

1 **Inherited variation in circadian rhythm genes and risks of prostate cancer and three other**
2 **cancer sites in combined cancer consortia**

3 Fangyi Gu ^{*†1,2}, Han Zhang ^{*1}, Paula L. Hyland¹, Sonja Berndt¹, Susan M. Gapstur³, William
4 Wheeler⁴, the ELLIPSE consortium[‡], Christopher I. Amos⁵, Stephane Bezieau⁶, Heike
5 Bickeböllner⁷, Hermann Brenner⁸⁻¹⁰, Paul Brennan¹¹, Jenny Chang-Claude¹², David V Conti¹³,
6 Jennifer Ann Doherty¹⁴, Stephen B Gruber¹³, Tabitha A Harrison¹⁵, Richard B Hayes¹⁶, Michael
7 Hoffmeister⁸, Richard S Houlston¹⁷, Rayjean J. Hung¹⁸, Mark A. Jenkins¹⁹, Peter Kraft²⁰, Kate
8 Lawrenson²¹, James McKay¹¹, Sarah Markt²⁰, Lorelei Mucci²⁰, Catherine M. Phelan²², Conghui
9 Qu¹⁵, Angela Risch²³⁻²⁶, Mary Anne Rossing¹⁵, H.-Erich Wichmann²⁷⁻²⁹, Jianxin Shi¹, Eva
10 Schernhammer^{20,30,31}, Kai Yu¹, Maria Teresa Landi¹, Neil E. Caporaso¹

11 ***Co-first authors; †Corresponding author**

12 **‡The list of authors in the supplementary materials**

13 ¹Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA

14 ²Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY,
15 USA

16 ³Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA

17 ⁴Information Management Services, Inc., Rockville, MD, USA

18 ⁵Geisel School of Medicine at Dartmouth, Lebanon, NH, USA

19 ⁶Service de Génétique Médicale, CHU Nantes, Nantes, France

20 ⁷Department of Genetic Epidemiology, University Medical Center Göttingen, Göttingen,
21 Germany

22 ⁸Division of Clinical Epidemiology and Aging Research, German Cancer Research Center
23 (DKFZ), Heidelberg, Germany

24 ⁹Division of Preventive Oncology, National Center for Tumor Diseases (NCT) and German
25 Cancer Research Center (DKFZ), Heidelberg, Germany

26 ¹⁰German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg,
27 Germany

28 ¹¹International Agency for Research on Cancer, Lyon, France

1 ¹²Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg,
2 Germany

3 ¹³Keck School of Medicine, University of South California, Los Angeles, CA, USA

4 ¹⁴Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA

5 ¹⁵Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA,
6 USA

7 ¹⁶New York University School of Medicine, Department of Population Health, New York, NY,
8 USA

9 ¹⁷Division of Genetics and Epidemiology, the Institute of Cancer Research, Sutton, Surrey, UK

10 ¹⁸Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Canada

11 ¹⁹Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health,
12 University of Melbourne, Parkville, Victoria, Australia

13 ²⁰Department of Epidemiology, Harvard T.H Chan School of Public Health, Boston, MA, USA

14 ²¹Cedars-Sinai Medical Center, Los Angeles, CA, USA

15 ²²Department of Cancer Epidemiology, Population Sciences Division, Moffitt Cancer Center,
16 Tampa, FL, USA

17 ²³Division of Molecular Biology, University of Salzburg, Austria

18 ²⁴Cancer Cluster Salzburg, Austria

19 ²⁵Translational Lung Research Center, Heidelberg, Germany within the German Center for Lung
20 Research (DZL), Germany

21 ²⁶Division of Epigenomics and Cancer Risk Factors, DKFZ German Cancer Research Center,
22 Heidelberg, Germany

23 ²⁷Institute of Medical Informatics, Biometry and Epidemiology, University of Munich, Germany

24 ²⁸Helmholtz Center Munich, Institute of Epidemiology II, Germany

25

26 ²⁹Institute of Medical Statistics and Epidemiology, Technical University Munich, Germany

27

28 ³⁰Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical
29 School, Boston, MA, USA

1 ³¹Department of Epidemiology, Medical University of Vienna, Vienna, Austria

2 **Novelty & Impact:** We found a significant association of circadian rhythm and melatonin
3 pathway genes with prostate cancer risk, at the gene and pathway level, after taking multiple
4 comparisons into account. The sample size is the largest to our knowledge, with a further
5 replication in an independent data. This study provides evidence in support of a role for circadian
6 rhythm and melatonin pathways in prostate carcinogenesis.

7 Word count: 267 in abstract; 2701 in text.

8

1 **ABSTRACT**

2
3 Circadian disruption has been linked to carcinogenesis in animal models, but the evidence in
4 humans is inconclusive. Genetic variation in circadian rhythm genes provides a tool to
5 investigate such associations. We examined associations of genetic variation in nine core
6 circadian rhythm genes and six melatonin pathway genes with risk of colorectal, lung, ovarian
7 and prostate cancers using data from the Genetic Associations and Mechanisms in Oncology
8 (GAME-ON) network. The major results for prostate cancer were replicated in the Prostate,
9 Lung, Colorectal and Ovarian (PLCO) cancer screening trial, and for colorectal cancer in the
10 Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO). The total number of
11 cancer cases and controls was 15,838/18,159 for colorectal, 14,818/14,227 for prostate,
12 12,537/17,285 for lung and 4,369/9,123 for ovary. For each cancer site, we conducted gene-
13 based and pathway-based analyses by applying the summary-based Adaptive Rank Truncated
14 Product method (sARTP) on the summary association statistics for each SNP within the
15 candidate gene regions. Aggregate genetic variation in circadian rhythm and melatonin pathways
16 were significantly associated with the risk of prostate cancer in data combining GAME-ON and
17 PLCO, after Bonferroni correction ($P_{\text{pathway}} < 0.00625$). The two most significant genes were
18 *NPAS2* ($P_{\text{gene}} = 0.0062$) and *AANAT* ($P_{\text{gene}} = 0.00078$); the latter being significant after Bonferroni
19 correction. For colorectal cancer, we observed a suggestive association with the circadian rhythm
20 pathway in GAME-ON ($P_{\text{pathway}} = 0.021$); this association was not confirmed in GECCO
21 ($P_{\text{pathway}} = 0.76$) or the combined data ($P_{\text{pathway}} = 0.17$). No association was observed for ovarian and
22 lung cancer. These findings support a potential role for circadian rhythm and melatonin pathways
23 in prostate carcinogenesis. Further functional studies are needed to better understand the
24 underlying biologic mechanisms.

1 **Keywords:** circadian rhythm, melatonin, prostate cancer, cancer

2

1 INTRODUCTION

2 Circadian rhythm is driven by an internal biological clock, which enables humans to sustain an
3 approximate 24-hour cycle of biological processes¹, and regulates diverse cancer-related
4 biological functions such as metabolism, immune regulation, DNA repair and cell cycle control².
5 Disruption of circadian rhythm has been linked to carcinogenesis at the system, cell and
6 molecular levels². Based on sufficient evidence in experimental animals for the carcinogenicity
7 of light exposure during the biological night, and limited epidemiological studies showing
8 increased risk of breast cancer among female nightshift workers and flight attendants employed
9 at least ten years, shift work with disrupted circadian rhythm has been categorized as a probable
10 carcinogen to humans by the International Agency for Research on Cancer³. However, evidence
11 for cancers other than breast is limited. Increased cancer risks in other organs have been
12 observed in mouse models with ablated circadian rhythm genes, such as the blood⁴, liver⁴, ovary
13 ⁴, intestine⁵, colon⁵ and skin⁶, possibly due to constitutively elevated cell proliferation⁶,
14 impaired DNA repair⁷ and apoptosis⁸, and inefficient immune response^{9,10}. There is growing
15 evidence from epidemiologic studies that other types of cancers including prostate¹¹⁻¹⁴, colon¹⁵
16 and non-Hodgkin lymphoma¹⁶ also may be associated with rotating and night shift work.

17 A few candidate gene studies have examined associations between genes involved in
18 circadian processes and several cancer sites¹⁷⁻²⁹, especially breast^{21, 24-26, 29}. In this study, we
19 examined associations of the core genes involved in the circadian rhythm and melatonin
20 pathways with the risk of prostate, colorectal, lung and ovarian cancer in population of European
21 descent, taking advantage of the large study populations from the Genetic Associations and
22 Mechanisms in Oncology (GAME-ON) GWAS consortia. We conducted a pathway-level

1 analysis, aggregating association evidence across multiple genes. Potentially interesting findings
2 were further replicated in independent populations of European descent.

