

Breast cancer risk, nightwork, and circadian clock gene polymorphisms

Thérèse Truong^{1,2}, Benoît Liquet^{3,4}, Florence Menegaux^{1,2}, Sabine Plancoulaine^{2,5}, Pierre Laurent-Puig⁶, Claire Mulot⁶, Emilie Cordina-Duverger^{1,2}, Marie Sanchez^{1,2}, Patrick Arveux⁷, Pierre Kerbrat⁸, Sylvia Richardson³ and Pascal Guénel^{1,2}

¹Inserm, CESP Center for Research in Epidemiology and Population Health, U1018, Environmental Epidemiology of Cancer, Villejuif, France

²Université Paris-Sud, UMRS 1018, Villejuif, France

³Biostatistical Unit, MRC, Cambridge, UK

⁴School of Mathematics and Physics, The University of Queensland, St Lucia, Queensland, Australia

⁵Inserm, CESP Center for Research in Epidemiology and Population Health, U1018, Epidemiology of Diabetes, Obesity and Chronic Kidney Disease Over Lifecourse, Villejuif, France

⁶Université Paris Descartes, INSERM UMR-S775 EPIGENETEC, Paris, France

⁷Département d'informatique médicale, Center Georges-François Leclerc, Dijon, France

⁸Center Eugène Marquis, Rennes, France

Correspondence should be addressed to P Guénel
Email
pascal.guenel@inserm.fr

Abstract

Night shift work has been associated with an increased risk of breast cancer pointing to a role of circadian disruption. We investigated the role of circadian clock gene polymorphisms and their interaction with nightwork in breast cancer risk in a population-based case–control study in France including 1126 breast cancer cases and 1174 controls. We estimated breast cancer risk associated with each of the 577 single nucleotide polymorphisms (SNPs) in 23 circadian clock genes. We also used a gene- and pathway-based approach to investigate the overall effect on breast cancer of circadian clock gene variants that might not be detected in analyses based on individual SNPs. Interactions with nightwork were tested at the SNP, gene, and pathway levels. We found that two SNPs in *RORA* (rs1482057 and rs12914272) were associated with breast cancer in the whole sample and among postmenopausal women. In this subpopulation, we also reported an association with rs11932595 in *CLOCK*, and with *CLOCK*, *RORA*, and *NPAS2* in the analyses at the gene level. Breast cancer risk in postmenopausal women was also associated with overall genetic variation in the circadian gene pathway ($P=0.04$), but this association was not detected in premenopausal women. There was some evidence of an interaction between *PER1* and nightwork in breast cancer in the whole sample ($P=0.024$), although the effect was not statistically significant after correcting for multiple testing ($P=0.452$). Our results support the hypothesis that circadian clock gene variants modulate breast cancer risk.

Key Words

- ▶ breast cancer
- ▶ case–control study
- ▶ circadian rhythm
- ▶ nightwork

Endocrine-Related Cancer
(2014) 21, 629–638

Introduction

The human circadian system is controlled by a central pacemaker located in the suprachiasmatic nucleus of the anterior hypothalamus and synchronized or modulated by external cues such as light or external temperature (Fu & Lee 2003, Ko & Takahashi 2006, Rana &

Mahmood 2010, Savvidis & Koutsilieris 2012, Foster & Kreitzman 2014). This circadian clock is regulated endogenously at the molecular level by periodic transcription of genes that form a network of self-regulated feedback loop. The circadian rhythm pathway plays a

key role in the maintenance of various endocrine, physiological, and behavioral functions including cell cycle regulation, hormone secretion, body temperature, and sleep/wake cycle.

The disruption of the natural circadian rhythm due to exposure to light at night has long been suspected to have an effect on breast cancer risk (Stevens *et al.* 2014). In 2007, results from experimental and epidemiological studies led the International Agency for Research on Cancer to classify shift work that involves circadian disruption as probably carcinogenic to humans (group 2A) (Straif *et al.* 2007). These findings were supported by recent epidemiological studies reporting an increased risk of breast cancer among women working in night shifts (Lie *et al.* 2011, Hansen & Lassen 2012), including our population-based case-control study in France (Menegaux *et al.* 2013). Mechanistic hypotheses proposed to explain the potential carcinogenic effects of circadian rhythm disruption (Fritschi *et al.* 2011) include exposure to light at night that suppresses the nocturnal peak of melatonin, sleep disruption with detrimental effects on the immune function, and confusion of the circadian master clock contributing to asynchronous cell proliferation.

Genetic variation in genes involved in the circadian rhythm pathway has been the focus of particular attention in recent years. Experimental studies have reported that mutations in circadian genes may produce a lack of circadian coordination in mice (Evans & Davidson 2013). In humans, single nucleotide polymorphisms (SNPs) in circadian genes have been associated with advanced sleep-phase syndrome, diurnal preference, or sleep quality (Pedrazzoli *et al.* 2010). Several authors have also investigated the genetic variants in circadian genes in relation to breast cancer risk, but only the core circadian genes and a limited number of SNPs in each gene were investigated in epidemiological studies (Zhu *et al.* 2005, 2008, Hoffman *et al.* 2010a,b,c, Fu *et al.* 2012), and none of the reported associations were replicated. Additionally, it has been hypothesized that the increased risk of breast cancer among women working at night could be modulated by genetic polymorphisms in the circadian pathway genes. A few studies highlighted interactions between nightwork and specific circadian gene variants in relation to breast cancer (Monsees *et al.* 2012, Grundy *et al.* 2013, Zienolddiny *et al.* 2013), but again none of the reported interactions were replicated.

It is possible that breast cancer risk is not linked to only a few SNPs in circadian genes, but rather to the combined effects of genetic variants that may collectively alter the functioning of the circadian clock. Using genome-wide

data, and grouping the genetic variants by biological pathways, Li *et al.* (2011) reported suggestive evidence of an association of breast cancer with genes involved in the circadian rhythm. Thus, conducting analyses at the gene- or pathway level may be helpful to examine the overall effect of circadian gene polymorphisms on breast cancer. Assessment of the overall significance of group of SNPs in a given gene (or pathway) circumvents some of the multiple testing problems and offers the potential to highlight a gene (or a pathway) with several SNPs (or genes) individually weakly associated with the disease.

