Research Article

Candidate SNP Markers of Chronopathologies Are Predicted by a Significant Change in the Affinity of TATA-Binding Protein for Human Gene Promoters

Petr Ponomarenko,[1] Dmitry Rasskazov,[2] Valentin Suslov,[2] Ekaterina Sharypova,[2] Ludmila Savinkova,[2] Olga Podkolodnaya,[2] Nikolay L. Podkolodny,[2] Natalya N. Tverdokhleb,[2,3] Irina Chadaeva,[2] Mikhail Ponomarenko,[2,3] and Nikolay Kolchanov[2,3]

[1]*Children's Hospital Los Angeles, University of Southern California, Los Angeles, CA 90027, USA*
[2]*Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences, Novosibirsk 630090, Russia*
[3]*Department of Natural Sciences, Novosibirsk State University, Novosibirsk 630090, Russia*

Correspondence should be addressed to Mikhail Ponomarenko; pon@bionet.nsc.ru

Variations in human genome (e.g., single nucleotide polymorphisms, SNPs) may be associated with hereditary diseases, their complications, comorbidities, and drug responses. Using Web service SNP_TATA_Comparator presented in our previous paper, here we analyzed immediate surroundings of known SNP markers of diseases and identified several candidate SNP markers that can significantly change the affinity of TATA-binding protein for human gene promoters, with circadian consequences. For example, rs572527200 may be related to asthma, where symptoms are circadian (worse at night), and rs367732974 may be associated with heart attacks that are characterized by a circadian preference (early morning). By the same method, we analyzed the 90 bp proximal promoter region of each protein-coding transcript of each human gene of the circadian clock core. This analysis yielded 53 candidate SNP markers, such as rs181985043 (susceptibility to acute Q fever in male patients), rs192518038 (higher risk of a heart attack in patients with diabetes), and rs374778785 (emphysema and lung cancer in smokers). If they are properly validated according to clinical standards, these candidate SNP markers may turn out to be useful for physicians (to select optimal treatment for each patient) and for the general population (to choose a lifestyle preventing possible circadian complications of diseases).

1. Introduction

Diurnal (circadian) oscillations of the expression level have been reliably identified in ~10000 genes of placental mammals [1]. The circadian clock of mammals is a system of self-sustained oscillators that function under the control of a central circadian pacemaker located in suprachiasmatic nuclei of the hypothalamus [2]. They synchronize all processes in living organisms, from gene transcription to behavior, thus ensuring their temporal adaptation to 24-hour days on Earth [1]. The minimal set of 12 genes—*CLOCK, ARNTL, ARNTL2, PER1, PER2, CRY1, CRY2, CSNK1E, CSNK1D,*

RORA, RORC, and *NR1D1*—forms the core of the molecular genetic mechanism of the circadian clock, whose functioning is based on feedback relations among its components [3, 4] and on the relations of these genes with the entry points for external signals, which modulate parameters of the circadian clock in response to such external stimuli as light and food [5]. Via the retinohypothalamic tract, the central circadian oscillator imposes a rhythm on peripheral oscillators, which share their molecular genetic structure but work in each cell in accordance with their own specific rhythms of organs, tissues, and systems of tissues [1]. All these oscillators set the rhythm for a multitude of genes via expression of

tissue-specific transcription factors (short-term regulation) or chromatin remodeling (long-term regulation) [6, 7]. Indeed, transcriptomic studies have shown that genes that are subject to circadian control are characterized by overrepresentation of short GC-rich and TA-rich motifs for binding of transcription factors (e.g., TBP-binding motifs) [8, 9] in comparison with genome-wide average values of these parameters. In addition, an empirical study [10] revealed that CLOCK-ARNTL is a pioneer-like transcription factor that interacts with nucleosomes for rhythmic chromatin opening. Adjustment of the peripheral oscillation to the general circadian rhythm synchronizes the functioning of various systems of organs, whereas their desynchronosis can worsen or cause pathological changes in systems that are not interacting directly (e.g., autoimmune disorders may be caused by desynchronosis of the immune defense of the body from exotoxins and excretory/metabolizing systems dealing with analogous endotoxins [11]). Chronopharmacology is concerned with identification of circadian optima for diagnosis [12] and treatment [13, 14].

Experiments on genetic animal models have shown that, in addition to changes in parameters of the circadian clock (amplitude, a phasic response to external signals, or the period of free-flowing rhythm), the mutant animals develop such disorders as metabolic syndrome; disturbances in the system of gluconeogenesis or lipogenesis, in renal function, or in thermogenesis; and development of tumors [15, 16]. Furthermore, research in the field of genetic epidemiology uncovered associations of single nucleotide polymorphisms (SNPs) of circadian clock genes with a wide range of pathological states [17, 18]. A large number of such SNPs are located in noncoding regions of genes (these regions are responsible for regulation of expression). Functional annotation of regulatory SNPs and analysis of their manifestations at the level of gene expression are worthwhile tasks because many of such SNPs may be markers of clinical disorders.

During the "pregenomic era," association of an SNP with a disease used to be a lucky finding [19–22], whereas, now, in the "postgenomic era," identification of such associations is one of the goals of the 1000 Genomes Project [23]. Database dbSNP collects and ranks variants of each SNP by their prevalence [24]. The most frequent variant is entered into the reference human genome GRCh38 (NCBI) or hg38 (UCSC) (the terms used by the UCSC Genome Browser [25]) as an ancestral variant in the Ensembl database [26]. Minor alleles of SNPs in genes involved in a given pathological process can be found by means of the Web service *UCSC Genome Browser* [25], which visualizes a whole-genome map. Subsequent routine genotyping of these alleles in representative cohorts of patients and among healthy volunteers reveals (among minor alleles of SNPs) biomedical markers that are statistically significantly associated with the pathology in question [27]; this procedure takes up a lot of time and work. Computational (bioinformatic) analysis of many millions of unannotated SNPs from the 1000 Genomes Project may accelerate and cheapen the search for biomedical SNP markers.

Thus, the greatest success was achieved in the case of SNPs located in protein-coding regions of genes [28] because of the invariant (predictable) disruptions in the structure-function relations of the proteins encoded by these genes [29]. Moreover, advanced computer-based simulations of molecular dynamics and structures allowed researchers to predict in detail which SNPs would change the proteins. For example, molecular dynamics simulations provide deep analysis of the SNP-caused alterations in the amino acid arrangement that can affect the native three-dimensional atomic conformation of protein structure in order to estimate the most probable conformational modifications [30]. As an alternative/addition to molecular dynamics simulations for conformational sampling of proteins, so-called normal mode-based simulations guarantee multiscale modeling of protein conformational changes [31]. Besides, global minima of molecular docking for native and mutant structures can account for various substrate conformations and help identify an individual conformation with the most favorable binding energy [32]. In the case of drug resistance, computations of shape complementarity—between either widely used or promising new drugs and a binding pocket of a protein altered by an SNP—bring together the advantages of protein structure (or dynamics) simulations and the ability to dock one structure with another [33]. Finally, the alignment of multiple protein structures and/or sequences holds a key to the above calculations for comprehensive SNP analysis of protein-coding gene regions [34]. Meanwhile, the smallest progress was observed with regulatory SNPs because their manifestations may vary from cell to cell, from tissue to tissue, from patient to patient, and from subpopulation to subpopulation [26]. That is why computer-based prediction of candidate regulatory SNP markers of human diseases is a challenging problem for current functional genomics, genetics, and bioinformatics.

In our previous study [35], we described a freely available Web service, SNP_TATA_Comparator (created by us), and demonstrated its practical use on more than 40 biomedical SNP markers in the binding sites for TATA-binding protein (TBP) between positions −70 and −20 relative to the transcription start (the region where all such empirically proven sites are located [36, 37]). Recently, we showed suitability of this Web service for prediction of candidate SNP markers of complications of Mendelian diseases in obesity [38] and of autoimmune complications of these diseases [39] as well as SNP markers that can either enhance or weaken biological activity of oncogene inhibitors during cancer chemotherapy [40] (hereinafter, we use the term "Mendelian disease" according to the notation in database Online Mendelian Inheritance in Man, OMIM® [28]).