3 4 **METHODS**

5 **Study populations**

6 Our initial analyses used data from 20 GWAS studies on four common cancer sites within the
7 National Cancer Institute GAME-ON Network (<http://epi.grants.cancer.gov/gameon/>)³⁰,
8 including 12,537 lung cancer cases and 17,285 controls from the Transdisciplinary Research for
9 Cancer of Lung (TRICL) consortium; 5,100 colorectal cases and 4,831 controls from the
10 ColoRectal Transdisciplinary Study (CORECT); 10,218 prostate cancer cases and 11,286
11 controls from the Elucidating Loci in Prostate Cancer Susceptibility (ELLIPSE) consortium; as
12 well as 4,369 ovarian cancer cases and 9,123 controls from the Follow-up of Ovarian Cancer
13 Genetic Association and Interaction Studies (FOCI) (Table 1). For colorectal and prostate cancer,
14 potentially interesting findings were carried forward and replicated in additional independent
15 data: 10,738 cases and 13,328 controls from the Genetics and Epidemiology of Colorectal
16 Cancer Consortium for colorectal cancer (GECCO)³¹; 4,600 cases and 2,940 controls from the
17 Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial for prostate cancer³². All
18 participants were of European descent, and most of the studies were conducted using Illumina
19 genotyping platforms (Table 1). Details of the genotyping and quality control steps were
20 published previously³⁰⁻³². All participating studies obtained approval from the institutional ethics
21 review board, and informed consents were obtained from each study participant by the individual
22 study coordinating center.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Candidate genes

For the circadian rhythm pathway, we included nine well-established core circadian rhythm genes that generate the mammalian circadian rhythm³³ and were selected for a previous cancer study to represent the circadian rhythm pathway²⁴: *CLOCK* and its paralogue *NPAS2* (neuronal PAS domain protein 2); *ARNTL* (aryl hydrocarbon receptor nuclear translocator-like; a.k.a. *Bmal1*); *CKIε* (casein kinase I ε; a.k.a. *CSNK1E*); Cryptochrome 1 (*CRY1*); *CRY2*; and three Period homologs (*PER1*, *PER2* and *PER3*).

Due to a close integration of melatonin to the circadian system, we also included four genes involved in melatonin biosynthesis (http://www.kegg.jp/kegg-bin/show_module?M00037)³⁴ and two melatonin receptor genes: arylalkylamine N-acetyltransferase (*AANAT*, a gene encoding the rate limiting enzyme in the melatonin biosynthesis), *TPH1* (tryptophan hydroxylase 1), *TPH2*, and *DDC* (aromatic-L-amino-acid decarboxylase); *MTNR1α* (melatonin receptor 1α), and *MTNR1β*. Another gene involved in the melatonin biosynthesis, *ASMT* (Acetylserotonin O-methyltransferase) was not included because we have no access to the data of the x chromosome where this gene is located.

Statistical analyses

The analytical methods of original studies and the cancer-specific results have been described previously^{31, 32, 35-38} and summarized in Table 1. Briefly each original study provided log odds ratios and standard errors on each SNP and each cancer risk, mostly adjusting for age, principal

1 components (PCs), and sex (if applicable). For each cancer site, fixed-effect meta-analyses were
2 conducted to combine summary association statistics of participating studies by the cohort
3 consortium. The genotypes were imputed based on data of European populations from the 1000
4 Genomes Project (March 2012 reference panel)³⁹, using either MaCH⁴⁰ or IMPUTE⁴¹. We
5 extracted both the genotyped and imputed SNPs of the genetic regions from 20 kb upstream to
6 10 kb downstream of each candidate gene.

7 We conducted gene- and pathway-based meta-analyses using the summary based
8 adaptive rank truncated product (sARTP) method, which combines SNP-level association
9 evidence across SNPs in a gene or a pathway⁴². The sARTP method automatically adjusts for
10 the size of the gene (i.e., number of SNPs in a gene) and the size of the pathway (i.e., number of
11 genes in a pathway) through a resampling procedure. The final gene- and pathway-level p-values
12 were estimated from the resampled null distribution through one million resampling steps. The
13 sARTP method accounts for the linkage disequilibrium (LD) between SNPs to maintain proper
14 type I error. The LDs between SNPs were estimated from the 503 European subjects (CEU, TSI,
15 FIN, GBR, IBS) in the 1000 Genome Project (phase 3, v5, 2013/05/02)³⁹. We excluded SNPs
16 with MAF < 5% and applied LD filtering to highly correlated SNP pairs ($r^2 > 0.95$). We also
17 conducted a sensitivity analysis using a more stringent threshold for LD pruning ($r^2 > 0.8$).

18 For prostate and colorectal cancer that have pathway p-values less than 0.05, we
19 replicated our findings in PLCO and GECCO. We also repeated the gene- and pathway-based
20 analyses on data combining the initial and replication studies.

21 To eliminate the impact of potential systematic biases in SNP-level association, we
22 adjusted for the genomic control inflation factor ($\lambda=1.015$) for data from the CORECT^{37,42}.
23 The genomic control inflation factors for GECCO, ELLIPSE, PLCO, TRICL and FOCI were

1 close to or smaller than 1.0, thus were not adjusted in our analyses. To take potential false-
2 positives from multiple-comparisons into account (two pathways, or 15 genes) for each of the
3 four cancer sites, pathways with p-value < 0.00625 ($0.05 / (2 \times 4)$) and genes with p-value $<$
4 0.00083 ($0.05 / (15 \times 4)$) were considered significant.

5 For prostate cancer, where we found significant associations with genetic variations of
6 circadian and melatonin pathways after the Bonferroni correction, secondary analyses for
7 aggressive prostate cancer were conducted at the gene and pathway level, using data combining
8 six studies of ELLIPSE and PLCO (4,446 cases and 12,724 controls). For the SNPs with the
9 smallest p-values in the genes with $P_{\text{gene}} \leq 0.05$ on the risk of overall prostate cancer, we also
10 checked their SNP associations with aggressive prostate cancer.

11

12 RESULTS

13 We found suggestive associations between genetic variation in both circadian rhythm and
14 melatonin pathways and prostate cancer risk based on data of GAME-ON, with ($P_{\text{pathway}}=0.014$
15 and 0.024 , respectively (Table 2). These associations were not statistically significant in PLCO
16 alone ($P_{\text{pathway}}=0.28$ and 0.21), but were enhanced in the combined data of GAME-ON and
17 PLCO ($P_{\text{pathway}}=0.0016$ and 0.0060) (Table 2), both being significant after Bonferroni correction.
18 *NPAS2* in the circadian rhythm pathway ($P_{\text{gene}}=0.0062$) and *AANAT* ($P_{\text{gene}}=0.00078$) in the
19 melatonin pathway contributed the most to the association with the risk of prostate cancer, with
20 *AANAT* survived Bonferroni correction (Table 3). Other genes with the gene-level p-values at
21 borderline significance were *CLOCK* ($P_{\text{gene}}=0.021$), *CRY2* ($P_{\text{gene}}=0.043$), *DDC* ($P_{\text{gene}}=0.050$),
22 *PER2* ($P_{\text{gene}}=0.060$), and *PER1* ($P_{\text{gene}}=0.063$) (Table 3). A sensitivity analysis with more

1 stringent threshold in LD pruning ($r^2 > 0.8$) produced consistent pathway-level and gene-level
2 results (data not shown). SNPs with p-value < 0.01 in *NPAS2* and *AANAT* are presented in Table
3 4.

4 With a much smaller number of aggressive prostate cancer cases (4,446 cases, 12,724
5 controls), we did not observe significant association of aggressive prostate cancer with either
6 pathway ($P_{\text{pathway}}=0.29$ and 0.66), but we observed a suggestive association with *PER3*
7 ($P_{\text{gene}}=0.03$) (Supplementary Table 2). For SNPs that have the smallest p-values in genes
8 *CLOCK*, *CRY2*, *NPAS2*, *AANAT*, and *DDC* ($P_{\text{gene}} \leq 0.05$ with overall prostate cancer), the log
9 odds ratios (β) estimated for overall and aggressive prostate cancer are comparable and have the
10 same direction (Supplementary Table 3).

11 For colorectal cancer (Table 2), we observed a suggestive association with circadian
12 rhythm pathway in GAME-ON ($P_{\text{pathway}}=0.021$), but not in GECCO ($P_{\text{pathway}}=0.76$) or in the
13 combined data ($P_{\text{pathway}}=0.17$) (Supplementary Table 4). No association was observed for ovarian
14 cancer and lung cancer (Table 2, Supplementary Table 5).

15

1 **DISCUSSION**

2 We found common genetic variations in the circadian rhythm and melatonin pathways were
3 associated with prostate cancer risk in the population of European descent. These associations
4 were initially identified in the GAME-ON consortium, and further confirmed in the data
5 combining the GAME-ON and PLCO studies. Our findings suggest that the circadian rhythm
6 and melatonin pathways may be involved in prostate carcinogenesis.