In this paper, we investigated the role of circadian genes and their interaction with nightwork in breast cancer at multiple genomic levels, extending interaction tests to gene and pathway levels. We used an extended list of circadian genes and an increased number of tag-SNPs in order to achieve a high coverage of the genetic variation in these genes.

Materials and methods

Study population

The CECILE study is a population-based case-control study conducted in Côte d'Or and Ille-et-Vilaine, two French administrative areas (*départements*) located in the eastern and western parts of France respectively. The design of the CECILE study was described previously in detail (Menegaux *et al.* 2013).

Women aged 25–75 years, resident in one of these two areas, and diagnosed with histologically confirmed breast cancer between 2005 and 2007 were eligible for this study. Of the 1553 eligible cases identified during the study period, 163 refused to participate, 151 women could not be contacted, and seven died before the interview. Finally, 1232 (79%) incident breast cancer cases were included in the study. Information on estrogen (ER) and progesterone (PR) receptor status was obtained from the pathology report.

Controls selected from the general population were women with no history of breast cancer frequency matched to the cases by the 10-year age group and study area. For recruiting the controls, phone numbers of private homes were randomly selected from the telephone directory where unlisted numbers had previously been recreated. In order to prevent differential participation rates across categories of socioeconomic status (SES), the distribution by SES of the control group was determined *a priori* to mirror the distribution by SES among women of the same age in the general population available from the census data, using quotas by SES applied during the

recruitment process. Of the 1731 eligible controls identified by phone numbers, who met the inclusion criteria, 1317 (76%) accepted to participate in an in-person interview and were included in the study.

The local ethical committee approved the study protocol and all subjects signed informed consent.

Data collection

A standardized questionnaire was administered during in-person interviews by trained interviewers to obtain information on demographic and socioeconomic characteristics, reproduction, medical history, family history of cancer, diet, lifestyle factors, and residential and occupational history over the lifetime.

For each job held for at least 6 consecutive months, we obtained a description of the work tasks, work places, occupational exposures, and work schedules. Nightwork was defined as work for at least 1 h between 1100 and 0500 h during all or part of each job period.

At the end of the interview, women were invited to give a blood sample. In case of refusal, the woman was asked to provide a buccal cell sample. Of the 2549 subjects included in the analysis, 2135 (1080 cases and 1055 controls) gave a blood sample and 348 (121 cases and 227 controls) provided a buccal cell sample.

Genotyping and quality control

We designed a custom microarray to examine associations between common genetic variants in candidate biological pathways and cancers including breast cancer. For this study, 28 candidate pathways, including 648 genes, were selected in January 2011 from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and from a literature review. A total of 8716 SNPs were selected using the freely available Tagger Software (Haploview; de Bakker *et al.* 2005) to capture SNPs within 5 kb of each gene (pairwise approach with $r^2 \geq 0.8$) with a minimum minor allele frequency (MAF) of 0.05 in the Caucasian population (CEU) genotyped by the HapMap Project (Data Release 21/Phase II, NCBI Build 36.1, assembly dbSNPb126). The customized microarray was processed by Integragen (Evry, France) using Illumina technology and Infinium iSelect custom genotyping assay. This microarray has a high coverage of the genes and the pathways of interest when compared with microarrays used in genome-wide association studies (GWAS).

Of the 8716 tag-SNPs selected to be included in the iSelect beadchip, 987 did not pass the assay design, 703

failed genotyping (no call), and ten had completion rate $<95\%$. Out of the 7016 remaining SNPs, eight failed in the Hardy–Weinberg equilibrium test at $P < 0.00001$ (Bonferroni correction based on 7016 tests). We also genotyped ten CEPH control samples and when more than one discrepancy between the genotypes was obtained and the HapMap publicly available genotypes were found for a variant, this variant was excluded ($n=12$). We excluded 143 SNPs that had MAF <0.01 among controls from analyses. After all exclusions, there were 6853 SNPs in 639 genes. In this paper, we used the genotype data obtained in the circadian rhythm pathway that included 577 validated SNPs in 23 genes (*CLOCK*, *ARNTL*, *NPAS2*, *CRY1*, *CRY2*, *PER1*, *PER2*, *PER3*, *RORA*, *RORB*, *RORC*, *BHLHE40*, *BHLHE41*, *SKP1*, *FBXW11*, *CUL1*, *TIMELESS*, *FBXL3*, *NR1D1*, *CSNK1D*, *CSNK1E*, *RBX1*, and *BTRC*).

Of the 2549 subjects in the CECILE study, sufficient DNA could be extracted only from 2432 subjects (1181 cases and 1251 controls). Of them, we further excluded 103 individuals with the genotyping completion rate $<90\%$ and 29 non-Caucasians, leaving 1126 cases and 1174 controls available for the analyses.

Statistical analysis

Odds ratios (ORs) denoting the association between breast cancer and the SNPs in circadian genes were estimated using unconditional logistic regression assuming a log-additive genetic model. Each SNP was coded as 0, 1, and 2, denoting the number of minor alleles in the genotype. In the interaction analyses with nightwork, we divided nightwork into two classes (<2 and ≥ 2 years), to be comparable to previous studies (Monsees *et al.* 2012, Grundy *et al.* 2013), and we tested the interaction with each of the 577 SNPs in the circadian genes using the likelihood ratio test comparing models with and without an interaction term. Interactions with longer durations of nightwork were also tested, but the results were essentially unchanged and are not shown.