In the present work, we applied our Web service SNP_TATA_Comparator [35] to unannotated SNPs in binding sites of TBP which are located near known SNP markers of Mendelian human diseases and, for this reason, can also cause the same pathologies if these SNPs change the affinity of TBP for the same promoters of the same human genes. Furthermore, we found some data on biochemical markers of chronopathologies (where these markers have the effects on gene expression which are identical to the effects of the above SNPs) and clinical studies on the prevalence of these chronopathologies as complications of the Mendelian
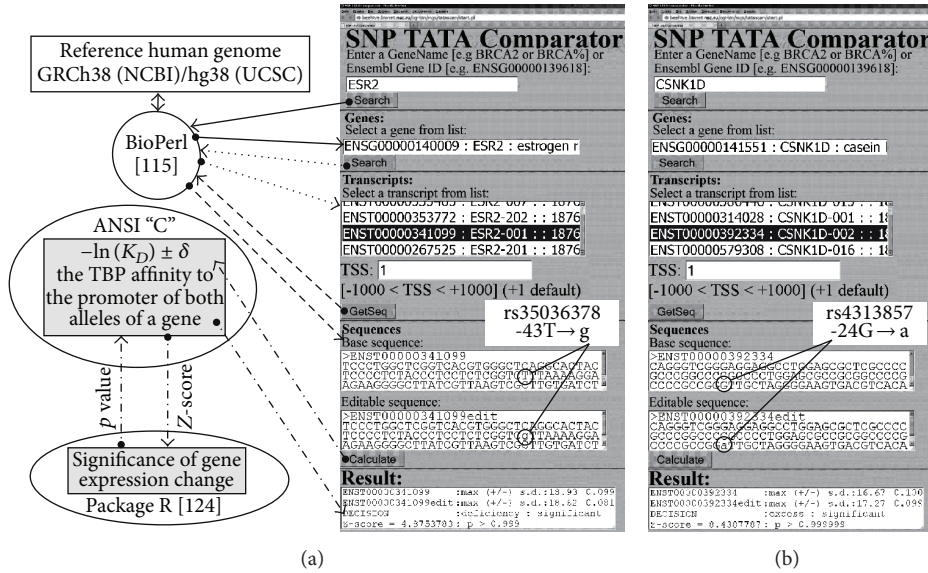
FIGURE 1: Examples of the predictions by SNP_TATA_Comparator [35] for statistically significant alterations in the affinity of TATA-binding protein for human gene promoters. (a) The known biomedical SNP marker rs35036378 located within a promoter of the human *ESR2* gene associated with a Mendelian disease. This SNP is now predicted (in this study) to be a candidate SNP marker of circadian comorbidities and complications of diseases that one can see in Table 1. (b) The candidate SNP marker rs4313857 identified in this study within the human *CSNK1D* gene (belongs to the circadian clock core), as shown in Table 2.

diseases caused by these SNPs. Finally, using SNP_TATA_Comparator [35], we analyzed all SNPs within 90 bp proximal promoter regions for all protein-coding transcripts of the genes of the circadian clock core. As a result, we identified 53 candidate SNP markers of human chronopathologies; validation of these markers in accordance with clinical standards may make these SNPs useful for predictive-preventive personalized medicine [41].

2. Methods

2.1. Web Service.
Web service SNP_TATA_Comparator [35] is a bioinformatics application freely available on the Web (Figure 1; URL: http://beehive.bionet.nsc.ru/cgi-bin/mgs/tatascan/start.pl), which allows a user (i) to find an ancestral variant of the promoter for a transcript under study (the "Base sequence" text box) from the reference human genome (solid, dashed, dotted, and boldfaced arrows; BioPerl [115] is used), (ii) to introduce a mutation of interest (the "Editable sequence" text box), and (iii) to assess (the "Calculate" button) the values of TBP's affinity for these two promoter variants, the relative mutation-related change in transcript levels, and statistical significance according to Z-score (the "Result" text box) as described in detail in our previous study [35].

2.2. The Bioinformatics Model.
For each proximal 90 bp DNA sequence $\{s_{-90} \cdots s_{-1}\}$ of a given gene promoter (where $s_i \in \{a, c, g, t\}$; s_0 is the transcription start site), our Web service SNP_TATA_Comparator [35] calculates the maximal value of $-\ln(K_D) \pm \delta$ of the estimate of TBP's binding affinity for the 26 bp window $\{s_{i-13} \cdots s_i \cdots s_{i+12}\}$ [116, 117] (where $-70 \leq i \leq -20$ in both DNA strands; K_D is the equilibrium dissociation constant of the TBP-DNA complex, expressed in moles per liter; M), as follows:

$$\ln(K_D) = 10.9 - 0.2 \{\ln(K_1([TA]; \mu)) \\ + \ln(K_2(PWM_{Bucher})) + \ln(K_3([WR]; [TV]))\}; \tag{1}$$

$$\delta = \frac{1}{78} \\ \cdot \sum_{\varphi \in \{a,t,g,c\}} \sum_{j=-13}^{12} \ln\left(\frac{K_D(s_{i-13} \cdots s_{i+j-1} s_{i+j} s_{i+j+1} \cdots s_{-1})}{K_D(s_{i-13} \cdots s_{i+j-1} \varphi s_{i+j+1} \cdots s_{-1})}\right), \tag{2}$$

where 10.9 (natural logarithm units) is empirical nonspecific TBP-DNA affinity, 10^{-5} M [118]; 0.2 is the stoichiometric coefficient; K_1 is an empirical estimate of the equilibrium constant of TBP sliding along DNA; the average values of TBP's affinity for double-stranded DNA were estimated using the minor groove width (μ) and the TA dinucleotide content, [TA] [119]. K_2 is an empirical estimate of the equilibrium constant of the primary corecognition between TBP and an appropriate TBP-biding site on DNA [$-\ln(K_2)$ is the maximal score of Bucher's position-weighted matrix: PWM$_{Bucher}$] [120]. K_3 is an empirical estimate of the equilibrium constant of stabilization of the TBP-DNA complex due to the bend of the axis of the DNA helix by an angle of 19° to 90° [121, 122] which depends on abundance of two TA-rich dinucleotides, WR $\in$ {AA, AG, TA, TG} and TV $\in$ {TA, TG, TC} [123]; δ is the standard deviation of K_D estimates for all the possible mononucleotide substitutions within the 26 bp DNA sliding window corresponding to the maximal K_D value found for the DNA sequence under study.

For two DNA sequences of the minor (mut) and ancestral (wt) alleles being compared, (1) and (2) yield $\{-\ln(K_D^{(mut)}) \pm$

TABLE 1: Candidate SNP markers of circadian complications and/or comorbidities of Mendelian diseases as predicted by a significant change in the affinity of TATA-binding protein for human gene promoters.

Gene	dbSNP [24] rel. 146				ID or [reference]	K_D, nM				Known Mendelian diseases (see [references]) and (hypothetically) circadian complications and comorbidities whose candidate SNP markers were predicted by us in [this work] (see Figure 2)	[References] or [this work] (see Figure 2)
	SNP wt→mut	5′ flank	$\frac{wt}{mut}$	3′ flank		$\frac{wt}{mut}$	Δ	Z	α		
	-29tDEL	gggctgggca	$\frac{t}{—}$	atacaacagt	rs397509430	$\frac{5}{29}$	→	34	10^{-6}		
	-29t→a, g, c	gggctgggca	$\frac{t}{a, g, c}$	atacaacagt	rs33980857	$\frac{5}{21}$	→	27	10^{-6}		
HBB	-28a→g	ggctgggcat	$\frac{a}{g}$	aaagtcaggg	rs34598529	$\frac{5}{18}$	→	24	10^{-6}	Malaria resistance and β-thalassemia and (hypothetically) circadian symptoms (worse at night) in restless legs syndrome caused by iron deficiency anemia cooccurring with thalassemia and for sensorineural hearing loss as a complication of deferoxamine-based therapy in thalassemia	[24, 42], [this work] [43–47]
	-27a→g, c	gctgggcata	$\frac{a}{g, c}$	aagtcagggc	rs33931746	$\frac{5}{11}$	→	14	10^{-6}		
	-30a→g, c	agggctgggc	$\frac{a}{g, c}$	taaaagtcag	rs33981098	$\frac{5}{9}$	→	10	10^{-6}		
	-31c→a, t, g	cagggctggg	$\frac{c}{a, t, g}$	ataaaagtca	rs34500389	$\frac{5}{6}$	→	3	10^{-2}		
HBB	-25aaDEL	ctgggcataa	$\frac{aa}{—}$	gtcagggcag	rs63750953	$\frac{5}{8}$	→	9	10^{-6}	(Hypothetically) malaria resistance, β-thalassemia, and circadian symptoms (worse at night) in restless legs syndrome and sensorineural hearing loss (for details, see above)	[This work] [42–47]
	-25a→c	tgggcataaa	$\frac{a}{c}$	gtcagggcag	rs281864525	$\frac{5}{7}$	→	7	10^{-6}		
HBD	-31a→g	caggaccagc	$\frac{a}{g}$	taaaaggcag	rs35518301	$\frac{4}{8}$	→	11	10^{-6}	Malaria resistance and δ-thalassemia and (hypothetically) circadian symptoms (worse at night) for restless legs syndrome and sensorineural hearing loss (for details, see above)	[24, 42], [this work] [43–47]
HBD	-30t→c	aggaccagca	$\frac{t}{c}$	aaaaggcagg	rs34166473	$\frac{4}{8}$	→	18	10^{-6}	(Hypothetically) malaria resistance, δ-thalassemia, and circadian symptoms (worse at night) for restless legs syndrome and sensorineural hearing loss (for details, see above)	[This work] [42–47]
MMP12	-27a→g	gatatcaact	$\frac{a}{g}$	tgagtcactc	rs2276109	$\frac{11}{14}$	→	3	10^{-2}	Low risk of systemic sclerosis, psoriasis, and asthma whose symptoms are circadian (worse at night)	[48–51]

TABLE 1: Continued.