7 Circadian disruption has been suggested as a prostate cancer risk factor based on
8 epidemiological observation of increased prostate cancer risks among shift workers¹¹⁻¹⁴, and
9 countries with more light exposure at night⁴³. In support of this hypothesis, three genetic
10 epidemiology studies found suggestive associations between SNPs in core circadian genes and
11 prostate cancer^{19, 23, 27} or aggressive prostate cancer²³ in Caucasian^{23, 27} and Asian¹⁹ populations,
12 although these studies had limited power (sample sizes < 2600) to adjust for multiple
13 comparisons. By taking advantage of the large study population from cancer consortia and using
14 a novel analytical tool, our study provided further evidence that the circadian rhythm and
15 melatonin pathways may be involved in prostate carcinogenesis in humans.

16 Although multiple genes are likely to contribute to pathway association signals, the most
17 significant genes were *NPAS2* and *AANAT*. Previous functional studies suggest that *NPAS2* plays
18 an important role in DNA damage response, cell cycle control and apoptosis by activating
19 diverse downstream genes^{44, 45}, consistent with a role as a tumor suppressor. In line with our
20 finding, the Thr allele of rs23051560 ($P=7.5 \times 10^{-4}$), a non-synonymous SNP (Ala394Thr) in the
21 *NPAS2*, has been suggestively associated with lower risks of breast cancer²⁸, prostate cancer¹⁹,
22 and NHL⁴⁶, three tumors that have been linked with circadian disruption in epidemiologic studies.

1 This SNP has also been suggested to modify the association of night shift work and breast cancer
2 risk, with Thr carriers more vulnerable to shift work effect²⁴. AANAT (aka., serotonin N-
3 acetyltransferase) is the rate limiting and originating enzyme for melatonin synthesis, through
4 which the suprachiasmatic nucleus via a sympathetic multisynaptic pathway regulates rhythmic
5 melatonin synthesis⁴⁷. Melatonin acts as a chronobiotic molecule, optimizing phase
6 relationships between oscillators in both central nervous system and peripheral organs,
7 reinforcing circadian rhythms of body functions, and entraining body rhythms to the
8 environmental light phase^{48, 49}.

9 A mechanism linking the circadian system, melatonin and prostate cancer may operate
10 through the neuroendocrine gonadal axis. The pineal gland and melatonin have a role in the
11 inhibition of the neuroendocrine gonadal axis⁵⁰; while sex hormones, such as androgen, are
12 essential on prostate development. Androgen has been a prostate cancer inducer in animals⁵¹,
13 and associated with increased prostate cancer risk in humans^{52, 53}. Therefore, it is possible that
14 an increase in androgen, subsequent to disrupted circadian rhythm and/or suppressed melatonin
15⁵⁴, may contribute to prostate carcinogenesis. Alternatively, melatonin may have a direct anti-
16 tumor effect, by controlling the p53 pathway, or its antimetabolic, antioxidant and immune-
17 modulatory activities¹. Both in vitro and in vivo studies provide evidence that melatonin inhibits
18 prostate tumor growth^{55, 56}, whereas melatonin suppression in rats increases tumor growth in a
19 dose-dependent manner⁵⁰. In agreement with the melatonin hypothesis, lower urinary 6-
20 sulfatoxymelatonin has been associated with an increased risk of advanced prostate cancer in a
21 prospective study⁵⁷.

1 Apart from mechanisms related to melatonin, the circadian clock may control cell
2 proliferation and apoptosis through regulating the expression of genes involved in these
3 processes at the transcription or translation level, such as *c-Myc* and *Mdm2*, *Trp53* and *Gadd45*,
4 *cyclins* etc.²

5 We did not find any significant association for the risk of aggressive prostate cancer at
6 the gene or pathway level. Given a much smaller number of aggressive prostate cancer cases,
7 and the fact that genetic effects are generally small on cancer risk, the statistical power of gene-
8 and pathway-based analyses was limited. However, we observed a suggestive association with
9 *PER3* ($P_{\text{gene}}=0.03$); a SNP (rs1012477) of this gene has been associated with prostate cancer
10 aggressiveness in a previous report²⁷. For SNPs with the smallest p-values associated with
11 overall prostate cancer within *CLOCK*, *CRY2*, *NPAS2*, *AANAT*, and *DDC*, the estimated effect
12 sizes for the risk of overall and aggressive prostate cancer are comparable and have the same
13 direction. Given the poor prognosis and public health impact of aggressive prostate cancer, more
14 focused study is needed for the role of circadian rhythm genes and prostate cancer
15 aggressiveness.

16 Our study did not find associations in the circadian rhythm or melatonin pathway genes
17 with colorectal, lung or ovarian cancer. Several important factors need to be considered before
18 concluding that circadian rhythm has no effect on these cancer sites. First, gene functions differ
19 by organs and although we studied the core genes in each pathway, there might be other critical
20 circadian-related genes missed in this study. *ROR α* , for example, suggested as an important
21 regulator for homeostasis in intestinal epithelium⁵⁸, as well as newly identified circadian genes⁵⁹
22 are worthwhile to be evaluated in the future. Second, the statistical power of gene- and pathway-

1 based analyses for studying ovarian cancer may be limited by small sample size compared with
2 other cancer sites considered in this paper. Third, for lung and colorectal cancer, where
3 environmental and life style risk factors play a dominant role, the contribution of disrupted
4 circadian rhythm might be small and/or may be indirectly associated with cancer through
5 modifying the toxicity of environmental carcinogens⁶⁰, or altering the DNA damage response^{6,7}.
6 Therefore, incorporating data on environmental carcinogens and measures of toxicity into the
7 study of circadian rhythm and cancer may be important. Fourth, although genetic variation does
8 not suffer from confounding bias by other life style factors, it may have a smaller impact on
9 circadian rhythm disruption than light exposure at night and night shift work. Therefore, future
10 studies of both environmental or life style inducers of circadian disruption coupled with
11 mechanistic or genetic marker studies in circadian rhythm pathways are needed.

12 In this study, like other candidate pathway-based analyses⁶¹, we assigned SNPs to each
13 of the circadian genes based on genomic location. Approaches that assign SNPs to a gene based
14 on functionality such as a genetic influence on gene expression or expression quantitative risk
15 loci (eQTL) might reveal more signals, but this type of approach relies heavily on the known
16 eQTL function of the SNPs in the tissue of interest and, in fact, the eQTL effects on gene
17 expression are typically tissue-specific⁶². We attempted to evaluate the involvement of the top
18 prostate cancer risk SNPs of *AANAT* and *NPAS2* as functional eQTLs using RNA-seq and SNP
19 data from ten normal brain tissues (GTEx). We observed modest eQTL effects on *AANAT* and
20 *NPAS2* mRNA levels by the top risk SNPs, but no risk eQTL survived correction for multiple
21 comparisons (data not shown). Importantly, published data suggest that the target tissue for
22 melatonin synthesis is the pineal gland, while for circadian rhythm it is the suprachiasmatic
23 nucleus (SCN)¹. RNA-seq data for these normal brain tissues are not available in GTEx or to

1 our knowledge from any other publically available database. Thus, whether the observed prostate
2 cancer risk SNPs of *AANAT* and *NPAS2* circadian genes are functional eQTLs, and whether the
3 changes in mRNA levels in the pineal gland and SCN are associated with prostate cancer
4 susceptibility remains to be determined.

5
6 Our study has many strengths. Using genetic markers to examine circadian hypotheses
7 minimizes the bias due to potential confounders, and therefore is a valuable complement to
8 traditional epidemiologic studies (e.g., in night shift workers). We used an analytical tool that
9 combines signals across SNPs within genes and pathways, and therefore found significant results
10 that would have been detectable by single SNP analysis. To our knowledge, the sample sizes in
11 our study are the largest to date for colorectal, lung, and prostate cancer. The data quality of the
12 included GWAS studies is well established. To control potential false positive findings, we
13 adjusted for multiple comparisons, and replicated our findings in independent data.

14 In summary, our study suggests that common genetic variation in and around circadian
15 rhythm and melatonin pathways may be involved in human prostate carcinogenesis, in support of
16 circadian disruption as a potential human carcinogen.

1 **Acknowledgement**

2 We thank Dr. Andrew Bergen and Shailesh Kumar (NIH/NHLBI) for the discussion on
3 functional annotation and circadian rhythm. We recognize the following contributors from
4 CORECT: Stephanie L. Schmit, Fredrick R. Schumacher, Christopher K. Edlund, Gad Rennert,
5 Eric Jacobs, Peter T. Campbell, John L. Hopper, Daniel D. Buchanan, Li Li, Michael Woods,
6 Graham Giles. Other contributors from GECCO are listed in the supplementary materials.