We also investigated the associations of breast cancer with genes, observed as a combination of SNPs, and with the circadian rhythm pathway as a whole, observed as the combination of the 23 circadian genes. These analyses were conducted using the Adaptive Rank Truncated Product (ARTP) method (Yu *et al.* 2009) that can combine association signals from the SNPs in a given gene (or from the genes in a pathway) to provide a P value at the gene or pathway level respectively. In this method, a gene (or a pathway) P value is calculated based on the product of the top K smallest SNP-level P values within a gene (or the

top K' smallest gene-level P values within a pathway). Different truncation points K (or K') were used, defined according to the number of SNPs in the gene and the number of genes in the pathways. We chose the truncation points proposed by Yu *et al.* (2009), i.e. five truncation points defined at every 5% of the top SNPs for the gene-level association and ten truncation points defined at every 5% of the top genes for the pathway-level association. The smallest P value of the product over the different truncation points was used as the test statistic and its significance was assessed using permutations. As we tested multiple genes using the ARTP method, the P value associated with each gene was corrected to account for multiple testing. The P value associated with the whole pathway does not require correction for multiple testing.

To study the interaction between nightwork and genetic factors at the gene and pathway levels, we modified the ARTP method so that the SNP-level interaction P values (available at <http://cran.r-project.org/web/packages/PIGE/index.html>) can be used as an input.

All analyses were adjusted for age, study area, age at menarche (<12, 12, 13, 14, 15 years and more), age at first full-term pregnancy (<22, 22–24, 25–27, and >27 years), parity (nulliparous, 1, 2, 3, and 4+ children), current use of menopausal hormone therapy (yes or no), BMI (<18.5, 18.5–24, 25–30, and >30 kg/m²), alcohol consumption (≤3 drinks/week, 4–7 drinks/week, 8–14 drinks/week, and >14 drinks/week), and tobacco consumption (never, former, or current smokers). All analyses were stratified by menopausal status, as defined in a previous paper (Cordina-Duverger *et al.* 2013). We also conducted analysis according to the ER status (ER positive and ER negative) or PR status (PR positive and PR negative). The P values at the SNP and the gene levels were corrected for multiple testing for the number of SNPs and for the number of genes, respectively, using the false discovery rate (FDR) method (Benjamini & Hochberg 1995). All statistical tests were two sided and statistical significance required a FDR P value of 0.05 or less. Statistical analyses were conducted using SAS version 9.1 (SAS Institute, Cary, NC, USA) and R version 2.15.1.

Results

Characterization of study population and nightwork

The distribution by age, study area, and well-established breast cancer risk factors is given in Table 1. Consistent with the literature, family history of breast cancer in

first-degree relatives, early age at menarche, late age at first full-term pregnancy, low parity, current use of menopausal hormone therapy, and low BMI in premenopausal women were associated with an increased incidence of breast cancer. No association with alcohol consumption or BMI in postmenopausal women was apparent.

Table 2 provides information on the association of breast cancer with nightwork (ever/never) and duration of nightwork (<2 and ≥2 years) used for the study of interaction with genetic variants.

Main genetic effects

Analysis of individual SNPs ORs and P values associated with each of the 577 SNPs in the 23 circadian genes are given in Supplementary Table 1, see section on supplementary data given at the end of this article. In the whole sample, two SNPs in *RORA* were statistically significant after FDR correction: rs1482057 (OR=0.75 (0.65–0.86), FDR P value=0.025) and rs12914272 (OR=0.79 (0.70–0.89), FDR P value=0.026). These two SNPs were weakly correlated ($r^2=0.45$).

Stratification by menopausal status (Supplementary Table 1) indicated that breast cancer in postmenopausal women was associated with rs1482057 (OR=0.71 (0.59–0.85), FDR P value=0.058) and rs12914272 (OR=0.76 (0.65–0.89), FDR P value=0.072) in *RORA* and with rs11932595 in *CLOCK* (OR=0.74 (0.64–0.86), FDR P value=0.058), with P values at the limit of statistical significance. No association was apparent in premenopausal women. Analyses by ER or PR status did not show any association with the 577 SNPs (results not shown).

Analyses at the gene and pathway levels The P values obtained from the ARTP method for the genes and the circadian clock pathway are given in Table 3. In the total sample, *RORA* was associated with breast cancer ($P=0.025$), but this association was no longer statistically significant after correcting for multiple testing (FDR P value=0.575). The circadian rhythm pathway as a whole was not associated with breast cancer ($P=0.578$).

Among postmenopausal women, breast cancer was associated with *CLOCK*, *RORA*, and *NPAS2* ($P=0.001$, $P=0.019$, and $P=0.031$ respectively), but only the association with *CLOCK* remained statistically significant after correcting for multiple testing (FDR P value=0.023). Breast cancer was also associated with an overall genetic variation in the circadian rhythm pathway in this subgroup ($P=0.041$).

Among premenopausal women, the association with *PER3* (P value=0.023) became non-significant after correcting for multiple testing (FDR P value=0.536). The circadian pathway as a whole was not associated with breast cancer (P value=0.564).

No association was reported after stratification of the case group by ER or PR status (not shown).

Interaction with nightwork

Testing interaction at the SNP level The P values for interaction between the SNPs and nightwork are given in [Supplementary Table 2](#), see section on [supplementary data](#) given at the end of this article. After FDR correction, no interaction term was statistically significant.