Gene	ID or [reference]	dbSNP [24] rel. 146 SNP wt→mut	5′ flank	wt/mut	3′ flank	K_D, nM wt/mut	Δ	Z	α	Known Mendelian diseases (see [references]) and (hypothetically) circadian complications and comorbidities whose candidate SNP markers were predicted by us in [this work] (see Figure 2)	[References] or [this work] (see Figure 2)
MMP12	rs572527200	-30a→g	gatgatatca	a/g	ctatgagtca	11/14	→	3	10^{-2}	(Hypothetically) low risk of systemic sclerosis, psoriasis, and asthma whose symptoms are circadian (worse at night)	[This work] [48–51]
IL1B	rs1143627	-31c→t	ttttgaaagc	c/t	ataaaacag	5/2	↑	15	10^{-6}	Gastric cancer in Helicobacter pylori infection, hepatocellular carcinoma in hepatitis C virus infection, non-small cell lung cancer, chronic gastritis and gastric ulcer in H. pylori infection, Graves' disease, greater body fat in older men, recurrent major depressive disorder whose diagnosis and therapy are characterized by circadian optima for use, and (hypothetically) bipolar disorder whose diagnosis and therapy are characterized by circadian optima for use depending on the diet	[24, 52–59], [this work] [60, 61]
IL1B	rs549858786	-28a→t	tgaaagccat	a/t	aaaacagcga	5/6	→	8	10^{-6}	(Hypothetically) rheumatoid arthritis that can disrupt the circadian rhythm of IL1B gene expression	[This work] [62, 63]
F3	rs563763767	-21c→t	ccctttatag	c/t	gcgcggggca	3/2	↑	6	10^{-6}	Venous thromboembolism and myocardial infarction that are characterized by their circadian preference for the early morning in the elderly	[64–66]
F7	See [67]	-33a→c	ccttggaggc	a/c	gagaactttg	53/62	→	3	10^{-2}	Moderate bleeding tendency whose symptoms are circadian: worse with chronic change of time zones and in winter	[67, 68]
F7	rs749691733	-21c→t	agaactttgc	c/t	cgtcagtccc	53/66	→	4	10^{-3}	(Hypothetically) moderate bleeding tendency whose symptoms are circadian: worse with chronic change of time zones and in winter	[This work] [67, 68]
F7	rs367732974	-19g→a	aactttgccc	g/a	tcagtccat	53/47	↑	2	0.05	(Hypothetically) heart attacks that are characterized by circadian preference for the early morning in the elderly and circadian (postprandial) flare-ups of thrombogenesis	[This work] [69, 70]
F7	rs549591993	-13c→a	gcccgtcagt	c/a	ccatggggaa	53/25	↑	13	10^{-6}		
F7	rs77947114	-23g→a	agagaacttt	g/a	ccgctgagtc	53/19	↑	19	10^{-6}		
F7	rs770113559	-38g→a	gtcacccttg	g/a	aggcagagaa	53/41	↑	5	10^{-6}		
F7	rs754814507	-54c→t	cctccccat	c/t	cctctgtcac	53/45	↑	3	10^{-3}		

TABLE 1: Continued.

Gene	ID or [reference]	dbSNP [24] rel. 146 SNP wt→mut	5′ flank	wt/mut	3′ flank	K_D, nM wt/mut	Δ	Z	α	Known Mendelian diseases (see [references]) and (hypothetically) circadian complications and comorbidities whose candidate SNP markers were predicted by us in [this work] (see Figure 2)	[References] or [this work] (see Figure 2)
NOS2	See [71]	-51t→c	gtataaatac	t/c	tcttggctgc	2/1	↑	3	10^{-2}	Resistance to malaria and high risk of epilepsy that damages the hypothalamus and the circadian clock as a whole and (hypothetically) remission of panic disorder whose symptoms are circadian (worse late in the evening)	[71–73], [this work] [74]
DHFR	rs10168	-26g→a	ctgcacaaat	g/a	gggacgaggg	15/9	↑	9	10^{-6}	In leukemia, resistance to methotrexate therapy that is characterized by a circadian optimum for use	[13, 75]
DHFR	rs750793297	-25g→t	tgcacaaatg	g/t	ggacgagggg	15/13	↑	3	0.01	(Hypothetically) in leukemia, resistance to methotrexate therapy that is characterized by a circadian optimum for use	[This work] [13, 75]
DHFR	rs766799008	-28a→g	ctgcacaaat	a/g	tggggacgag	15/19	↓	3	10^{-3}		
DHFR	rs764508464	-a28DEL	ctgcacaaat	a/—	tggggacgag	15/37	↓	17	10^{-6}	(Hypothetically) enhancement of an apparent bioactivity of methotrexate therapy that may be characterized by a circadian optimum for use	[This work] [13, 40, 75]
DHFR	rs754122321	-31c→g	ctcgcctgca	c/g	aaatggggac	15/25	↓	9	10^{-3}		
StAR	rs16887226	-33c→t	cagccttcag	c/t	gggggacatt	10/10	=	0	>0.5	Hypertension in diabetes (EMSA: an unknown TF-binding site is disrupted rather than a TBP-binding site) and (hypothetically) weak circadian resistance to exotoxins (deficiency of the mediator between the circadian and immune systems)	[76] [this work] [11]
StAR	rs544850971	-22a→g	tcagcggggg	a/g	catttaagac	10/12	↓	5	10^{-2}	(Hypothetically) hypertension in diabetes and weak circadian resistance to exotoxins	[This work] [11, 76]
CETP	See [77]	-54[18 bp]DEL	cgtggggct	[18]/—	gggctccagg	4/7	↓	7	10^{-6}	Hyperalphalipoproteinemia reduces the risk of atherosclerosis that is characterized by circadian (postprandial) flare-ups	[77–79]
CETP	rs17231520	-68g→a	ggggctgggc	g/a	gacatacata	4/2	↑	10	10^{-6}	(Hypothetically) risk of atherosclerosis that is characterized by circadian (postprandial) flare-ups and hypoalphalipoproteinemia causing chronopathologies of the liver	
CETP	rs56903466	-53g→a	atacatatac	g/a	ggctccaggc	4/3	↑	4	10^{-3}		[This work] [80]
CETP	rs757176551	-49c→g	catatacggg	c/g	tccaggctga	4/2	↑	10	10^{63}		
APOA1	See [81]	-35a→c	tgcagacata	a/c	ataggccctg	3/4	↓	5	10^{-3}	Hypoalphalipoproteinemia causing chronopathologies in the liver, as well as hematuria, fatty liver, and obesity, and (hypothetically) atherosclerosis that is characterized by circadian (postprandial) flare-ups	[80, 81] [this work] [77–80]

TABLE 1: Continued.

Gene	ID or [reference]	SNP wt→mut	5' flank	wt/mut	3' flank	K_D, nM wt/mut	Δ	Z	α	Known Mendelian diseases (see [references]) and (hypothetically) circadian complications and comorbidities whose candidate SNP markers were predicted by us in [this work] (see Figure 2)	[References] or [this work] (see Figure 2)
		dbSNP [24] rel. 146									
CYP2B6	rs34223104	-28t→c	gatgaaattt	t/c	ataacagggt	4/10	↓	15	10^{-6}	Better metabolic activation of anticancer prodrug cyclophosphamide that is characterized by a circadian optimum for use	[14, 82]
CYP2B6	rs63558831	-26t→c	tgaaatttta	t/c	aacagggtgc	4/10	↓	13	10^{-6}	(Hypothetically) better metabolic activation of cyclophosphamide that is characterized by a circadian optimum for use	[This work] [14, 82]
INS	rs5505	-9c→t	agatcactgt	c/t	cttctgccat	53/44	↑	4	10^{-6}	Type 1 diabetes after neonatal diabetes mellitus and risk of hyperinsulinemia that disturbs circadian rhythms of the reproductive system, of blood pressure, and of tumor-host balance	[24, 83–85]
INS	rs56307167	-28c→t	tcagccctgc	c/t	tgtctcccag	53/44	↑	4	10^{-3}	(Hypothetically) type 1 diabetes after neonatal diabetes mellitus and risk of hyperinsulinemia that disturbs circadian rhythms of the reproductive system, of blood pressure, and of tumor-host balance	[This work] [24, 83–85]
INS	rs11557611	-8c→t	gatcactgtc	c/t	ttctgccatg	53/60	↓	2	0.05	(Hypothetically) hypothalamic amenorrhea	[This work] [86]
ESR2	rs35036378 (see Figure 1(a))	-43t→g	ccctcggtc	t/g	ttaaaaggaa	6/8	↓	5	10^{-3}	Preventive therapy for an ESR2-deficient pT1 primary tumor against its progression to breast cancer by means of tamoxifen when this treatment is characterized by a circadian optimum for use and (hypothetically) disturbances in circadian (daytime) behavioral activity	[87–89] [this work] [90]
ESR2	rs766797386	-32g→t	ttaaaaggaa	g/t	aaggggctta	6/7	↓	3	10^{-2}	(Hypothetically) preventive therapy for an ESR2-deficient primary pT1 tumor against its progression to breast cancer by means of tamoxifen when this treatment is characterized by a circadian optimum for use; disturbances in circadian (daytime) behavioral activity	[This work] [87–90]

wt, ancestral allele; mut, minor allele; K_D, an estimate [35] of the dissociation constant (K_D) of the TBP-DNA complex in vitro [91]; Δ, a gene expression change in comparison with the norm (=): overexpression (↑) and underexpression (↓); Z, Z-score; $\alpha = 1 - p$, significance (where p is a probability rate of acceptance of H_0 hypothesis "$K_D^{(mut)} \neq K_D^{(wt)}$"; see Figure 1); TF, transcription factor; EMSA, electrophoretic mobility shift assay.