7 **Funding:**

8 TRICL (Transdisciplinary Research for Cancer of Lung) and International Lung Cancer
9 Consortium (ILCCO): National Institute of Health U19 CA148127-01 (PI: Amos),
10 1U19CA148127-02 (PI: Bickeböllner), Canadian Cancer Society Research Institute (no. 020214,
11 PI: Hung).

12 DRIVE (Discovery, Biology, and Risk of Inherited Variants in Breast Cancer): National Institute
13 of Health U19 CA148065.

14 CORECT (ColoRectal Transdisciplinary Study): National Institute of Health U19 CA148107;
15 R01 CA81488, P30 CA014089.

16 ELLIPSE (Elucidating Loci in Prostate Cancer Susceptibility): This work was support by the
17 GAME-ON U19 initiative for prostate cancer (ELLIPSE), U19 CA148537.

18 FOCI (Follow-up of Ovarian Cancer Genetic Association and Interaction Studies): National
19 Institutes of Health U19 CA148112-01 (PI: Sellers) and R01-CA149429 (Phelan).

20 GECCO (Genetics and Epidemiology of Colorectal Cancer Consortium): National Cancer
21 Institute, National Institutes of Health, US Department of Health and Human Services (U01
22 CA137088; R01 CA059045). ASTERISK: a Hospital Clinical Research Program (PHRC) and
23 supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises
24 Françaises dans la Lutte contre le Cancer (GEFLUC), the Association Anne de Bretagne
25 Génétique and the Ligue Régionale Contre le Cancer (LRCC). DACHS: German Research
26 Council (Deutsche Forschungsgemeinschaft, BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, and CH
27 117/1-1), and the German Federal Ministry of Education and Research (01KH0404 and
28 01ER0814). DALs: National Institutes of Health (R01 CA48998 to MLS); HPFS is supported by
29 the National Institutes of Health (P01 CA 055075, UM1 CA167552, R01 137178, R01 CA
30 151993, and P50 CA 127003), NHS by the National Institutes of Health (R01 CA137178, P01
31 CA 087969, R01 CA151993, and P50 CA 127003), and PHS by the National Institutes of Health
32 (R01 CA042182). OFCCR: National Institutes of Health, through funding allocated to the
33 Ontario Registry for Studies of Familial Colorectal Cancer (U01 CA074783); see CFR section.
34 Additional funding toward genetic analyses of OFCCR includes the Ontario Research Fund, the
35 Canadian Institutes of Health Research, and the Ontario Institute for Cancer Research, through

1 generous support from the Ontario Ministry of Research and Innovation. PLCO: Intramural
2 Research Program of the Division of Cancer Epidemiology and Genetics and supported by
3 contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS.
4 Additionally, a subset of control samples were genotyped as part of the Cancer Genetic Markers
5 of Susceptibility (CGEMS) Prostate Cancer GWAS (Yeager M, et al. *Nat Genet.*
6 2007;39(5):645–649), Colon CGEMS pancreatic cancer scan (PanScan) (Amundadottir L, et al.
7 *Nat Genet.* 2009;41(9):986–990 and Petersen GM, et al. *Nat Genet.* 2010;42(3):224–228), and
8 the Lung Cancer and Smoking study. The prostate and PanScan study datasets were accessed
9 with appropriate approval through the dbGaP online resource (<http://cgems.cancer.gov/data/>)
10 accession numbers phs000207.v1.p1 and phs000206.v3.p2, respectively, and the lung datasets
11 were accessed from the dbGaP website (<http://www.ncbi.nlm.nih.gov/gap>) through accession
12 number phs000093.v2.p2. Funding for the Lung Cancer and Smoking study was provided by
13 National Institutes of Health (NIH), Genes, Environment, and Health Initiative (GEI) Z01 CP
14 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438. For the lung study, the GENEVA
15 Coordinating Center provided assistance with genotype cleaning and general study coordination,
16 and the Johns Hopkins University Center for Inherited Disease Research conducted genotyping.
17 PMH: National Institutes of Health (R01 CA076366 to PA Newcomb). VITAL: National
18 Institutes of Health (K05 CA154337). WHI: The WHI program is funded by the National Heart,
19 Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human
20 Services through contracts HHSN268201100046C, HHSN268201100001C,
21 HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and
22 HHSN271201100004C.

23 CAPS GWAS study was supported by the Swedish Cancer Foundation (grant no 09-0677, 11-
24 484, 12-823), the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linneus
25 Centre (Contract ID 70867902) financed by the Swedish Research Council, Swedish Research
26 Council (grant no K2010-70X-20430-04-3, 2014-2269).

27
28 CRUK GWAS: This work was supported by the Canadian Institutes of Health Research,
29 European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-
30 F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C5047/A3354,
31 C5047/A10692, C16913/A6135, and The National Institute of Health (NIH) Cancer Post-Cancer
32 GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative). We would also
33 like to thank the following for funding support: The Institute of Cancer Research and The
34 Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign
35 UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network
36 UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR
37 funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The
38 Royal Marsden NHS Foundation Trust. The Prostate Cancer Program of Cancer Council Victoria
39 also acknowledge grant support from The National Health and Medical Research Council,
40 Australia (126402, 209057, 251533, 396414, 450104, 504700, 504702, 504715, 623204, 940394,
41 614296.), VicHealth, Cancer Council Victoria, The Prostate Cancer Foundation of Australia, The
42 Whitten Foundation, PricewaterhouseCoopers, and Tattersall's. EAO, DMK, and EMK
43 acknowledge the Intramural Program of the National Human Genome Research Institute for their
44 support.

1 REFERENCE

- 2 1. Brzezinski A. Melatonin in humans. *N Engl J Med* 1997;**336**: 186-95.
- 3 2. Fu L, Lee CC. The circadian clock: pacemaker and tumour suppressor. *Nat Rev Cancer* 2003;**3**:
4 350-61.
- 5 3. Straif K, Baan R, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, Altieri A, Benbrahim-Tallaa L,
6 Cogliano V. Carcinogenicity of shift-work, painting, and fire-fighting. *Lancet Oncol* 2007;**8**: 1065-6.
- 7 4. Lee S, Donehower LA, Herron AJ, Moore DD, Fu L. Disrupting circadian homeostasis of
8 sympathetic signaling promotes tumor development in mice. *PLoS One* 2010;**5**: e10995.
- 9 5. Wood PA, Yang X, Taber A, Oh EY, Ansell C, Ayers SE, Al-Assaad Z, Carnevale K, Berger FG,
10 Pena MM, Hrushesky WJ. Period 2 mutation accelerates ApcMin/+ tumorigenesis. *Mol Cancer Res*
11 2008;**6**: 1786-93.
- 12 6. Geyfman M, Kumar V, Liu Q, Ruiz R, Gordon W, Espitia F, Cam E, Millar SE, Smyth P, Ihler A,
13 Takahashi JS, Andersen B. Brain and muscle Arnt-like protein-1 (BMAL1) controls circadian cell
14 proliferation and susceptibility to UVB-induced DNA damage in the epidermis. *Proc Natl Acad Sci U S A*
15 2012;**109**: 11758-63.
- 16 7. Fu L, Pelicano H, Liu J, Huang P, Lee C. The circadian gene Period2 plays an important role in
17 tumor suppression and DNA damage response in vivo. *Cell* 2002;**111**: 41-50.
- 18 8. Gery S, Komatsu N, Baldjyan L, Yu A, Koo D, Koeffler HP. The circadian gene per1 plays an
19 important role in cell growth and DNA damage control in human cancer cells. *Mol Cell* 2006;**22**: 375-82.
- 20 9. Gibbs J, Ince L, Matthews L, Mei J, Bell T, Yang N, Saer B, Begley N, Poolman T, Pariollaud M,
21 Farrow S, DeMayo F, et al. An epithelial circadian clock controls pulmonary inflammation and
22 glucocorticoid action. *Nat Med* 2014;**20**: 919-26.
- 23 10. Rahman SA, Castanon-Cervantes O, Scheer FA, Shea SA, Czeisler CA, Davidson AJ, Lockley SW.
24 Endogenous circadian regulation of pro-inflammatory cytokines and chemokines in the presence of
25 bacterial lipopolysaccharide in humans. *Brain Behav Immun* 2015;**47**: 4-13.
- 26 11. Kubo T, Ozasa K, Mikami K, Wakai K, Fujino Y, Watanabe Y, Miki T, Nakao M, Hayashi K,
27 Suzuki K, Mori M, Washio M, et al. Prospective cohort study of the risk of prostate cancer among
28 rotating-shift workers: findings from the Japan collaborative cohort study. *Am J Epidemiol* 2006;**164**:
29 549-55.
- 30 12. Conlon M, Lightfoot N, Kreiger N. Rotating shift work and risk of prostate cancer.
31 *Epidemiology* 2007;**18**: 182-3.
- 32 13. Parent ME, El-Zein M, Rousseau MC, Pintos J, Siemiatycki J. Night work and the risk of cancer
33 among men. *Am J Epidemiol* 2012;**176**: 751-9.
- 34 14. Papantoniou K, Castano-Vinyals G, Espinosa A, Aragonés N, Perez-Gomez B, Burgos J,
35 Gomez-Acebo I, Llorca J, Peiro R, Jimenez-Moleon JJ, Arredondo F, Tardon A, et al. Night shift work,
36 chronotype and prostate cancer risk in the MCC-Spain case-control study. *Int J Cancer* 2015;**137**: 1147-
37 57.
- 38 15. Schernhammer ES, Laden F, Speizer FE, Willett WC, Hunter DJ, Kawachi I, Fuchs CS, Colditz
39 GA. Night-shift work and risk of colorectal cancer in the nurses' health study. *J Natl Cancer Inst* 2003;**95**:
40 825-8.
- 41 16. Lahti TA, Partonen T, Kyronen P, Kauppinen T, Pukkala E. Night-time work predisposes to
42 non-Hodgkin lymphoma. *Int J Cancer* 2008;**123**: 2148-51.
- 43 17. Alexander M, Burch JB, Steck SE, Chen CF, Hurley TG, Cavicchia P, Ray M, Shivappa N, Guess J,
44 Zhang H, Youngstedt SD, Creek KE, et al. Case-control study of the PERIOD3 clock gene length
45 polymorphism and colorectal adenoma formation. *Oncol Rep* 2015;**33**: 935-41.
- 46 18. Cao Q, Gery S, Dashti A, Yin D, Zhou Y, Gu J, Koeffler HP. A role for the clock gene per1 in
47 prostate cancer. *Cancer Res* 2009;**69**: 7619-25.