Table 1 Characteristics of cases and controls in the CECILE study

	Cases $n=1126$ (%)	Controls $n=1174$ (%)	OR	95% CI
Age at reference date (years)				
25–34	41 (3.6)	40 (3.4)		
35–44	163 (14.5)	162 (13.8)		
45–54	350 (31.1)	355 (30.2)		
55–64	338 (30.0)	336 (28.6)		
65–74	234 (20.8)	281 (23.9)		
Study area (département)				
Côte d'Or	349 (31.0)	409 (34.8)		
Ille-et-Vilaine	777 (69.0)	765 (65.2)		
Hormone receptor status				
ER+/PR+	714 (63.4)			
ER+/PR–	149 (13.2)			
ER–/PR+	10 (0.9)			
ER–/PR–	158 (14.0)			
Unknown	95 (8.4)			
Family history of breast cancer in first-degree relatives				
No	933 (82.9)	1053 (89.7)	1.00	Reference
Yes	193 (17.1)	121 (10.3)	1.82	1.43–2.32
Age at menarche (years)				
<12	14 (1.2)	12 (1.0)	1.00	0.77–1.3
12	201 (17.8)	194 (16.5)	1.00	Reference
13	270 (24.0)	258 (22.0)	0.94	0.74–1.2
14	258 (22.9)	262 (22.3)	0.88	0.69–1.14
15+	212 (18.8)	232 (19.8)	0.76	0.58–0.99
Age at FFTP (years)				
<22	241 (24.0)	311 (28.4)	0.91	0.72–1.15
22–24	290 (28.8)	345 (31.5)	1.00	Reference
25–27	228 (22.7)	258 (23.6)	1.07	0.84–1.35
27+	247 (24.5)	181 (16.5)	1.63	1.27–2.09
Parity				
Nulliparous	120 (10.7)	79 (6.7)	1.00	Reference
1	182 (16.2)	153 (13.0)	0.79	0.56–1.14
2	448 (39.8)	411 (35.0)	0.71	0.52–0.97
3	269 (23.9)	365 (31.1)	0.47	0.34–0.65
4+	107 (9.5)	166 (14.1)	0.42	0.28–0.61
Current use of MHT				
No	984 (87.4)	1062 (90.5)	1.00	Reference
Yes	142 (12.6)	112 (9.5)	1.40	1.06–1.83
Alcohol (drinks/week)				
≤3	875 (77.7)	886 (75.5)	1.00	Reference
4–7	148 (13.1)	163 (13.9)	0.91	0.72–1.16
8–14	59 (5.2)	80 (6.8)	0.74	0.52–1.06
>14	44 (3.9)	45 (3.8)	0.98	0.64–1.50
BMI (kg/m ²)				
Women <50 years				
<18.5	26 (7.4)	12 (3.0)	2.15	1.06–4.36
18.5–24	252 (72.0)	254 (65.0)	1.00	Reference
25–30	52 (14.9)	82 (21.0)	0.65	0.44–0.97
30+	20 (5.7)	43 (11.0)	0.49	0.28–0.85
Women ≥50 years				
<18.5	14 (1.8)	18 (2.3)	0.81	0.39–1.65
18.5–24	413 (53.2)	399 (51.0)	1.00	Reference
25–30	226 (29.1)	239 (30.5)	0.95	0.75–1.19
30+	119 (15.3)	126 (16.1)	0.95	0.71–1.26

OR, odds ratio, adjusted for age and study area; ER, estrogen receptor; PR, progesterone receptor; FFTP, first full-term pregnancy; MHT, menopausal hormone therapy.

Table 2 Odds ratios for breast cancer associated with nightwork in the CECILE study

	Cases <i>n</i> = 1126 (%)	Controls <i>n</i> = 1174 (%)	OR ^a	95% CI	<i>P</i> value
Nightwork					
Never	973 (86.4)	1045 (89.0)	1.00	Reference	
Ever	153 (13.6)	129 (11.0)	1.32	(1.02–1.72)	0.04
Total duration of nightwork periods (years)					
<2	990 (87.9)	1067 (90.9)	1.00	Reference	
≥2	136 (12.1)	107 (9.11)	1.42	(1.08–1.88)	0.01

^aOR, odds ratio adjusted for age, study area, parity, age at first full-term pregnancy, age at menarche, family history of breast cancer, current use of hormonal replacement therapy, BMI, and tobacco and alcohol consumption.

Testing interaction at the gene and pathway levels

The *P* values for genes by strata of nightwork and the *P* values of interaction between nightwork and circadian genes are given in Table 4. The *P* values for *RORA* (*P*=0.024) among women who never worked or worked at night for <2 years, and for *CUL1* (*P*=0.040) and *PER1* (*P*=0.061) among women who worked at night for 2 or more years were no longer significant after correcting for multiple testing. The interaction *P* values of <0.05 were observed for *PER1*, *CUL1*, and *ARNTL*, but none was significant after correcting for multiple testing. No interaction was observed between nightwork and the

circadian pathway as a whole. Analyses stratified by menopausal, ER status or PR status generated non-statistically significant results (not shown).

Discussion

Our findings provide new evidence that circadian genes may play a role in breast cancer etiology, as the circadian gene pathway was associated with postmenopausal breast cancer risk in our data (*P*=0.041). Genes that contribute most to the overall association with postmenopausal breast cancer were *CLOCK* (*P*=0.02 after correcting for

Table 3 Gene and pathway levels of breast cancer association with *P* values in the circadian rhythm pathway

Genes	Region	Number of SNPs	All (1126 ca/1174 co)		Premenopausal women (450 ca/437 co)		Postmenopausal women (676 ca/737 co)	
			<i>P</i> value ^a	FDR <i>P</i> value ^b	<i>P</i> value ^a	FDR <i>P</i> value ^b	<i>P</i> value ^a	FDR <i>P</i> value ^b
<i>PER3</i>	1p36	15	0.449	0.991	0.023	0.536	0.175	0.663
<i>RORC</i>	1q21	14	0.978	0.991	0.894	0.935	0.997	0.997
<i>NPAS2</i>	2q11	62	0.200	0.739	0.714	0.935	0.031	0.238
<i>PER2</i>	2q37	11	0.772	0.991	0.334	0.854	0.978	0.997
<i>BHLHE40</i>	3p26	9	0.593	0.991	0.834	0.935	0.329	0.663
<i>CLOCK</i>	4q12	11	0.205	0.739	0.271	0.854	0.001	0.023
<i>SKP1</i>	5q31	4	0.980	0.991	0.626	0.935	0.831	0.997
<i>FBXW11</i>	5q35	8	0.870	0.991	0.449	0.935	0.375	0.663
<i>CUL1</i>	7q36	23	0.131	0.739	0.160	0.854	0.051	0.293
<i>RORB</i>	9q21	34	0.683	0.991	0.261	0.854	0.241	0.663
<i>BTRC</i>	10q24	13	0.225	0.739	0.119	0.854	0.360	0.663
<i>ARNTL</i>	11p15	24	0.991	0.991	0.137	0.854	0.444	0.729
<i>CRY2</i>	11p11	9	0.222	0.739	0.725	0.935	0.322	0.663
<i>BHLHE41</i>	12p12	4	0.912	0.991	0.779	0.935	0.528	0.810
<i>TIMELESS</i>	12q13	7	0.744	0.991	0.851	0.935	0.136	0.626
<i>CRY1</i>	12q23	7	0.129	0.739	0.324	0.854	0.226	0.663
<i>FBXL3</i>	13q22	7	0.944	0.991	0.733	0.935	0.912	0.997
<i>RORA</i>	15q22	288	0.025	0.575	0.861	0.935	0.019	0.218
<i>PER1</i>	17p13	5	0.751	0.991	0.217	0.854	0.997	0.997
<i>NR1D1</i>	17q21	8	0.832	0.991	0.623	0.935	0.822	0.997
<i>CSNK1D</i>	17q25	3	0.677	0.991	0.446	0.935	0.687	0.988
<i>CSNK1E</i>	22q13	9	0.855	0.991	0.937	0.937	0.954	0.997
<i>RBX1</i>	22q13	2	0.385	0.991	0.800	0.935	0.321	0.663
Pathway		577	0.578		0.564		0.041	