TABLE 2: Candidate SNP markers within the human genes of the circadian clock core as predicted by a significant change in the affinity of TATA-binding protein for human gene promoters.

Gene	ID	dbSNP [24] rel. 146 SNP wt→mut	5′ flank	wt/mut	3′ flank	K_D, nM wt/mut	Δ	Z	α	Circadian complications and comorbidities whose candidate SNP markers were predicted in [this work] (clinical data, laboratory animals, or cellular model)	References found (see Figure 2)
CLOCK	rs192518038	-57**g**→t	aggacctaag	**g**/t	ctagcgctct	63/29	↑	14	10^{-6}	(Hypothetically) higher risk of heart attacks that are characterized by circadian preference for the early morning in the elderly with diabetes (CLOCK-mutant mice)	[This work] [92]
	rs53733415	-63**c**→t	gcctccagga	**c**/t	ctaaggctag	63/45	↑	7	10^{-6}		
ARNTL	rs534789405	-28**g**→a	cggattggct	**g**/a	ggggcggccg	184/89	↑	12	10^{-6}	(Hypothetically) malignant pleural mesothelioma (13 malignant pleural mesothelioma cell lines in comparison with the nontumorigenic mesothelial cell line MeT-5A)	[This work] [93]
	rs758737644	-39**c**→t	tgcactgtta	**c**/t	acattctgtt	10/3	↑	4	10^{-3}		
	rs54903l146	-46**t**→a	caaaacttat	**t**/a	ggggtgctatg	6/5	↑	4	10^{-3}		
	rs776246315	-21**g**→a, (t)	ttccagccgc	**g**/a, (t)	tgagtccagg	49/31	↑	8	10^{-6}	(Hypothetically) inhibition of type 1 diabetes (murine model)	[This work] [94]
	rs140915764	-22**c**→t	tttccagccg	**c**/t	gtgagtccag	49/32	↑	7	10^{-6}		
	rs756988598	-23**g**→a	gtttccagcc	**g**/a	cgtgagtcca	49/39	↑	4	10^{-3}		
ARNTL2	rs753093730	-33**g**→a, t	ctgccatag	**g**/a, t	taaagtgttg	7/5	↑	3	10^{-3}		
	rs536395877	-46**c**→a	ttgttgtact	**c**/a	tgctgccat	7/5	↑	5	10^{-3}		
	rs769981079	-42**g**→a, t	ccagtgcatt	**g**/a, t	ctcctgtggt	49/26	↑	12	10^{-6}		
	rsl1899732	-46**c**→t	agaaccagtg	**c**/t	attgctcctg	49/14	↑	18	10^{-6}		
	rs746050396	-57**g**→a	gttgagagag	**g**/a	agaaccagtg	49/36	↑	5	10^{-6}		
ARNTL2	rs369143719	-48**c**→t	tcttgttgta	**c**/t	tctgctgccc	7/9	↓	4	10^{-3}	(Hypothetically) suppression of diabetes protection (murine model)	[This work] [95]
	rs374142420	-51**g**→c	atgtcttgtt	**g**/c	tactctgctg	7/10	↓	4	10^{-3}		
	rs770635249	-56[**ttg**]DEL	aataaatgtc	**ttg**/—	ttgtactctg	7/10	↓	4	10^{-3}		
CRY1	rs747100146	INS-49tt	gataggagtt	—/tt	aattattccta	6/7	↓	2	0.05	(Hypothetically) susceptibility to arthritis ($CRY1^{-/-}$ $CRY2^{-/-}$ knockout mice)	[This work] [96]

Table 2: Continued.

Gene	ID	dbSNP [24] rel. 146 SNP wt→mut	5′ flank	wt/mut	3′ flank	K_D, nM wt/mut	Δ	Z	α	Circadian complications and comorbidities whose candidate SNP markers were predicted in [this work] (clinical data, laboratory animals, or cellular model)	[References] found (see Figure 2)
CRY2	rs753656899	-16a→g	gcggggacta	a/g	gggtggagtt	27/56	↓	12	10^{-6}	(Hypothetically) susceptibility to arthritis and to mood disorders (CRY1$^{-/-}$CRY2$^{-/-}$ and CRY2$^{-/-}$ knockout mice, resp.)	[This work] [96, 97]
	rs575588903	-9a→g	ctaagggtgg	a/g	gttgcggcgt	27/25	↑	2	0.05		
CRY2	rs757256843	-14g→t	ggggactaag	g/t	gtggagttgc	27/13	↑	10	10^{-6}	(Hypothetically) resistance to chemotherapy and poor prognosis in colorectal cancer (colorectal cancer samples)	[This work] [98]
	rs760179689	-24g→a	ccctgtgggc	g/a	gggactaagg	27/22	↑	1	10^{-3}		
	rs529410313	-47c→a	agctgtcagt	c/a	ttgcaagtca	22/18	↑	3	10^{-2}		
	rsl37890200	-17c→t	gccaataagg	c/t	ggagagtgtg	21/14	↑	6	10^{-6}	(Hypothetically) longer survival among patients with gastric cancer	[This work] [99]
PER1	rs773740924	-26c→a	ctcgccctgg	c/a	caataaggcg	21/14	↑	7	10^{-6}		
	rs2518024	-60g→a	gtgtctgga	g/a	ttaaacagc	17/8	↑	12	10^{-6}		
PER1	rs796629786	-13t→g	tcggcgcccc	t/g	aagccaataa	56/118	↓	13	10^{-6}	(Hypothetically) prostate cancer; ethanol hepatotoxicity, defects in hippocampal development, and resulting impairment of spatial learning capacity (PER1$^{-/-}$ knockout mice)	[This work] [17, 100–103]
	rs3027175	-67c→t	ccagcaggtg	c/t	tctggagtta	17/19	↓	2	0.05		
PER2	rs780846747	-20c→t	cacaccttgt	c/t	aagtagaaga	10/5	↑	12	10^{-6}	(Hypothetically) susceptibility to acute Q fever in male patients; suppression of tumor growth (cell line S-180)	[This work] [104, 105]
	rs750596430	-13c→t	agccaggcag	c/t	agcggcgcgg	106/60	↑	10	10^{-6}		
	rs535050458	-20c→t	gccagcgagc	c/t	aggcagcagc	106/72	↑	7	10^{-6}		
RORA	rs54717 0533	-21c→a,(t)	cgccagcgag	c/a,(t)	caggcagcag	106/72	↑	7	10^{-6}	(Hypothetically) emphysema and its progression to lung cancer among smokers (mice, smoking machine TE-10)	[This work] [106]
	rs551503425	-50c→t	tgccaatatc	c/t	aagggttgcc	19/6	↑	15	10^{-6}		
	rs374778785	-56a→t	attatccccc	a/t	tactcctccc	34/28	↑	4	10^{-3}		

TABLE 2: Continued.

| Gene | dbSNP [24] rel. 146 | | | | | K_D, nM | | Δ | Z | α | Circadian complications and comorbidities whose candidate SNP markers were predicted in [this work] (clinical data, laboratory animals, or cellular model) | [References] found (see Figure 2) |
	ID	SNP wt→mut	5' flank	wt/mut	3' flank	wt	mut					
RORA	rs764749271	-28[12 bp]DEL	ccttctcctt	[12 bp]/—	tttttttt	28	31	↓	2	0.05	(Hypothetically) severe cerebellar ataxia (RORA$^{-/-}$ knockout mice)	[This work] [107]
	rs762440045	-55a→c	acatgggagtc	a/c	gctccggcag	106	245	↓	15	10^{-6}		
RORC	rs568650510	-15c→t	actcctttc	c/t	ctgcctgctg	55	25	↑	14	10^{-6}	(Hypothetically) neurological manifestations of Behçet syndrome and asthma whose symptoms are circadian (worse at night) (18 patients and 30 pediatric patients, resp.)	[This work] [51, 108, 109]
	rs777965060	-12c→t	aggccctctg	c/t	ttgacccca	80	69	↑	3	0.05		
	rs369188273	-16c→t	ccctccctc	c/t	gcgccgctc	288	195	↑	7	10^{-6}		
	rs558609213	-18c→t	tcttttcttg	c/t	atccctgcag	30	9	↑	21	10^{-6}		
	rs201914168	-20g→t	cctgacagag	g/t	ccctctgctt	80	53	↑	7	10^{-6}		
CSNK1E	rs374953337	-25c→t	tgaccctga	c/t	agaggccctc	80	27	↑	17	10^{-6}	(Hypothetically) higher risk of c-MYC-dependent carcinogenesis, testicular cancer, and overproduction of cerebral β-amyloid	[This work] [110–112]
	rs368019196	-31c→t	tgcctctgac	c/t	cctgacagag	80	68	↑	3	10^{-2}		
	rs775747363	-40[gcc]DEL	ctggctgcct	gcc/—	tctgacccct	80	67	↑	4	10^{-3}		
	rs780432736	-51g→a	gccgggccgtg	g/a	ctggctgcct	80	71	↑	2	0.05		
	rs747636670	-52g→a	ggccggggcgt	g/a	gctggctgcc	80	56	↑	5	10^{-6}		

TABLE 2: Continued.