- 1 19. Chu LW, Zhu Y, Yu K, Zheng T, Yu H, Zhang Y, Sesterhenn I, Chokkalingam AP, Danforth KN,
2 Shen MC, Stanczyk FZ, Gao YT, et al. Variants in circadian genes and prostate cancer risk: a population-
3 based study in China. *Prostate Cancer Prostatic Dis* 2008;**11**: 342-8.
- 4 20. Couto P, Miranda D, Vieira R, Vilhena A, De Marco L, Bastos-Rodrigues L. Association
5 between CLOCK, PER3 and CCRN4L with nonsmall cell lung cancer in Brazilian patients. *Mol Med Rep*
6 2014;**10**: 435-40.
- 7 21. Grundy A, Schuetz JM, Lai AS, Janoo-Gilani R, Leach S, Burstyn I, Richardson H, Brooks-Wilson
8 A, Spinelli JJ, Aronson KJ. Shift work, circadian gene variants and risk of breast cancer. *Cancer Epidemiol*
9 2013;**37**: 606-12.
- 10 22. Jim HS, Lin HY, Tyrer JP, Lawrenson K, Dennis J, Chornokur G, Chen Z, Chen AY, Permut-Wey
11 J, Aben KK, Anton-Culver H, Antonenkova N, et al. Common Genetic Variation in Circadian Rhythm
12 Genes and Risk of Epithelial Ovarian Cancer (EOC). *J Genet Genome Res* 2015;**2**.
- 13 23. Markt SC, Valdimarsdottir UA, Shui IM, Sigurdardottir LG, Rider JR, Tamimi RM, Batista JL,
14 Haneuse S, Flynn-Evans E, Lockley SW, Czeisler CA, Stampfer MJ, et al. Circadian clock genes and risk of
15 fatal prostate cancer. *Cancer Causes Control* 2015;**26**: 25-33.
- 16 24. Monsees GM, Kraft P, Hankinson SE, Hunter DJ, Schernhammer ES. Circadian genes and
17 breast cancer susceptibility in rotating shift workers. *Int J Cancer* 2012;**131**: 2547-52.
- 18 25. Rabstein S, Harth V, Justenhoven C, Pesch B, Plottner S, Heinze E, Lotz A, Baisch C,
19 Schiffermann M, Brauch H, Hamann U, Ko Y, et al. Polymorphisms in circadian genes, night work and
20 breast cancer: results from the GENICA study. *Chronobiol Int* 2014;**31**: 1115-22.
- 21 26. Truong T, Liquet B, Menegaux F, Plancoulaine S, Laurent-Puig P, Mulot C, Cordina-Duverger E,
22 Sanchez M, Arveux P, Kerbrat P, Richardson S, Guenel P. Breast cancer risk, nightwork, and circadian
23 clock gene polymorphisms. *Endocr Relat Cancer* 2014;**21**: 629-38.
- 24 27. Zhu Y, Stevens RG, Hoffman AE, Fitzgerald LM, Kwon EM, Ostrander EA, Davis S, Zheng T,
25 Stanford JL. Testing the circadian gene hypothesis in prostate cancer: a population-based case-control
26 study. *Cancer Res* 2009;**69**: 9315-22.
- 27 28. Zhu Y, Stevens RG, Leaderer D, Hoffman A, Holford T, Zhang Y, Brown HN, Zheng T. Non-
28 synonymous polymorphisms in the circadian gene NPAS2 and breast cancer risk. *Breast Cancer Res Treat*
29 2008;**107**: 421-5.
- 30 29. Zienolddiny S, Haugen A, Lie JA, Kjuus H, Anmarkrud KH, Kjaerheim K. Analysis of
31 polymorphisms in the circadian-related genes and breast cancer risk in Norwegian nurses working night
32 shifts. *Breast Cancer Res* 2013;**15**: R53.
- 33 30. Hung RJ, Ulrich CM, Goode EL, Brhane Y, Muir K, Chan AT, Marchand LL, Schildkraut J, Witte
34 JS, Eeles R, Boffetta P, Spitz MR, et al. Cross Cancer Genomic Investigation of Inflammation Pathway for
35 Five Common Cancers: Lung, Ovary, Prostate, Breast, and Colorectal Cancer. *J Natl Cancer Inst* 2015;**107**.
- 36 31. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, Berndt SI, Bezieau S,
37 Brenner H, Butterbach K, Caan BJ, Campbell PT, et al. Identification of Genetic Susceptibility Loci for
38 Colorectal Tumors in a Genome-Wide Meta-analysis. *Gastroenterology* 2013;**144**: 799-807.e24.
- 39 32. Berndt SI, Wang Z, Yeager M, Alavanja MC, Albanes D, Amundadottir L, Andriole G, Beane
40 Freeman L, Campa D, Cancel-Tassin G, Canzian F, Cornu JN, et al. Two susceptibility loci identified for
41 prostate cancer aggressiveness. *Nat Commun* 2015;**6**: 6889.
- 42 33. Kondratov RV, Gorbacheva VY, Antoch MP. The role of mammalian circadian proteins in
43 normal physiology and genotoxic stress responses. *Curr Top Dev Biol* 2007;**78**: 173-216.
- 44 34. Konturek SJ, Konturek PC, Brzozowski T. Melatonin in gastroprotection against stress-
45 induced acute gastric lesions and in healing of chronic gastric ulcers. *J Physiol Pharmacol* 2006;**57 Suppl**
46 **5**: 51-66.

1 35. Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson
2 BE, Haiman CA, Schleutker J, Hamdy FC, et al. Identification of seven new prostate cancer susceptibility
3 loci through a genome-wide association study. *Nat Genet* 2009;**41**: 1116-21.

4 36. Pharoah PD, Tsai YY, Ramus SJ, Phelan CM, Goode EL, Lawrenson K, Buckley M, Fridley BL,
5 Tyrer JP, Shen H, Weber R, Karevan R, et al. GWAS meta-analysis and replication identifies three new
6 susceptibility loci for ovarian cancer. *Nat Genet* 2013;**45**: 362-70, 70e1-2.

7 37. Schumacher FR, Schmit SL, Jiao S, Edlund CK, Wang H, Zhang B, Hsu L, Huang SC, Fischer CP,
8 Harju JF, Idos GE, Lejbkowitz F, et al. Genome-wide association study of colorectal cancer identifies six
9 new susceptibility loci. *Nat Commun* 2015;**6**: 7138.

10 38. Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeboller H, Risch A, McKay JD,
11 Wang Y, Dai J, Gaborieau V, McLaughlin J, et al. Influence of common genetic variation on lung cancer
12 risk: meta-analysis of 14 900 cases and 29 485 controls. *Hum Mol Genet* 2012;**21**: 4980-95.