All analyses were adjusted for age, study area, age at menarche, age at first full-term pregnancy, parity, current use of hormone therapy (only in postmenopausal women), BMI, and alcohol and tobacco consumption.

^a*P* values were calculated using the ARTP method.

^bFDR, false discovery rate-corrected *P* values (correcting for the number of genes tested).

Table 4 P values of stratified and interaction analyses at gene and pathway levels by the duration of nightwork

Genes	Duration of nightwork				Interaction	
	<2 years (990 ca/1067 co)		≥2 years (136 ca/107 co)		P value ^a	FDR P value ^b
	P value ^a	FDR P value ^b	P value ^a	FDR P value ^b		
<i>PER3</i>	0.623	0.995	0.732	0.878	0.821	0.858
<i>RORC</i>	0.988	0.995	0.376	0.878	0.233	0.627
<i>NPAS2</i>	0.219	0.883	0.343	0.878	0.181	0.627
<i>PER2</i>	0.678	0.995	0.850	0.878	0.763	0.854
<i>BHLHE40</i>	0.442	0.995	0.777	0.878	0.594	0.854
<i>CLOCK</i>	0.299	0.883	0.878	0.878	0.953	0.953
<i>SKP1</i>	0.663	0.995	0.353	0.878	0.111	0.627
<i>FBXW11</i>	0.746	0.995	0.837	0.878	0.742	0.854
<i>CUL1</i>	0.287	0.883	0.040	0.644	0.050	0.452
<i>RORB</i>	0.306	0.883	0.760	0.878	0.248	0.627
<i>BTRC</i>	0.393	0.995	0.739	0.878	0.703	0.854
<i>ARNTL</i>	0.987	0.995	0.102	0.644	0.059	0.452
<i>CRY2</i>	0.160	0.883	0.712	0.878	0.572	0.854
<i>BHLHE41</i>	0.995	0.995	0.696	0.878	0.780	0.854
<i>TIMELESS</i>	0.878	0.995	0.605	0.878	0.667	0.854
<i>CRY1</i>	0.188	0.883	0.112	0.644	0.510	0.854
<i>FBXL3</i>	0.958	0.995	0.303	0.878	0.300	0.627
<i>RORA</i>	0.024	0.552	0.650	0.878	0.274	0.627
<i>PER1</i>	0.975	0.995	0.061	0.644	0.024	0.452
<i>NR1D1</i>	0.807	0.995	0.531	0.878	0.423	0.811
<i>CSNK1D</i>	0.633	0.995	0.721	0.878	0.770	0.854
<i>CSNK1E</i>	0.528	0.995	0.572	0.878	0.233	0.627
<i>RBX1</i>	0.307	0.883	0.195	0.878	0.157	0.627
Pathway	0.574		0.621		0.265	

All analyses were adjusted for age, study area, age at menarche, age at first full-term pregnancy, parity, current use of hormone therapy, BMI, and alcohol and tobacco consumption.

^aP values were calculated using the ARTP method.

^bFDR, false discovery rate-corrected P values (correcting for the number of genes tested).

multiple testing), as well as *RORA* and *NPAS2*, although the association with these two genes was not statistically significant after correcting for multiple testing. In the analysis of individual SNPs, the strongest associations with postmenopausal breast cancer were observed for rs11932595 in *CLOCK* and rs1482057 in *RORA*. No such associations were observed in premenopausal women. No clear interaction between genes and nightwork was apparent, as the P values close to or below the limit of statistical significance for *CUL1*, *ARNTL*, and *PER1* became non-significant after correcting for multiple testing.

Biological mechanisms

The circadian genes highlighted in this study are key regulators of the circadian system and may play a role in carcinogenesis. *CLOCK* and *NPAS2* are important components of the mammalian circadian oscillator. The molecular oscillator is based on molecular feedback loops within a positive limb (*CLOCK*, *NPAS2*, and *ARNTL*) and a negative limb (*PER* and *CRY*). The transcription of period (*PER1*, *PER2*,

and *PER3*) and cryptochrome (*CRY1* and *CRY2*) genes is activated by heterodimers composed of *ARNTL* and either of the two related proteins *CLOCK* or *NPAS2*. *ARNTL*–*CLOCK* or *ARNTL*–*NPAS2* heterodimers are transcriptional factors that induce the expression of *PER* and *CRY* genes by binding to their promoters, which in turn will act on the *ARNTL*–*CLOCK*/*NPAS2* complex to repress their own transcription (Lévi et al. 2010, Rana & Mahmood 2010). The *ARNTL*–*CLOCK*/*NPAS2* complex also regulates the transcription of several clock-controlled genes with various biological functions, including some that are relevant for carcinogenesis such as cell cycle control (Hoffman et al. 2010a,b,c).