Gene	ID	SNP wt→mut	5′ flank	wt/mut	3′ flank	K_D, nM wt/mut	Δ	Z	α	Circadian complications and comorbidities whose candidate SNP markers were predicted in [this work] (clinical data, laboratory animals, or cellular model)	[References] found (see Figure 2)
	rs775283367	-11c→g	ctcttaccta	c/g	gtcagctctt	8/10	→	4	10^{-3}		
CSNK1E	rs746761879	-18t→c	ccgactactc	t/c	tacctacgtc	8/11	→	5	10^{-6}	(Hypothetically) increased responsiveness to opioids (CSNK1E$^{-/-}$ knockout mice)	[This work] [113]
	rs777083641	-22[ag]DEL	ctggctgcct	ag/—	tctgacccct	80/93	→	3	10^{-2}		
	rs2899302	-59g→c	cgagaaaact	g/c	cgcgaggcct	288/335	→	3	10^{-2}		
CSNK1D	rs4313857 (see Figure 1(b))	-24g→a	gcccgcgcgg	g/a	ttgctagggg	57/32	↑	8	10^{-6}	(Hypothetically) breast cancers (27 surgically resected tumor samples)	[This work] [114]
	rs571866458	-38c→a	gagccggacc	c/a	gcagtagcgg	54/42	↑	4	10^{-3}		
	rs540139460	-56g→a	gcagggtcgg	g/a	aggaggcctg	253/145	↑	10	10^{-6}		
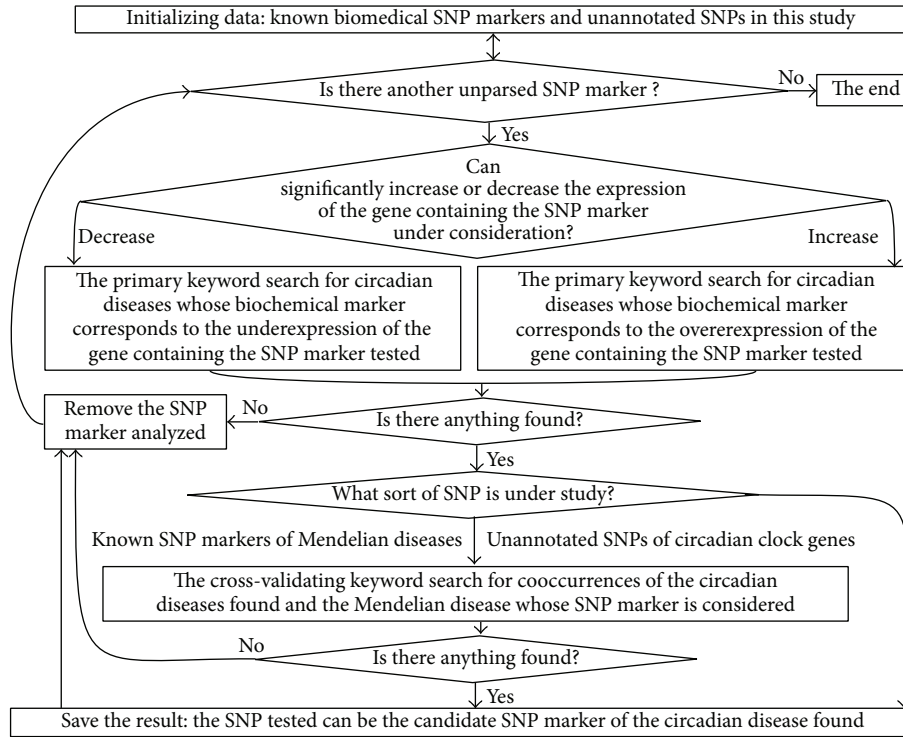
FIGURE 2: A flow chart showing our manual keyword search for chronopathologies (whose known biochemical markers correspond to our predictions of a significant change in the affinity of TATA-binding protein for human gene promoters).

$\delta_{(\text{mut})}\}$ and $\{-\ln(K_D^{(\text{wt})}) \pm \delta_{(\text{wt})}\}$, respectively. Our Web service SNP_TATA_Comparator [35] compares them using Fisher's Z-score [124]:

$$Z = \frac{\left| \ln \left(K_D^{(\text{mut})} / K_D^{(\text{wt})} \right) \right|}{\sqrt{\delta_{(\text{mut})}^2 + \delta_{(\text{wt})}^2}}. \tag{3}$$

Using the standard statistical package R [124], we transform Z-score into p value of the probability of acceptance of H_0 hypothesis "$K_D^{(\text{mut})} \neq K_D^{(\text{wt})}$" (where $\alpha = 1 - p$ is the statistical significance). Two cases, "$K_D^{(\text{mut})} < K_D^{(\text{wt})}$" and "$K_D^{(\text{mut})} > K_D^{(\text{wt})}$," correspond, respectively, to overexpression and underexpression of the gene under study [125]. For more details, see our previous article [35].

2.3. Keyword Search. Figure 2 shows this keyword search for data on known biochemical markers of chronopathologies; these markers correspond to predictions of SNP_TATA_Comparator (Figure 1) regarding a relative mutation-induced change in gene expression. For each known or candidate SNP marker causing either significant overexpression or underexpression of the human gene containing the SNP, we performed a manual keyword search using various combinations of the terms "overexpression," "deficiency," "circadian," and many others corresponding to chronopathologies in public databases, as described in detail elsewhere [126]. In the case of genes of the circadian clock core, the obtained data are in Table 2 as results of this study. For SNP markers of Mendelian diseases, we conducted an additional

keyword search for data on the prevalence of the uncovered chronopathologies as complications of these diseases; this procedure is some sort of cross-validation of the rough qualitative rates without statistical testing (Table 1).

Our heuristic interpretation of the keyword search results is shown in *italics* in the second rightmost column of Tables 1 and 2 and labeled with the word "*(Hypothetically)*" in front. We cite the studies (found during our manual keyword search) within the rightmost column of these tables, shown as [references] in *italics* and labeled with the phrase "*[This work].*"

3. Results and Discussion

3.1. The Results on Candidate SNP Markers of Circadian Complications of Mendelian Diseases. These results are presented in Table 1. Let us review in detail these more comprehensively studied SNP markers in order to briefly describe, in a similar fashion, the candidate SNP markers in the genes of human circadian clock core which were identified for the first time (in our study).

Genes HBB and HBD encode β- and δ-chains of hemoglobin, respectively. In the binding sites for TBP in their promoters, these two genes contain the greatest number (seven) of known SNP markers (rs35518301, rs397509430, rs33981098, rs34598529, rs33931746, rs33980857, and rs34500389) of thalassemia and resistance to malaria [24, 42], as a result of a hemoglobin deficiency (Table 1). A primary search by keywords uncovered a hemoglobin deficiency as a biochemical marker of circadian (nocturnal) aggravation of restless legs syndrome [43] and sensorineural hearing

loss [44]. A cross-validating search by keywords revealed that iron deficiency anemia substantially contributes to the pathogenesis of restless legs syndrome and cooccurs with thalassemia [45, 46], whereas sensorineural hearing loss is a complication of thalassemia in children during treatment with deferoxamine [47]. We found three additional unannotated SNPs (rs63750953, rs281864525, and rs34166473) that can also reduce expression of genes *HBB* and *HBD* and may serve as candidate SNP markers of these chronopathologies.

The MMP12 gene codes for matrix metalloproteinase 12 and, in its promoter, contains a known SNP marker (rs2276109) of a lower risk of systemic sclerosis [48], psoriasis [49], and asthma [50]. A keyword search yielded circadian (nocturnal) aggravation of asthma symptoms [51]. Here we found an unannotated SNP (rs572527200) with the same effects on the TBP-promoter affinity.

Gene IL1B encodes interleukin 1β and, in its promoter, contains one of the most widely studied SNP markers (rs1143627) of stomach ulcer, chronic gastritis, gastric cancer, hepatocellular carcinoma, non-small cell lung cancer, Graves' disease, and excess body fat in older men [52–57] as well as major depressive disorder [58] with a circadian optimum for diagnosis and treatment [59] that can be shifted by a high-fat or high-carbohydrate diet [60]. The primary search by keywords uncovered association of *IL1B* overexpression (with "-31T") with a bipolar disorder [61] that also has a circadian optimum for diagnosis and treatment depending on the diet [60]. Near this known SNP marker, we found unannotated rs549858786, which was found to lower *IL1B* expression (Table 1). The primary keyword search produced an IL1B protein deficiency as a biochemical marker of rheumatoid arthritis [62], for which an additional keyword search yielded a study showing that this disease is associated with disturbances of the circadian rhythm of *IL1B* expression [63].