13 39. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S,
14 McVean GA, Abecasis GR. A global reference for human genetic variation. *Nature* 2015;**526**: 68-74.

15 40. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to
16 estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;**34**: 816-34.

17 41. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-
18 wide association studies by imputation of genotypes. *Nat Genet* 2007;**39**: 906-13.

19 42. Zhang H, Wheeler W, Hyland PL, Yang Y, Shi J, Chatterjee N, Yu K. A Powerful Procedure for
20 Pathway-Based Meta-analysis Using Summary Statistics Identifies 43 Pathways Associated with Type II
21 Diabetes in European Populations. *PLoS Genet* 2016;**12**: e1006122.

22 43. Rybnikova NA, Haim A, Portnov BA. Is prostate cancer incidence worldwide linked to artificial
23 light at night exposures? Review of earlier findings and analysis of current trends. *Arch Environ Occup*
24 *Health* 2017;**72**: 111-22.

25 44. Hoffman AE, Zheng T, Ba Y, Zhu Y. The circadian gene NPAS2, a putative tumor suppressor, is
26 involved in DNA damage response. *Mol Cancer Res* 2008;**6**: 1461-8.

27 45. Yi CH, Zheng T, Leaderer D, Hoffman A, Zhu Y. Cancer-related transcriptional targets of the
28 circadian gene NPAS2 identified by genome-wide ChIP-on-chip analysis. *Cancer Lett* 2009;**284**: 149-56.

29 46. Zhu Y, Leaderer D, Guss C, Brown HN, Zhang Y, Boyle P, Stevens RG, Hoffman A, Qin Q, Han X,
30 Zheng T. Ala394Thr polymorphism in the clock gene NPAS2: a circadian modifier for the risk of non-
31 Hodgkin's lymphoma. *Int J Cancer* 2007;**120**: 432-5.

32 47. Sengupta ATG. THE PINEAL GLAND, 2011.

33 48. Hardeland R, Madrid JA, Tan DX, Reiter RJ. Melatonin, the circadian multioscillator system
34 and health: the need for detailed analyses of peripheral melatonin signaling. *J Pineal Res* 2012;**52**: 139-
35 66.

36 49. Pevet P, Agez L, Bothorel B, Saboureaux M, Gauer F, Laurent V, Masson-Pevet M. Melatonin
37 in the multi-oscillatory mammalian circadian world. *Chronobiol Int* 2006;**23**: 39-51.

38 50. Toma JG, Amerongen HM, Hennes SC, O'Brien MG, McBlain WA, Buzzell GR. Effects of
39 olfactory bulbectomy, melatonin, and/or pinealectomy on three sublines of the Dunning R3327 rat
40 prostatic adenocarcinoma. *J Pineal Res* 1987;**4**: 321-38.

41 51. Brown CE, Warren S, Chute RN, Ryan KJ, Todd RB. Hormonally induced tumors of the
42 reproductive system of parabiosed male rats. *Cancer Res* 1979;**39**: 3971-6.

43 52. Gann PH, Hennekens CH, Ma J, Longcope C, Stampfer MJ. Prospective study of sex hormone
44 levels and risk of prostate cancer. *J Natl Cancer Inst* 1996;**88**: 1118-26.

45 53. Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG, Lieber MM, Cespedes
46 RD, Atkins JN, Lippman SM, Carlin SM, Ryan A, et al. The influence of finasteride on the development of
47 prostate cancer. *N Engl J Med* 2003;**349**: 215-24.

- 1 54. Papantoniou K, Pozo OJ, Espinosa A, Marcos J, Castano-Vinyals G, Basagana X, Juanola Pages
2 E, Mirabent J, Martin J, Such Faro P, Gasco Aparici A, Middleton B, et al. Increased and mistimed sex
3 hormone production in night shift workers. *Cancer Epidemiol Biomarkers Prev* 2015;**24**: 854-63.
- 4 55. Xi SC, Siu SW, Fong SW, Shiu SY. Inhibition of androgen-sensitive LNCaP prostate cancer
5 growth in vivo by melatonin: association of antiproliferative action of the pineal hormone with mt1
6 receptor protein expression. *Prostate* 2001;**46**: 52-61.
- 7 56. Siu SW, Lau KW, Tam PC, Shiu SY. Melatonin and prostate cancer cell proliferation: interplay
8 with castration, epidermal growth factor, and androgen sensitivity. *Prostate* 2002;**52**: 106-22.
- 9 57. Sigurdardottir LG, Markt SC, Rider JR, Haneuse S, Fall K, Schernhammer ES, Tamimi RM,
10 Flynn-Evans E, Batista JL, Launer L, Harris T, Aspelund T, et al. Urinary melatonin levels, sleep disruption,
11 and risk of prostate cancer in elderly men. *Eur Urol* 2015;**67**: 191-4.
- 12 58. Mukherji A, Kobiita A, Ye T, Chambon P. Homeostasis in intestinal epithelium is orchestrated
13 by the circadian clock and microbiota cues transduced by TLRs. *Cell* 2013;**153**: 812-27.
- 14 59. Lane JM, Vlasac I, Anderson SG, Kyle SD, Dixon WG, Bechtold DA, Gill S, Little MA, Luik A,
15 Loudon A, Emsley R, Scheer FA, et al. Genome-wide association analysis identifies novel loci for
16 chronotype in 100,420 individuals from the UK Biobank. *Nat Commun* 2016;**7**: 10889.
- 17 60. Levi F, Okyar A, Dulong S, Innominato PF, Clairambault J. Circadian timing in cancer
18 treatments. *Annu Rev Pharmacol Toxicol* 2010;**50**: 377-421.
- 19 61. Sung H, Yang HH, Zhang H, Yang Q, Hu N, Tang ZZ, Su H, Wang L, Wang C, Ding T, Fan JH,
20 Qiao YL, et al. Common genetic variants in epigenetic machinery genes and risk of upper gastrointestinal
21 cancers. *Int J Epidemiol* 2015;**44**: 1341-52.
- 22 62. Hernandez DG, Nalls MA, Moore M, Chong S, Dillman A, Trabzuni D, Gibbs JR, Ryten M,
23 Arepalli S, Weale ME, Zonderman AB, Troncoso J, et al. Integration of GWAS SNPs and tissue specific
24 expression profiling reveal discrete eQTLs for human traits in blood and brain. *Neurobiol Dis* 2012;**47**:
25 20-8.

26

Table 1. Summary of study populations and designs for each cancer site

Consortium Name	Cancer Site	No. study*	Cases	Controls	Genotyping Platform	Reference Panel	Covariants
Initial data of GAME-ON							
CORECT	Colorectal	6	5100	4831	Affymetrix Axiom	1000 Genome [†]	age, sex, first 4 principal components (PCs) ³⁷
TRICL	Lung	6	12537	17285	Illumina 317K/550K/610K Illumina	1000 Genome [†]	age, sex, PCs ³⁸
FOCI	Ovary	3	4369	9123	317K/370K/550K/610K/670K/2.5M	1000 Genome [†]	study, first 5 PCs ³⁶
ELLIPSE	Prostate	5	10218	11286	Illumina, Affymetrix	1000 Genome [†]	age, study, PCs ³⁵
Replication data							
PLCO	Prostate	1	4600	2940	Illumina HumanOmni2.5 Beadchip Illumina 550K/610K/CytoSNP/Omni;	1000 Genome [†]	age, 2 significant PCs ³² age, sex (when applicable), center/region (when applicable), batch (when applicable), smoking status (when applicable), first 3 PCs ³¹
GECCO	Colorectal	21	10738	13328	Affymetrix for one study	1000 Genome [†]	

1 *Contributed studies are listed in the supplementary table 1; [†]1000 Genome March 2012 reference panel

2 CORECT: ColoRectal Transdisciplinary Study

3 TRICL: Transdisciplinary Research for Cancer of Lung

4 FOCI: Follow-up of Ovarian Cancer Genetic Association and Interaction Studies

5 ELLIPSE: Elucidating Loci in Prostate Cancer Susceptibility

6 PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

7 GECCO: Genetics and Epidemiology of Colorectal Cancer Consortium

1 Table 2. Pathway results for each cancer site

Cancer	Data	Circadian rhythm pathway		Melatonin pathway	
		N.SNP	P-value	N.SNP	P-value
Prostate	GAME-ON	520	0.014	258	0.024
	PLCO	521	0.28	223	0.21
	Combined data	521	0.0016*	263	0.0060*
Colorectal	GAME-ON	653	0.021	352	0.24
	GECCO	670	0.76	376	0.066
	Combined data	842	0.17	459	0.091
Lung	GAME-ON	510	0.71	243	0.22
Ovary	GAME-ON	521	0.14	263	0.26