These heterodimers also induce a secondary regulatory loop activating transcription of retinoic acid-related orphan nuclear receptors, *REV-ERBA* and *RORA*. *RORA* activates transcription of *ARNTL*, whereas *REV-ERBA* represses the transcription process (Lévi et al. 2010, Rana & Mahmood 2010). It has been reported that *RORA* is a potential tumor suppressor gene and its inactivation in different types of cancer may contribute to breast cancer development and progression (Zhu et al. 2006, Xiong et al. 2012).

It has been demonstrated that sex hormones may influence the expression of circadian genes (Nakamura et al. 2001, 2008, Mostafaie et al. 2009) and that circadian genes may play a role in hormone regulation (Urlep & Rozman 2013). Although, the exact mechanisms underlying this interplay are largely unknown, the differences we observed between post- and premenopausal women may be explained in part by differences in circulating levels of estrogen and progesterone in these two groups.

Main genetic effects

Analyses of individual SNPs Two SNPs in *RORA* (rs1482057 and rs12914272) were associated with breast cancer risk in the whole sample and among postmenopausal women. *RORA* has not been investigated in previous studies on circadian genes and breast cancer risk (Zhu et al. 2005, 2008, Hoffman et al. 2010a,b,c, Fu et al. 2012, Monsees et al. 2012, Grundy et al. 2013), with the exception of a study among Norwegian nurses, which did not find any association for variants in these gene (Zienolddiny et al. 2013). However, out of the four SNPs in *RORA* examined in this study, none was in linkage disequilibrium with the two SNPs we highlighted. It should be noted that rs1482057 and rs12914272 are located in the intronic region of *RORA*. It is possible that these SNPs are markers of a causal variant that remained to be discovered. Our findings in *RORA* need to be replicated in independent studies.

Of the 11 SNPs in *CLOCK* included in our analysis, six had non-corrected *P* values <0.01 (rs10462028, rs1801260, rs3792603, rs3805147, rs11932595, and rs7698022) among postmenopausal women. After adjusting for multiple testing, only the association with rs11932595 remained at the limit of statistical significance. This later variant, as well as rs7698022, has been reported to be associated with breast cancer in a previous hospital-based case-control study in Connecticut (Hoffman et al. 2010a), but the ORs were in the other direction. These opposite findings are surprising but cannot be compared easily due to the lack of stratification by menopausal status in the study conducted in Connecticut, where cases were more likely to be postmenopausal than controls ($P < 0.001$). In another hospital-based study in Canada (Grundy et al. 2013), rs11932595 and rs7698022 were unrelated to breast cancer risk, before or after stratification by menopausal status. A possible explanation for these negative findings could be the smaller number of subjects and the decreased statistical power, as compared with our study. We believe that the associations of breast cancer with several SNPs in the *CLOCK* gene, although moderate, are in favor of a real effect of this gene on breast

tumorigenesis. Pooled analyses of studies with appropriate stratification by menopausal status would be helpful for further insights into the role of *CLOCK* in breast cancer.

Previous studies also reported that variants in *NPAS2* (Zhu et al. 2008), *CRY2* (Hoffman et al. 2010b,c), *TIMELESS* (Fu et al. 2012), and *ARNTL* (*BMAL1*) (Zienolddiny et al. 2013) were associated with breast cancer. These associations were study specific and were not replicated in our data.

Analyses at the gene and pathway levels In our study, we used a gene- and pathway-based approach to investigate the overall effect of circadian clock gene polymorphisms on breast cancer risk. This approach is useful to detect the combined effects of genetic polymorphisms that are weakly associated with the disease but may not be detected in single-SNP analyses and may provide additional insights into the mechanisms underlying disease susceptibility. The multiple genetic variants in circadian pathway highlighted in previous studies and in our own study are likely to reflect the genetic complexity of the circadian system and its possible role in tumorigenesis. As the circadian pacemaker requires multiple molecular interactions to generate the circadian rhythms, single-SNP analyses may not be sufficient to analyze the association between circadian genes and breast cancer, which is also a polygenic complex trait.

Our results suggest that genetic variation in the circadian rhythm pathway as a whole is related to breast cancer susceptibility among postmenopausal women. This association was mostly driven by *CLOCK*, *RORA*, and *NPAS2*, the genes with the lowest *P* values. In a previous combined analysis of three GWAS where 285 954 SNPs were classified into 212 KEGG pathways (Li et al. 2011), the variants in the circadian rhythm pathway (characterized by 89 SNPs) were more frequently associated with breast cancer than the genetic variants in other pathways. Evidence for an association with this pathway (P value = 0.02) was mostly due to 13 SNPs in *NPAS2* and *PER1* having *P* values <0.05. Our results that are based on an exhaustive list of genes and a larger number of SNPs to achieve a high coverage of the genetic variation in the pathway provided additional evidence and strong support for a role of circadian genes in breast cancer.

Gene-environment interaction

We conducted interaction analyses to test the hypothesis that circadian genes may influence breast cancer susceptibility particularly in the subgroup of women with a history

of nightwork. Although we did not report any significant interaction between SNPs or circadian genes and nightwork after correcting for multiple testing, the interaction *P* values for *CUL1*, *ARNTL*, and *PER1* were top ranking and at the limit of statistical significance before FDR correction. The finding with *PER1* may be of particular interest, as a variant in *PER1* (rs2735611) was previously associated with an extreme morning preference (Carpen et al. 2006), a condition that was associated with an increased breast cancer risk among Danish military women working in night shifts (Hansen & Lassen 2012). In our study, this variant was tagged by rs2253820 ($r^2 = 1$ in CEU HapMap), but the interaction with nightwork was non-significant. However, our findings are consistent with the Danish study as we both point to a role of *PER1* in women with a history of nightwork for 2 or more years. Further studies with larger number of samples are required to explore these interactions.