The F3 gene encodes tissue thromboplastin (factor III) and, in its promoter, contains a known SNP marker (rs563763767) of an elevated risk of venous thromboembolism and myocardial infarction [64]. A keyword search produced clinical data on circadian aggravation of their symptoms (in the early morning) in the elderly [65]; these data are in agreement with basic research on a murine model of aging [66].

The F7 gene codes for serum prothrombin conversion accelerator (factor VII); in its promoter, some researchers [67] found a biomedical SNP marker: a substitution of the ancestral nucleotide A for minor nucleotide C at position -33 relative to the transcription start site (hereafter -33A→ C); this is a marker of moderate bleeding (as a result of underexpression of this gene). An additional database search revealed laboratory data on possible circadian aggravation of this disorder's symptoms during chronic changes of time zones and in the winter (data from a mouse model) [68]. Here we found an unannotated SNP (rs749691733) with the same effects on the TBP-promoter interaction. In addition, near this known SNP marker, we found five unannotated SNPs (rs367732974, rs549591993, rs777947114, rs770113559, and rs754814507) that can cause *F7* overexpression (Table 1). A keyword search produced an elevated F7 protein level as a biochemical marker

of heart attacks characterized by a circadian preference for the early morning in the elderly [69] and for circadian (postprandial) development of thrombogenesis [70]. Therefore, we propose rs367732974, rs549591993, rs777947114, rs770113559, and rs754814507 as candidate SNP markers of these two chronopathologies.

Gene NOS2 encodes inducible NO synthase; in its promoter, one study [71] uncovered an SNP marker (-51T→C) of resistance to malaria [71] and of a high risk of epilepsy [72] (as a result of overexpression of this gene). A keyword search yielded a review article [73] about epilepsy-associated hypothalamic damage that can impair the circadian clock system of the body as a whole [73]. Besides, we found some data [74] suggesting that excess NO is a biochemical marker of a remission of panic disorder that is characterized by circadian (late evening) aggravation of symptoms. Thus, we propose the SNP "NOS2: -51T→C" as a candidate marker of these chronopathologies.

Gene DHFR codes for dihydrofolate reductase; its promoter contains a known SNP marker (rs10168) of methotrexate resistance [75] that is characterized by a therapeutic optimum of its use [13]. Here we found an unannotated SNP (rs750793297) with the same effects on the TBP-promoter complex. Additionally, near this known SNP marker, we found three unannotated SNPs (rs766799008, rs764508464, and rs754122321) that can cause DHFR underexpression (Table 1). According to our recent paper [40], these SNPs can elevate an apparent bioactivity of methotrexate-based antitumor chemotherapy [13, 75].

The StAR gene encodes steroidogenic acute regulatory protein and contains an SNP marker (rs16887226) of hypertension in diabetes (as a result of lowered expression of this gene because of impaired binding of its promoter with an unknown transcription factor, not TBP) [76]. A keyword search produced associations with lowered resistance to endotoxins for underexpression of the StAR protein, which is a mediator of mutual synchronicity of the immune system and circadian system [11]. Near this known SNP marker, we found the unannotated SNP rs544850971, which can lower *StAR* expression (Table 1) and therefore can be a candidate SNP marker of the above-mentioned disorders.

Gene CETP codes for cholesterol ester transfer protein; in its promoter, it contains a known biomedical SNP marker: deletion of the region $G_{-72}GGCGGACATACATATAC_{-54}$ (18 bp long) at position -54 relative to the transcription start site (hereafter: -54[18 bp]DEL); this is a marker of hyperalphalipoproteinemia that lowers the risk of atherosclerosis [77, 78]. A keyword search uncovered clinical data on circadian pathogenesis (postprandial flare-up) of this disorder in diabetes [79]. Near this known SNP marker, we found three unannotated SNPs (rs17231520, rs757176551, and rs569033466), which can increase *CETP* expression (Table 1) and thereby increase the risk of atherosclerosis [77–79] and of hypoalphalipoproteinemia which causes hepatic chronopathologies [80].

The APOA1 gene encodes apolipoprotein A1; in its promoter, some researchers [81] identified an SNP marker (-35A→C) of hematuria, hepatic steatosis, and obesity and of

hypoalphalipoproteinemia which impairs the peripheral circadian clock in the liver [80]. A keyword search yielded some data on a knockout mouse model (APOA1$^{-/-}$) regarding the risk of atherosclerosis [78] which develops in postprandial flare-ups in diabetes [79]. For this reason, we propose the SNP "APOA1: -31A→C" as a candidate marker of this chronopathology.

Gene CYP2B6 encodes cytochrome P450 2B6 and contains a known SNP marker (rs34223104) of improved bioactivation of cyclophosphamide [82] with a circadian therapeutic optimum [14]. According to empirical and computational data [82], this SNP disrupts a major variant of the TBP-binding site in the CYP2B6 promoter and in its place creates a binding site for the transcription factor (activator) C/EBP; this change shifts the TBP-binding site and transcription start by 30 bp in the $5' \rightarrow 3'$ direction and turns them into their minor alternative variants. In close proximity to this known SNP marker, we found the unannotated SNP rs563558831, which, in the same manner, lowers TBP's affinity for this promoter (Table 1) and therefore can be a candidate SNP marker of the same chronopathology.

The INS gene encodes insulin, and its promoter contains a known SNP marker (rs5505) of neonatal diabetes and hyperinsulinemia [24]. A keyword search uncovered hyperinsulinemia as a biochemical marker of aberrations in the circadian rhythms of (i) the reproductive system [83], (ii) blood pressure [84], and (iii) the tumor-host balance [85]. Near this known SNP marker, we found unannotated rs563207167, which can also cause hyperinsulinemia and therefore can be a candidate SNP marker of the same chronopathologies (Table 1). In addition, here we found unannotated rs11557611, which can cause hypoinsulinemia (Table 1). A keyword search showed that hypoinsulinemia is a biochemical marker of hypothalamic amenorrhea [86]. Consequently, rs1155761 may serve as a candidate SNP marker of this chronopathology (Table 1).

Gene ESR2 codes for estrogen receptor 2 (β) and, in its promoter, contains a known SNP marker (rs35036378) for prophylactic treatment (with tamoxifen) of an ESR2-deficient primary tumor pT1 [87] to prevent progression to breast cancer [88]; this treatment is characterized by a circadian optimum for its use [89]. A keyword search yielded basic research findings of circadian disturbances of daytime behavioral activity in ESR2-deficient female mice [90]. Near this known SNP marker, we found an unannotated SNP (rs35036378) with the same effects on the TBP-promoter affinity.

3.2. The Results on Candidate SNP Markers within the Circadian Clock Core. These results are shown in Table 2. Let us review in more detail the data in this table using the *PER1* gene as an example, which encodes a protein called period 1—a subunit of the heterodimeric PER-CRY complex—which is the main negative component of the circadian clock core: this complex inhibits the activity of transcription factor CLOCK/ARNTL [92–98].

As predicted by SNP_TATA_Comparator [35], only five of the 28 SNPs (that are known in the 90 bp proximal promoter regions for various protein-coding transcripts of this gene

[24]) can affect the affinity of TBP for its promoters: rs137890200, rs773740924, and rs2518024 can enhance the TBP-promoter affinity, whereas rs796629786 and rs3027175 can reduce it. A keyword search showed that strong expression of the *PER1* gene inhibits the proliferation of tumor cells [17, 99, 100]; for example, in patients with strong expression of this gene, if they have a gastric cancer, longer survival is observed [99]. This gene is studied as a tumor suppressor; one of its mechanisms of action is the influence on the sensitivity of cells to DNA damage-induced apoptosis [17, 101]. Downregulation of PER1 was detected in human tissues of malignant tumors of the stomach and prostate [100, 101]. It should also be noted that, in studies of knockout mouse models (*PER1*$^{-/-}$), researchers observed impairment of spatial (3D) learning capacity and enhanced manifestations of ethanol hepatotoxicity [102, 103]. Therefore, we can hypothesize that rs137890200, rs773740924, rs2518024, rs796629786, and rs3027175 of the *PER1* gene are candidate SNP markers, as we propose in Table 2. One can see similar results for the other genes of the human circadian clock core [92–114] in this table.