2 *Statistically significant after Bonferroni correction ($p < 0.05/8=0.00625$)

3 P-value <0.05 in bold

1 Table 3. Pathway-based and gene-based results between circadian rhythm-melatonin pathway genes and prostate cancer

Gene	Chr	GAME-ON (10218 cases, 11286 controls)		PLCO (4600 cases, 2941 controls)		Combined data (14818 cases, 14227 controls)	
		N.SNP	P-value	N.SNP	P-value	N.SNP	P-value
Circadian rhythm pathway							
ARNTL	11	80	0.41	80	0.40	80	0.29
CK1E	22	48	0.67	48	0.11	48	0.30
CLOCK	4	24	0.013	24	0.44	24	0.021
CRY1	12	35	0.27	35	0.87	35	0.55
CRY2	11	20	0.53	20	0.073	20	0.043
NPAS2	2	167	0.051	167	0.14	167	0.0062
PER1	17	29	0.24	30	0.12	30	0.063
PER2	2	50	0.090	50	0.57	50	0.060
PER3	1	67	0.020	67	0.94	67	0.24
Pathway-level		520	0.014	521	0.28	521	0.0016*
Melatonin pathway							
AANAT	17	34	0.071	38	0.043	38	0.00078*
DDC	7	84	0.033	77	0.63	84	0.050
MTNR1A	4	35	0.041	18	0.52	35	0.35
MTNR1B	11	23	0.94	7	0.92	23	0.96
TPH1	11	18	0.72	18	0.17	18	0.15
TPH2	12	64	0.081	65	0.12	65	0.21
Pathway-level		258	0.024	223	0.21	263	0.0060*

2 *Statistically significant after Bonferroni correction ($p < 0.05/8=0.00625$ at pathway level; $p < 0.05/60=0.00083$ at gene level)

3 P-value<0.05 in bold

4

1 **Table 4. SNPs in *AANAT* and *NPAS2* with prostate cancer with meta-analyses p-value < 0.01**

SNP	Loc	Allele		RAF*	GAME-ON (ELLIPSE)		PLCO		Fixed-effect meta-analyses	
		Ref	Effect		β	P	β	P	β	P
Gene: <i>AANAT</i>										
rs150316415	74475409	G	A	0.94	0.34	4.33×10^{-3}	0.25	2.15×10^{-3}	0.28	3.41×10^{-5}
rs3744045	74475024	G	A	0.08	-0.27	5.04×10^{-3}	-0.21	2.85×10^{-3}	-0.23	4.80×10^{-5}
rs61742551	74472998	G	A	0.98	N/A	N/A	0.41	8.12×10^{-4}	0.41	8.12×10^{-4}
rs9894765	74456426	G	C	0.24	-0.07	0.16	-0.10	2.11×10^{-2}	-0.09	7.14×10^{-3}
rs12945905	74456758	C	T	0.80	0.13	1.67×10^{-2}	0.07	0.14	0.09	8.08×10^{-3}
Gene: <i>NPAS2</i>										
rs1542178	101595475	G	A	0.67	-0.08	6.50×10^{-4}	-0.09	9.88×10^{-3}	-0.08	2.03×10^{-5}
rs2305160	101591304	G	A	0.67	-0.08	7.70×10^{-4}	-0.09	1.52×10^{-2}	-0.08	3.47×10^{-5}
rs2305159	101591443	C	A	0.32	-0.08	4.84×10^{-4}	-0.04	0.24	-0.07	3.37×10^{-4}
rs1542179	101595235	G	A	0.32	-0.08	5.50×10^{-4}	-0.04	0.28	-0.07	4.55×10^{-4}
rs4851392	101581976	G	A	0.74	-0.07	2.26×10^{-3}	-0.06	8.68×10^{-2}	-0.07	4.71×10^{-4}
rs13019460	101461099	G	C	0.21	-0.06	0.18	-0.13	1.70×10^{-3}	-0.10	1.24×10^{-3}
rs6747874	101578489	G	A	0.74	0.08	2.77×10^{-3}	0.05	0.19	0.07	1.27×10^{-3}
rs6747755	101578458	G	A	0.74	0.08	3.18×10^{-3}	0.05	0.19	0.07	1.46×10^{-3}
rs12622050	101579454	G	A	0.76	0.08	2.47×10^{-3}	0.05	0.27	0.07	1.65×10^{-3}
rs12619710	101579487	C	T	0.26	-0.07	3.56×10^{-3}	-0.05	0.21	-0.07	1.73×10^{-3}
rs2278728	101598312	C	T	0.32	-0.07	2.02×10^{-3}	-0.04	0.33	-0.06	1.80×10^{-3}
rs876060	101576964	T	A	0.24	-0.08	2.47×10^{-3}	-0.04	0.31	-0.07	1.92×10^{-3}
rs13012930	101460947	G	A	0.82	0.04	0.18	0.15	9.93×10^{-4}	0.08	2.56×10^{-3}
rs4851391	101579811	G	C	0.24	-0.07	6.25×10^{-3}	-0.05	0.26	-0.06	3.60×10^{-3}
rs4851377	101522266	C	T	0.46	-0.05	5.54×10^{-2}	-0.07	3.33×10^{-2}	-0.06	4.98×10^{-3}
rs13017728	101481348	G	T	0.09	-0.10	0.1.8	-0.15	1.24×10^{-2}	-0.13	5.42×10^{-3}
rs965519	101470349	G	A	0.18	-0.04	0.22	-0.13	2.53×10^{-3}	-0.07	6.15×10^{-3}

rs2309993	101499264	C	T	0.67	0.07	0.10	0.08	3.24×10^{-2}	0.07	7.25×10^{-3}
rs4851386	101566938	C	T	0.52	-0.05	3.58×10^{-2}	-0.06	9.42×10^{-2}	-0.05	7.48×10^{-3}
rs3739006	101566184	G	A	0.52	-0.04	4.22×10^{-2}	-0.06	8.14×10^{-2}	-0.05	7.91×10^{-3}
rs4851385	101566323	G	C	0.48	0.04	4.22×10^{-2}	0.06	8.14×10^{-2}	0.05	7.91×10^{-3}
rs3739005	101566070	C	T	0.48	0.05	3.46×10^{-2}	0.05	0.13	0.05	9.19×10^{-3}

1 *Reference allele frequency. The frequencies are calculated from 503 European subjects in the 1000 Genomes data.

Authors in the ELLIPSE consortium (Alphabetic order)

Ali Amin Al Olama^{1,2}, Demetrius Albanes³, Sara Benlloch⁴, Federico Canzian⁵, Stephen J Chanock³, Constance Turman⁶, Jenny L Denvan⁷, Doug Easton⁴, Ros Eeles⁸, Graham G Giles⁹, Edward L Giovannucci^{6,10}, Henrik Grönberg¹¹, Christopher A Haiman¹², Freddie C Hamdy¹³, Robert N Hoover³, David J Hunter⁵, Tim J Key¹⁴, Laurence N Kolonel¹⁵, Zsofia Kote-Jarai⁷, Loic Le Marchand¹⁵, Sara Lindstrom⁶, Jing Ma⁶, Mitchell J. Machiela³, David E Neal¹⁶, Elio Riboli¹⁷, Fredrick R Schumacher¹⁸, Afshan Siddiq¹⁹, Meir J Stampfer¹⁰, Victoria Stevens²⁰, Ruth C Travis¹⁴, Fredrik Wiklund¹¹, Jianfeng Xu²¹⁻²²

¹Centre of Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

²Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

³Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA

⁴School of Clinical Medicine, University of Cambridge, Cambridge, UK

⁵Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁶Department of Epidemiology, Harvard School of Public Health, Boston MA, USA

⁷University of Bristol, Bristol, UK

⁸American Cancer Society, Inc., Atlanta, GA, USA

⁹Cancer Epidemiology Centre, Cancer Council Victoria Inc., Victoria, Australia

¹⁰Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA

¹¹Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

¹²Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

¹³Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

¹⁴Nuffield Department of Population Health, University of Oxford, Oxford, UK

¹⁵Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA

¹⁶Cambridge Clinical Trial Center & Oncology, University of Cambridge, Cambridge, UK

¹⁷Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

¹⁸Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA

¹⁹Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, UK

²⁰Institute of Cancer Research, London, UK

²¹Fudan Institute of Urology, Huashan Hospital, Fudan University, Shanghai, China

²²Program for Personalized Cancer Care, NorthShore University Health System, Evanston, IL, USA

Additional Acknowledgement

GECCO: The authors would like to thank all those at the GECCO Coordinating Center for helping bring together the data and people that made this project possible. The authors acknowledge Dave Duggan and team members at TGEN (Translational Genomics Research Institute), the Broad Institute, and the Génome Québec Innovation Center for genotyping DNA samples of cases and controls, and for scientific input for GECCO.