In the Nurses' Health study II (NHS2) (Monsees et al. 2012), the *NPAS2* variant allele of rs2305160 (Ala394Thr) was associated with a reduced breast cancer risk in women, which cumulated <2 years of history of shift work. This finding was not replicated in the Norwegian study (Zienolddiny et al. 2013), and an inverse association was found in the Canadian study (Grundy et al. 2013), suggesting that the reported interaction may be a false positive. The Norwegian study (Zienolddiny et al. 2013) also suggested an association of SNPs in *ARNTL*, *RORB*, *CLOCK*, *NPAS2*, *CSNK1E*, and *PER3* with breast cancer among women with a high 'maximum number of consecutive night shifts over the lifetime'. In our study, none of the interactions with nightwork highlighted in the NHS2 and the Norwegian studies were confirmed. However, unlike those studies conducted among nurses, our study conducted in women of general population included a wide variety of nightwork patterns. We also used another exposure metrics for nightwork than the Norwegian study.

Studying interaction between circadian genes and night shift work is complex, because the most appropriate exposure metrics for nightwork is not known and it requires large numbers of subjects. In particular, the duration of nightwork may be a too crude measure of circadian disruption.

Strengths and limits of the study

Our study is the first that included exhaustive selection of SNPs and genes in the circadian rhythm pathway. We included 577 variants in this pathway while previous studies included <200 polymorphisms. We used statistical methods to perform gene- and pathway-level analyses that permitted to gain power by combining weak signals from

SNP-level analysis. We observed suggestive associations of breast cancer with specific gene variants in *RORA*, the genes *CLOCK*, *RORA*, and *NPAS2*, and the circadian clock gene pathway as a whole. No clear interaction between circadian genes and nightwork in relation to breast cancer risk was detected, but the study size was limited. However, to our knowledge, this is thus far the largest study examining the gene–environment interaction between circadian genes and nightwork conducted in a population of European ancestry. It is also the first study examining this interaction using a pathway approach. It was conducted in a carefully designed population-based case–control study, with detailed information on lifestyle and environmental and occupational risk factors. The proportion of nightworkers was similar to that expected among women in France (<http://travail-emploi.gouv.fr/etudes-recherches-statistiques-de,76/etudes-et-recherches,77/publications-dares,98/dares-analyses-dares-indicateurs,102/2011-009-le-travail-de-nuit-des,13024.html>), and all analyses were closely adjusted for potential confounders.

Conclusion

Further studies are warranted to clarify the role of specific combination of SNPs of the circadian rhythm pathway in breast cancer development. Pooling data from studies with information on genes and work time schedule are required.

Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/ERC-14-0121>.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

The CECILE study was supported by grants from the French National Institute of Cancer (INCa), the Fondation de France, the French Agency for Environmental and Occupational Health Safety (ANSES – ex-AFSSSET), the French National Research Agency (ANR), and the League against Cancer.

References

- de Bakker PI, Yelensky R, Péter I, Gabriel SB, Daly MJ & Altshuler D 2005 Efficiency and power in genetic association studies. *Nat Genet* **37** 1217–1223.
- Benjamini Y & Hochberg Y 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* **57** 289–300.