Using SNP_TATA_Comparator [35], we analyzed 231 SNPs within 90 bp proximal promoter regions for the protein-coding transcripts of 12 genes of the human circadian clock core; only 52 of these SNPs (22%) were found to be capable of statistically significant changes in the affinity of TBP for promoters of these genes. As one can see in Table 2, we failed to find candidate SNP markers of chronopathologies for only one of the 12 genes, namely, *NR1D1*. This result shows that preliminary computational (bioinformatic) analysis of unannotated SNPs from the 1000 Genomes Project can indeed accelerate and cheapen the search for biomedical SNP markers because of selection (for this expensive and labor-intensive procedure) of only those candidate markers whose molecular mechanisms of pathological manifestation are easily understandable within the framework of existing clinical observations, genetic knowledge, scientific theories, hypotheses, and empirical data from animal and cellular models of human diseases.

It is also worth noting that only 13 of the 52 candidate SNP markers identified here decrease affinity of TBP for promoters of the genes of the circadian clock core, whereas the other 39 SNPs enhance it. In Table 1, however, one can see the opposite distribution of the candidate SNP markers (identified here) of circadian complications of Mendelian diseases: the majority (26 of 41, 62%) of the candidate SNP markers significantly reduce affinity of TBP for the human gene promoters, whereas the remaining 15 SNPs enhance it, as predicted by SNP_TATA_Comparator [35]. This difference is statistically significant ($p < 0.0005$) according to Fisher's exact test for 2×2 design. It is noteworthy (Table 1) that the ratio of the prevalence of candidate SNP markers of increased versus decreased affinity TBP-promoter is in agreement with independent studies by other investigators [127, 128]. Indeed, overall, in the reference human genome, the proportion of SNPs which weaken the binding sites of transcription factors is significantly greater than the share of SNPs which enhance this binding [127]. Similarly, some researchers [128] reported that SNPs of the binding sites for transcription factor NF-κB or

RNA polymerase II (significantly more often) weaken rather than enhance the binding of these proteins to the mutated DNA in comparison with the reference genome. Taken together, these findings suggest that the reduced proportion of candidate SNP markers weakening the affinity TBP-promoter may be a specific characteristic of the 12 genes of the human circadian clock core. This phenomenon may reflect the pressure of natural selection for robustness of their functioning under the conditions of incessant genetic variability of the promoter region being analyzed.

Why is the robustness of the circadian clock core so important for humans? As shown in Table 2, overall, dysregulation of these genes' expression may be a marker of a wide range of pathological conditions in humans, for example, cancer, neurodegenerative disorders, lung diseases, and cardiovascular diseases. The reason for such diversity of chronopathologies is that the circadian clock synchronizes a large number of molecular biological and biochemical processes on the whole-body level and integrates various individual signals from each cell, tissue, and organ into a united hierarchical system of circadian rhythms of the human body.

3.3. How to Use Candidate SNP Markers of Chronopathologies. In this work, we used SNP_TATA_Comparator [35] to analyze 484 SNPs within 90 bp proximal promoter regions for protein-coding transcripts of human genes. Only 53 of these SNPs (11%) were found to be candidate SNP markers of chronopathologies (Tables 1 and 2). This finding does not mean that the remaining 431 of the 484 SNPs (89%) cannot be SNP markers of some human diseases. This is because each of these SNPs may influence a specific promoter-related nucleosome [129], DNA methylation sites in promoters, binding sites for histone modifications, and binding sites for transcription factors (e.g., rs16887226 and rs34223104). At present, there is a large number and variety of freely available Web services [130–149]. Most of them rank unannotated SNPs by their generalized statistical similarity with biomedical SNP markers of human diseases; these Web services evaluate this similarity by superimposing SNPs on gene maps and on data from massively parallel high-throughput sequencing of chromatin immunoprecipitation material (ChIP-Seq) from experiments with complexes of various proteins with genomic DNA. Accuracy of such assessments is constantly increasing due to improvements in empirical formulas for whole-genome evaluation of similarity among pathological manifestations of various SNPs and due to the increasing diversity, completeness, and number of whole-genome maps for various epigenetic states of cells from various tissues and organs in health [150], during infection [151] (or other diseases [152]), or after treatment [153], as we predicted [154] on the basis of the Central Limit Theorem.

As an unexpected clever generalization of this mainstream approach, the authors of Web server GenomeRunner [155] proposed to evaluate the difference between SNPs in addition to the widely accepted notion of assessments of the similarity between them. In this active field of research, the new trend is creation of Web navigation services that help users generate their own hypotheses and ideas regarding how the SNP of interest can affect the signs and symptoms of

diseases under study [156]. Another innovation that emerged here is Web service PredictSNP2 for translation from the numerical predictions to an effect of an SNP on human health which is suitable for precise computer calculations in qualitative categories that are accessible to the general population [157]. These breakthroughs mean that SNP-related predictions are becoming interesting not only to narrow specialists who treat patients with one or another disease but also to anyone who is willing to customize their lifestyle to minimize the risk of diseases.

Because statistical significance of our predicted candidate SNP markers (Tables 1 and 2) varies from high ($\alpha < 10^{-6}$) to minimally acceptable ($\alpha < 0.05$), the proposed markers should be properly validated using clinical standards before practical use. The results of this validation are dependent on climate, environmental conditions, and lifestyles and on the ethnic, social, age, and gender composition of cohorts [158]. Accordingly, we arranged the ancestral and minor alleles of each of candidate SNP markers of chronopathologies by the predicted K_D values of TBP-DNA affinity *in vitro* [91]. As shown in Tables 1 and 2, these K_D values vary from 1 to 335 nM, whereas the extent of their variation among alleles of a given SNP may be 1 nM, which is less than 0.3% of the K_D range. This level of allelic variations is too small for empirical measurement without an *a priori* known, fairly narrow range of K_D values to be measured. Thus, the predicted K_D values (Tables 1 and 2) are an integral part of each candidate SNP marker; without these data, an SNP marker cannot be validated in practice.

Finally, pathological manifestation of SNP markers of Mendelian diseases, as a rule, is limited to the consequences of changes in the expression of only those genes that contain these SNPs and can be useful only to physicians of the narrow specialties relevant to the diseases in question. Nonetheless, candidate SNP markers of chronopathologies are associated with consequences of desynchronoses either among the nervous, immune, digestive, respiratory, and other systems of the human body or between the human body and its environment (Tables 1 and 2). These data can be useful both for physicians and for the general population. For instance, the candidate SNP marker rs568650510 may be associated with an elevated risk of asthma whose symptoms are circadian (worse at night [51]; Table 2). Using this information, a physician can select the treatment timing (for asthma symptoms in a patient with minor alleles of these SNPs) that could reduce the risk of aggravation at night. By the same token, any person with the minor allele -15T of this SNP can choose a lifestyle that can reduce the systematic nocturnal influence of the environmental factor that causes the asthma symptoms. Similarly, rs367732974, rs549591993, rs192518038, and rs537333415 may help reduce the risk of a heart attack [69]; rs374778785 may be useful for lowering the risk of emphysema and lung cancer among smokers [106], whereas rs2899302 may help decide whether to use opioids [113].

4. Conclusions

Here, we predicted candidate SNP markers of chronopathologies (Tables 1 and 2); these SNPs can change affinity of

TATA-binding protein for human gene promoters. After proper validation of these candidate markers in accordance with clinical standards, these SNPs may turn out to be useful both for physicians (to select the best treatment for each patient) and for the general population (to choose a lifestyle preventing possible circadian comorbidities and complications).

Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential competing interests.

Acknowledgments

References

[1] R. Zhang, N. F. Lahens, H. I. Ballance, M. E. Hughes, and J. B. Hogenesch, "A circadian gene expression atlas in mammals: implications for biology and medicine," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 45, pp. 16219–16224, 2014.

[2] S. M. Reppert and D. R. Weaver, "Molecular analysis of mammalian circadian rhythms," *Annual Review of Physiology*, vol. 63, pp. 647–676, 2001.

[3] S. A. Brown, E. Kowalska, and R. Dallmann, "(Re)inventing the circadian feedback loop," *Developmental Cell*, vol. 22, no. 3, pp. 477–487, 2012.

[4] J. K. Kim and D. B. Forger, "A mechanism for robust circadian timekeeping via stoichiometric balance," *Molecular Systems Biology*, vol. 8, no. 1, article 630, 2012.

[5] L. Chen and G. Yang, "PPARs integrate the mammalian clock and energy metabolism," *PPAR Research*, vol. 2014, Article ID 653017, 6 pages, 2014.

[6] K. Padmanabhan, M. S. Robles, T. Westerling, and C. J. Weitz, "Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex," *Science*, vol. 337, no. 6094, pp. 599–602, 2012.

[7] K. Eckel-Mahan and P. Sassone-Corsi, "Epigenetic regulation of the molecular clockwork," *Progress in Molecular Biology and Translational Science*, vol. 119, pp. 29–50, 2013.

[8] K. Bozek, A. Relógio, S. M. Kielbasa et al., "Regulation of clock-controlled genes in mammals," *PLoS ONE*, vol. 4, no. 3, Article ID e4882, 2009.

[9] K. Bozek, A. L. Rosahl, S. Gaub, S. Lorenzen, and H. Herzel, "Circadian transcription in liver," *BioSystems*, vol. 102, no. 1, pp. 61–69, 2010.

