risk of Developing a Follicular Lymphoma France: 12118, filed December 2013).

Informed consent The study was approved by the committees on research ethics each of the participating centers and conducted in accordance with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Anderson JR, Armitage JO, Weisenburger DD (1998) Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma classification project. Ann Oncol 9:717–720
2. Jaffe ES (2013) Follicular lymphomas: a tapestry of common and contrasting threads. Haematologica 98:1163–1165
3. Kridel R, Sehn LH, Gascoyne RD (2012) Pathogenesis of follicular lymphoma. J Clin Invest 122:3424–3431
4. Schüler F, Dölken L, Hirt C et al (2009) Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals. Int J Cancer 124:958–963

5. Limpens J, Stad R, Vos C et al (1995) Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. Blood 85:2528–2536
6. Liu Y, Hernandez AM, Shibata D, Cortopassi GA (1994) BCL2 translocation frequency rises with age in humans. Proc Natl Acad Sci USA 91:8910–8914
7. Roulland S, Lebailly P, Lecluse Y, Briand M, Pottier D, Gauduchon P (2004) Characterization of the t(14;18) BCL2-IGH translocation in farmers occupationally exposed to pesticides. Cancer Res 64:2264–2269
8. Roulland S, Kelly RS, Morgado E et al (2014) t(14;18) Translocation: a predictive blood biomarker for follicular lymphoma. J Clin Oncol 32:1347–1355
9. Baccarelli A, Hirt C, Pesatori AC et al (2006) t(14;18) translocations in lymphocytes of healthy dioxin-exposed individuals from Seveso, Italy. Carcinogenesis 27:2001–2007
10. Schroeder JC, Olshan AF, Baric R et al (2001) Agricultural risk factors for t(14;18) subtypes of non-Hodgkin's lymphoma. Epidemiology 12:701–709
11. Bentham G, Wolfreys AM, Liu Y et al (1999) Frequencies of hprt–mutations and bcl-2 translocations in circulating human lymphocytes are correlated with United Kingdom sunlight records. Mutagenesis 14:527–532
12. Schüler F, Hirt C, Dolken G (2003) Chromosomal translocation t(14;18) in healthy individuals. Semin Cancer Biol 13:203–209
13. Bell DA, Liu Y, Cortopassi GA (1995) Occurrence of bcl-2 oncogene translocation with increased frequency in the peripheral blood of heavy smokers. J Natl Cancer Inst 87:223–224
14. Riboli E, Kaaks R (1997) The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 26:S6
15. Riboli E, Hunt K, Slimani N et al (2002) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 5:1113–1124
16. Roswall N, Freisling H, Bueno-de-Mesquita HB et al (2014) Anthropometric measures and bladder cancer risk: a prospective study in the EPIC cohort. Int J Cancer 135:2918–2929
17. Slimani N, Kaaks R, Ferrari P et al (2002) European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. Public Health Nutr 5:1125–1145
18. Slimani N, Deharveng G, Unwin I et al (2007) The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. Eur J Clin Nutr 61:1037–1056
19. Kaaks R, Riboli E (1997) Validation and calibration of dietary intake measurements in the EPIC project: methodological considerations. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 26:S15
20. Pasqualucci L, Dominguez-Sola D, Chiarenza A et al (2011) Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 471:189–195
21. Shaknovich R, Melnick A (2011) Epigenetics and B-cell lymphoma. Curr Opin Hematol 18:293–299
22. Okosun J, Bödör C, Wang J et al (2014) Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat Genet 46:176–181
23. Britton JA, Khan AE, Rohrmann S et al (2008) Anthropometric characteristics and non-Hodgkin's lymphoma and multiple myeloma risk in the European Prospective Investigation into Cancer and Nutrition (EPIC). Haematologica 93:1666–1677
24. Fisher SG, Fisher RI (2004) The epidemiology of non-Hodgkin's lymphoma. Oncogene 23:6524–6534
25. Chang CM, Wang SS, Dave BJ et al (2011) Risk factors for non-Hodgkin lymphoma subtypes defined by histology and t(14;18) in a population-based case–control study. International journal of cancer. J Int du Cancer 129:938–947
26. Wadsworth ME, Hardy RJ, Paul AA, Marshall SF, Cole TJ (2002) Leg and trunk length at 43 years in relation to childhood health, diet and family circumstances; evidence from the 1946 national birth cohort. Int J Epidemiol 31:383–390
27. Vineis P, Miligi L, Crosignani P et al (2000) Delayed infection, family size and malignant lymphomas. J Epidemiol Community Health 54:907–911
28. Greaves M (2006) Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer 6:193–203
29. Chiu BC-H, Lan Q, Dave BJ, Blair A, Zahm SH, Weisenburger DD (2008) The utility of t(14;18) in understanding risk factors for non-Hodgkin lymphoma. JNCI Monogr 2008:69–73
30. Chiu BC, Dave BJ, Blair A et al (2007) Cigarette smoking, familial hematopoietic cancer, hair dye use, and risk of t(14;18)-defined subtypes of non-Hodgkin's lymphoma. Am J Epidemiol 165:652–659
31. Chiu BC, Dave BJ, Ward MH et al (2008) Dietary factors and risk of t(14;18)-defined subgroups of non-Hodgkin lymphoma. Cancer Causes Control 19:859–867
32. Chiu BC, Dave BJ, Blair A, Gapstur SM, Zahm SH, Weisenburger DD (2006) Agricultural pesticide use and risk of t(14;18)-defined subtypes of non-Hodgkin lymphoma. Blood 108:1363–1369