- Carpen JD, von Schantz M, Smits M, Skene DJ & Archer SN 2006 A silent polymorphism in the PER1 gene associates with extreme diurnal preference in humans. *Journal of Human Genetics* **51** 1122–1125. (doi:10.1007/s10038-006-0060-y)
- Cordina-Duverger E, Truong T, Anger A, Sanchez M, Arveux P, Kerbrat P & Guénel P 2013 Risk of breast cancer by type of menopausal hormone therapy: a case–control study among post-menopausal women in France. *PLoS ONE* **8** e78016. (doi:10.1371/journal.pone.0078016)
- Evans JA & Davidson AJ 2013 Health consequences of circadian disruption in humans and animal models. *Progress in Molecular Biology and Translational Science* **119** 283–323. (doi:10.1016/B978-0-12-396971-2.00010-5)
- Foster RG & Kreitzman L 2014 The rhythms of life – what your body clock means to you. *Experimental Physiology* **99** 599–606. (doi:10.1113/expphysiol.2012.071118)
- Fritschi L, Glass DC, Heyworth JS, Aronson K, Girschik J, Boyle T, Grundy A & Erren TC 2011 Hypotheses for mechanisms linking shiftwork and cancer. *Medical Hypotheses* **77** 430–436. (doi:10.1016/j.mehy.2011.06.002)
- Fu L & Lee CC 2003 The circadian clock: pacemaker and tumour suppressor. *Nature Reviews. Cancer* **3** 350–361. (doi:10.1038/nrc1072)
- Fu A, Leaderer D, Zheng T, Hoffman AE, Stevens RG & Zhu Y 2012 Genetic and epigenetic associations of circadian gene TIMELESS and breast cancer risk. *Molecular Carcinogenesis* **51** 923–929. (doi:10.1002/mc.20862)
- Grundy A, Schuetz JM, Lai AS, Janoo-Gilani R, Leach S, Burstyn I, Richardson H, Brooks-Wilson A, Spinelli JJ & Aronson KJ 2013 Shift work, circadian gene variants and risk of breast cancer. *Cancer Epidemiology* **37** 606–612. (doi:10.1016/j.canep.2013.04.006)
- Hansen J & Lassen CF 2012 Nested case–control study of night shift work and breast cancer risk among women in the Danish military. *Occupational and Environmental Medicine* **69** 551–556. (doi:10.1136/oemed-2011-100240)
- Hoffman AE, Yi C-H, Zheng T, Stevens RG, Leaderer D, Zhang Y, Holford TR, Hansen J, Paulson J & Zhu Y 2010a CLOCK in breast tumorigenesis: genetic, epigenetic, and transcriptional profiling analyses. *Cancer Research* **70** 1459–1468. (doi:10.1158/0008-5472.CAN-09-3798)
- Hoffman AE, Zheng T, Ba Y, Stevens RG, Yi C-H, Leaderer D & Zhu Y 2010b Phenotypic effects of the circadian gene Cryptochrome 2 on cancer-related pathways. *BMC Cancer* **10** 110. (doi:10.1186/1471-2407-10-110)
- Hoffman AE, Zheng T, Yi C-H, Stevens RG, Ba Y, Zhang Y, Leaderer D, Holford T, Hansen J & Zhu Y 2010c The core circadian gene Cryptochrome 2 influences breast cancer risk, possibly by mediating hormone signaling. *Cancer Prevention Research* **3** 539–548. (doi:10.1158/1940-6207.CAPR-09-0127)
- Ko CH & Takahashi JS 2006 Molecular components of the mammalian circadian clock. *Human Molecular Genetics* **15** R271–R277. (doi:10.1093/hmg/ddl207)
- Lévi F, Okyar A, Dulong S, Innominato PF & Clairambault J 2010 Circadian timing in cancer treatments. *Annual Review of Pharmacology and Toxicology* **50** 377–421. (doi:10.1146/annurev.pharmtox.48.113006.094626)
- Li J, Humphreys K, Heikkinen T, Aittomäki K, Blomqvist C, Pharoah PDP, Dunning AM, Ahmed S, Hooning MJ, Martens JW et al. 2011 A combined analysis of genome-wide association studies in breast cancer. *Breast Cancer Research and Treatment* **126** 717–727. (doi:10.1007/s10549-010-1172-9)
- Lie J-AS, Kjuus H, Zienolddiny S, Haugen A, Stevens RG & Kjørheim K 2011 Night work and breast cancer risk among Norwegian nurses: assessment by different exposure metrics. *American Journal of Epidemiology* **173** 1272–1279. (doi:10.1093/aje/kwr014)
- Menegaux F, Truong T, Anger A, Cordina-Duverger E, Lamkarkach F, Arveux P, Kerbrat P, Févotte J & Guénel P 2013 Night work and breast cancer: a population-based case–control study in France (the CECILE study). *International Journal of Cancer. Journal International du Cancer* **132** 924–931. (doi:10.1002/ijc.27669)
- Monsees GM, Kraft P, Hankinson SE, Hunter DJ & Schernhammer ES 2012 Circadian genes and breast cancer susceptibility in rotating shift workers. *International Journal of Cancer* **131** 2547–2552. (doi:10.1002/ijc.27564)
- Mostafaie N, Kállay E, Sauerzapf E, Bonner E, Kriwanek S, Cross HS, Huber KR & Krugluger W 2009 Correlated downregulation of estrogen receptor beta and the circadian clock gene Per1 in human colorectal cancer. *Molecular Carcinogenesis* **48** 642–647. (doi:10.1002/mc.20510)
- Nakamura TJ, Shinohara K, Funabashi T & Kimura F 2001 Effect of estrogen on the expression of Cry1 and Cry2 mRNAs in the suprachiasmatic nucleus of female rats. *Neuroscience Research* **41** 251–255. (doi:10.1016/S0168-0102(01)00285-1)
- Nakamura TJ, Sellix MT, Menaker M & Block GD 2008 Estrogen directly modulates circadian rhythms of PER2 expression in the uterus. *American Journal of Physiology. Endocrinology and Metabolism* **295** E1025–E1031. (doi:10.1152/ajpendo.90392.2008)
- Pedrazzoli M, Secolin R, Esteves LOB, Pereira DS, Koike BDV, Louzada FM, Lopes-Cendes I & Tufik S 2010 Interactions of polymorphisms in different clock genes associated with circadian phenotypes in humans. *Genetics and Molecular Biology* **33** 627–632. (doi:10.1590/S1415-47572010005000092)
- Rana S & Mahmood S 2010 Circadian rhythm and its role in malignancy. *Journal of Circadian Rhythms* **8** 3. (doi:10.1186/1740-3391-8-3)
- Savvidis C & Koutsilieris M 2012 Circadian rhythm disruption in cancer biology. *Molecular Medicine* **18** 1249–1260. (doi:10.2119/molmed.2012.00077)
- Stevens RG, Brainard GC, Blask DE, Lockley SW & Motta ME 2014 Breast cancer and circadian disruption from electric lighting in the modern world. *CA: A Cancer Journal for Clinicians* **64** 207–218. (doi:10.3322/caac.21218)
- Straif K, Baan R, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, Altieri A, Benbrahim-Tallaa L & Coglianò V 2007 Carcinogenicity of shift-work, painting, and fire-fighting. *Lancet Oncology* **8** 1065–1066. (doi:10.1016/S1470-2045(07)70373-X)
- Urlep Z & Rozman D 2013 The interplay between circadian system, cholesterol synthesis, and steroidogenesis affects various aspects of female reproduction. *Frontiers in Endocrinology* **4** 111. (doi:10.3389/fendo.2013.00111)
- Xiong G, Wang C, Evers BM, Zhou BP & Xu R 2012 ROR α suppresses breast tumor invasion by inducing SEMA3F expression. *Cancer Research* **72** 1728–1739. (doi:10.1158/0008-5472.CAN-11-2762)
- Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, Kraft P & Chatterjee N 2009 Pathway analysis by adaptive combination of P-values. *Genetic Epidemiology* **33** 700–709. (doi:10.1002/gepi.20422)
- Zhu Y, Brown HN, Zhang Y, Stevens RG & Zheng T 2005 Period3 structural variation: a circadian biomarker associated with breast cancer in young women. *Cancer Epidemiology, Biomarkers & Prevention* **14** 268–270.
- Zhu Y, McAvoy S, Kuhn R & Smith DI 2006 RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene* **25** 2901–8. (doi:10.1038/sj.onc.1209314)
- Zhu Y, Stevens RG, Leaderer D, Hoffman A, Holford T, Zhang Y, Brown HN & Zheng T 2008 Non-synonymous polymorphisms in the circadian gene NPAS2 and breast cancer risk. *Breast Cancer Research and Treatment* **107** 421–5. (doi:10.1007/s10549-007-9565-0)
- Zienolddiny S, Haugen A, Lie J-AS, Kjuus H, Anmarkrud KH & Kjørheim K 2013 Analysis of polymorphisms in the circadian-related genes and breast cancer risk in Norwegian nurses working night shifts. *Breast Cancer Research* **15** R53. (doi:10.1186/bcr3445)

Received in final form 19 May 2014

Accepted 9 June 2014

Made available online as an Accepted Preprint

11 June 2014