ASTERISK: We are very grateful to Dr. Bruno Buecher without whom this project would not have existed. We also thank all those who agreed to participate in this study, including the patients and the healthy control persons, as well as all the physicians, technicians and students.

DACHS: We thank all participants and cooperating clinicians, and Ute Handte-Daub, Utz Benschaid, Muhabbet Celik and Ursula Eilber for excellent technical assistance.

HPFS, NHS and PHS: We would like to acknowledge Patrice Soule and Hardeep Ranu of the Dana Farber Harvard Cancer Center High-Throughput Polymorphism Core who assisted in the genotyping for NHS, HPFS, and PHS under the supervision of Dr. Immaculata Devivo and Dr. David Hunter, Qin (Carolyn) Guo and Lixue Zhu who assisted in programming for NHS and HPFS, and Haiyan Zhang who assisted in programming for the PHS. We would like to thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-Up Study, for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

PLCO: The authors thank Drs. Christine Berg and Philip Prorok, Division of Cancer Prevention, National Cancer Institute, the Screening Center investigators and staff of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, Mr. Tom Riley and staff, Information Management Services, Inc., Ms. Barbara O'Brien and staff, Westat, Inc., and Drs. Bill Kopp and staff, SAIC-Frederick. Most importantly, we acknowledge the study participants for their contributions to making this study possible. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by NCI.

PMH: The authors would like to thank the study participants and staff of the Hormones and Colon Cancer study.

WHI: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: <http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf>

Supplementary Table 1. Population and design of each contributed study

Cancer	Study	Locations	Design
Initial analytical data in GAME-ON			
Colon & Rectum (CORECT)	MECC	US	Cohort
	CFR	US	Cohort
	Kentucky	US	Pop. CC
	CPS-II/ACS	US	Cohort
	Melbourne	Australia	Cohort
	Newfoundland	Canada	Pop. CC
Lung (TRICL)	MDACC	US	Hospital CC
	ICR	UK	Hospital CC
	Toronto	Canada	Clinic CC
	IARC	Europe	Hospital CC
	GLC	German	Pop. CC
	NCI	US	Pop. CC and nested CC
Ovary (FOCI)	UKGWAS	UK	CC
	USGWAS	US, Canada, Poland	CC
	U19	US	CC
Prostate (ELLIPSE)	BPC3	US	CC, nested CC
	CRUK1	UK	CC
	CRUK2	UK	CC
	CAPS1	Sweden	CC
	CAPS2	Sweden	CC
Replication data Prostate (PLCO)	PLCO	US	Nested CC

Colon & Rectum
(GECCO)

ASTERISK	France	Hospital CC
COLO23	US	Pop. CC
DACHS1	Germany	Pop. CC
DACHS2	Germany	Pop. CC
DALS1	US	Pop. CC
DALS2	US	Pop. CC
HPFS1	US	Nested CC
HPFS2	US	Nested CC
HPFSad	US	Nested CC
MEC	US	Nested CC
NHS1	US	Nested CC
NHS2	US	Nested CC
NHSad	US	Nested CC
OFCCR	Canada	Pop.CC
PHS1P2	US	Nested CC
PLCO1	US	Nested CC
PLCO2	US	Nested CC
PMH	US	Pop. CC
VITAL	US	Nested CC
WHI1	US	Nested CC
WHI2	US	Nested CC

CC: case-control

Supplementary table 2. Gene- and pathway-based p-values for overall and aggressive prostate cancer

Gene	Chr	Combined results (14818 cases, 14227 controls)		Aggressive prostate (up to 4446 cases, 12724 controls)	
		N.SNPs	P-value	N.SNPs	P-value
Circadian rhythm pathway					
ARNTL	11	80	0.29	80	0.54
CK1E	22	48	0.30	48	0.58
CLOCK	4	24	0.021	24	0.093
CRY1	12	35	0.55	35	0.87
CRY2	11	20	0.043	20	0.57
NPAS2	2	167	0.0062	167	0.18
PER1	17	30	0.063	30	0.70
PER2	2	50	0.060	50	0.23
PER3	1	67	0.24	67	0.030
Pathway-level		521	0.0016*	521	0.29
Melanotin pathway					
AANAT	17	38	0.00078*	38	0.47
DDC	7	84	0.050	84	0.49
MTNR1A	4	35	0.35	35	0.22
MTNR1B	11	23	0.96	23	0.32
TPH1	11	18	0.15	18	0.96
TPH2	12	65	0.21	65	0.35
Pathway-level		263	0.0060*	263	0.66

*Statistically significant after Bonferroni correction ($p < 0.05/8=0.00625$ at pathway level; $p < 0.05/60=0.00083$ at gene level)

P<0.05 in bold

Supplementary Table 3. Comparison of SNP-based results between overall and aggressive prostate cancer*

Gene	SNP*	Allele		log(OR)	Overall		Aggressive		
		Ref**	Eff**		SE	P-value	log(OR)	SE	P-value
Circadian rhythm pathway									
CLOCK	rs62309758	T	C	-0.09	0.03	1.45E-03	-0.09	0.04	7.57E-03
CRY2	rs7108730	T	C	0.08	0.03	3.66E-03	0.06	0.04	1.05E-01
NPAS2	rs2305160	A	G	0.08	0.02	3.47E-05	0.06	0.03	3.00E-02
Melatonin pathway									
AANAT	rs150316415	G	A	0.28	0.07	3.41E-05	0.16	0.08	6.49E-02
DDC	rs12718611	G	A	-0.11	0.04	1.72E-03	-0.07	0.05	1.12E-01

*SNPs with the smallest p-value in the genes with $P_{\text{gene}} \leq 0.05$, based on association with overall prostate cancer.

** reference and effect alleles

Supplementary table 4. Gene- and pathway-based p-values for colorectal cancer in GAME-ON and replication samples

Gene	Chr	Game-ON (CORECT) (5100 cases, 4831 controls)		GECCO (10738 cases, 13328 controls)		Combined results (15838 cases, 18159 controls)	
		N.SNPs	P-value	N.SNPs	P-value	N.SNPs	P-value
Circadian rhythm pathway							
ARNTL	11	114	0.0044	113	0.78	140	0.028
CK1E	22	38	0.14	55	0.18	68	0.24
CLOCK	4	47	0.18	35	0.34	53	0.11
CRY1	12	56	0.81	47	0.83	73	0.95
CRY2	11	35	0.64	32	0.85	41	0.91
NPAS2	2	202	0.011	212	0.82	245	0.51
PER1	17	47	0.60	38	0.44	53	0.55
PER2	2	54	0.63	54	0.40	68	0.59
PER3	1	60	0.68	84	0.15	101	0.047
Pathway-level		653	0.021	670	0.76	842	0.17
Melatonin pathway							
AANAT	17	53	0.59	52	0.85	61	0.91
DDC	7	119	0.89	115	0.58	147	0.74
MTNR1A	4	60	0.18	61	0.86	72	0.30
MTNR1B	11	33	0.92	34	0.87	45	0.96
TPH1	11	20	0.029	22	0.27	27	0.068
TPH2	12	67	0.77	92	0.0064	107	0.013
Pathway-level		352	0.24	376	0.066	459	0.091

P<0.05 in bold. None of gene based or pathway based p values reached Bonferroni corrected significance

Supplementary table 5. Gene- and pathway-based p-values for lung and ovarian cancers in GAME-ON

Gene	Chr	Lung cancer (12537 cases, 17285 controls)		Ovarian cancer (4369 cases, 9123 controls)	
		N.SNP*	P-value	N.SNP*	P-value
Circadian rhythm pathway					
ARNTL	11	78	0.18	80	0.58
CK1E	22	47	0.35	48	0.024
CLOCK	4	24	0.19	24	0.20
CRY1	12	33	0.40	35	0.29
CRY2	11	18	0.52	20	0.13
NPAS2	2	165	0.56	167	0.046
PER1	17	29	0.35	30	0.87
PER2	2	50	0.87	50	0.54
PER3	1	66	0.90	67	0.68
Pathway-level		510	0.71	521	0.14
Melatonin pathway					
AANAT	17	30	0.63	38	0.14
DDC	7	82	0.089	84	0.10
MTNR1A	4	35	0.93	35	0.20
MTNR1B	11	21	0.85	23	0.64
TPH1	11	17	0.23	18	0.21
TPH2	12	58	0.048	65	0.75
Pathway-level		243	0.22	263	0.26

*SNP numbers after the LD pruning, using $r^2 > 0.95$

P<0.05 in bold. None of gene- or pathway-level p-values reached the Bonferroni correction threshold of significance.