[10] J. S. Menet, S. Pescatore, and M. Rosbash, "CLOCK: BMAL1 is a pioneer-like transcription factor," *Genes and Development*, vol. 28, no. 1, pp. 8–13, 2014.

[11] J. Wang, Y. Luo, K. Wang et al., "Clock-controlled StAR's expression and corticosterone production contribute to the endotoxemia immune response," *Chronobiology International*, vol. 32, no. 3, pp. 358–367, 2015.

[12] P. Marckmann, B. Sandström, and J. Jespersen, "Dietary effects on circadian fluctuation in human blood coagulation factor VII and fibrinolysis," *Atherosclerosis*, vol. 101, no. 2, pp. 225–234, 1993.

[13] S. Ohdo, K. Inoue, E. Yukawa, S. Higuchi, S. Nakano, and N. Ogawa, "Chronotoxicity of methotrexate in mice and its relation to circadian rhythm of DNA synthesis and pharmacokinetics," *Japanese Journal of Pharmacology*, vol. 75, no. 3, pp. 283–290, 1997.

[14] V. Y. Gorbacheva, R. V. Kondratov, R. Zhang et al., "Circadian sensitivity to the chemotherapeutic agent cyclophosphamide depends on the functional status of the CLOCK/BMAL1 transactivation complex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 9, pp. 3407–3412, 2005.

[15] C. H. Ko and J. S. Takahashi, "Molecular components of the mammalian circadian clock," *Human Molecular Genetics*, vol. 15, no. 2, pp. R271–R277, 2006.

[16] S. Sahar and P. Sassone-Corsi, "Regulation of metabolism: the circadian clock dictates the time," *Trends in Endocrinology and Metabolism*, vol. 23, no. 1, pp. 1–8, 2012.

[17] N. M. Kettner, C. A. Katchy, and L. Fu, "Circadian gene variants in cancer," *Annals of Medicine*, vol. 46, no. 4, pp. 208–220, 2014.

[18] O. A. Podkolodnaya, "Molecular and genetic aspects of interactions of the circadian clock and the energy-producing substrate metabolism in mammals," *Russian Journal of Genetics*, vol. 50, no. 2, pp. 111–122, 2014.

[19] G. V. Vasiliev, V. M. Merkulov, V. F. Kobzev, T. I. Merkulova, M. P. Ponomarenko, and N. A. Kolchanov, "Point mutations within 663-666 bp of intron 6 of the human TDO2 gene, associated with a number of psychiatric disorders, damage the YY-1 transcription factor binding site," *FEBS Letters*, vol. 462, no. 1-2, pp. 85–88, 1999.

[20] J. V. Ponomarenko, T. I. Merkulova, G. V. Vasiliev et al., "rSNP_Guide, a database system for analysis of transcription factor binding to target sequences: application to SNPs and site-directed mutations," *Nucleic Acids Research*, vol. 29, no. 1, pp. 312–316, 2001.

[21] J. V. Ponomarenko, G. V. Orlova, T. I. Merkulova et al., "rSNP_Guide: an integrated database-tools system for studying SNPs and site-directed mutations in transcription factor binding sites," *Human Mutation*, vol. 20, no. 4, pp. 239–248, 2002.

[22] J. V. Ponomarenko, T. I. Merkulova, G. V. Orlova et al., "rSNP Guide, a database system for analysis of transcription factor binding to DNA with variations: application to genome annotation," *Nucleic Acids Research*, vol. 31, no. 1, pp. 118–121, 2003.

[23] O. Delaneau, J. Marchini, and 1000 Genomes Project Consortium, "Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel," *Nature Communications*, vol. 5, article 3934, 2014.

[24] S. T. Sherry, M.-H. Ward, M. Kholodov et al., "DbSNP: the NCBI database of genetic variation," *Nucleic Acids Research*, vol. 29, no. 1, pp. 308–311, 2001.

[25] M. Haeussler, B. J. Raney, A. S. Hinrichs et al., "Navigating protected genomics data with UCSC Genome Browser in a box," *Bioinformatics*, vol. 31, no. 5, pp. 764–766, 2015.

[26] D. R. Zerbino, S. P. Wilder, N. Johnson, T. Juettemann, and P. R. Flicek, "The ensembl regulatory build," *Genome Biology*, vol. 16, no. 1, article 56, 2015.

[27] A. Abbas, M. Lechevrel, and F. Sichel, "Identification of new single nucleotid polymorphisms (SNP) in alcohol dehydrogenase class IV ADH7 gene within a French population," *Archives of Toxicology*, vol. 80, no. 4, pp. 201–205, 2006.

[28] J. S. Amberger, C. A. Bocchini, F. Schiettecatte, A. F. Scott, and A. Hamosh, "OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders," *Nucleic Acids Research*, vol. 43, no. 1, pp. D789–D798, 2015.

[29] H. Mitsuyasu, K. Izuhara, X. Q. Mao et al., "Ile50Val variant of IL4R alpha upregulates IgE synthesis and associates with atopic asthma," *Nature Genetics*, vol. 19, no. 2, pp. 119–120, 1998.

[30] V. Rajendran, "Structural analysis of oncogenic mutation of isocitrate dehydrogenase 1," *Molecular BioSystems*, vol. 12, no. 7, pp. 2276–2287, 2016.

[31] M. Lopus, D. M. Paul, and R. Rajasekaran, "Unraveling the deleterious effects of cancer-driven stk11 mutants through conformational sampling approach," *Cancer Informatics*, vol. 15, pp. 35–44, 2016.

[32] D. Meshach Paul and R. Rajasekaran, "Exploration of structural and functional variations owing to point mutations in α-NAGA," *Interdisciplinary Sciences: Computational Life Sciences*, 2016.

[33] V. Rajendran and R. Sethumadhavan, "Drug resistance mechanism of PncA in Mycobacterium tuberculosis," *Journal of Biomolecular Structure and Dynamics*, vol. 32, no. 2, pp. 209–221, 2014.

[34] B. Senthilkumar and R. Rajasekaran, "Analysis of the structural stability among cyclotide members through cystine knot fold that underpins its potential use as a drug scaffold," *International Journal of Peptide Research and Therapeutics*, pp. 1–11, 2016.

[35] M. Ponomarenko, D. Rasskazov, O. Arkova et al., "How to use SNP_TATA_comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter," *BioMed Research International*, vol. 2015, Article ID 359835, 17 pages, 2015.

[36] L. K. Savinkova, M. P. Ponomarenko, P. M. Ponomarenko et al., "TATA box polymorphisms in human gene promoters and associated hereditary pathologies," *Biochemistry*, vol. 74, no. 2, pp. 117–129, 2009.

[37] M. Ponomarenko, V. Mironova, K. Gunbin, and L. Savinkova, "Hogness box," in *Brenner's Encyclopedia of Genetics*, S. Maloy and K. Hughes, Eds., vol. 3, pp. 491–494, Academic Press, Elsevier Inc, San Diego, Calif, USA, 2nd edition, 2013.

[38] O. V. Arkova, M. P. Ponomarenko, D. A. Rasskazov et al., "Obesity-related known and candidate SNP markers can significantly change affinity of TATA-binding protein for human gene promoters," *BMC Genomics*, vol. 16, supplement 13, article S5, 2015.

[39] M. P. Ponomarenko, O. Arkova, D. Rasskazov, P. Ponomarenko, L. Savinkova, and N. Kolchanov, "Candidate SNP markers of gender-biased autoimmune complications of monogenic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters," *Frontiers in Immunology*, vol. 7, article 130, 2016.

[40] I. I. Turnaev, D. A. Rasskazov, O. V. Arkova et al., "Hypothetical SNP markers that significantly affect the affinity of the TATA-binding protein to *VEGFA, ERBB2, IGF1R, FLT1, KDR*, and *MET* oncogene promoters as chemotherapy targets," *Molecular Biology*, vol. 50, no. 1, pp. 141–152, 2016.

[41] G. M. Trovato, "Sustainable medical research by effective and comprehensive medical skills: overcoming the frontiers by predictive, preventive and personalized medicine," *EPMA Journal*, vol. 5, no. 1, article 14, 2014.

[42] R. M. Bannerman, L. M. Garrick, P. Rusnak-Smalley, J. E. Hoke, and J. A. Edwards, "Hemoglobin deficit: an inherited hypochromic anemia in the mouse," *Proceedings of the Society for Experimental Biology and Medicine*, vol. 182, no. 1, pp. 52–57, 1986.

[43] E. L. Unger, C. J. Earley, and J. L. Beard, "Diurnal cycle influences peripheral and brain iron levels in mice," *Journal of Applied Physiology*, vol. 106, no. 1, pp. 187–193, 2009.

[44] A.-H. Sun, Z.-M. Wang, S.-Z. Xiao, Z.-J. Li, J.-Y. Li, and L.-S. Kong, "Red cell basic ferritin concentration in sensorineural hearing loss," *ORL Journal for Oto-Rhino-Laryngology and Its Related Specialties*, vol. 53, no. 5, pp. 270–272, 1991.

[45] R. P. Allen, S. Auerbach, H. Bahrain, M. Auerbach, and C. J. Earley, "The prevalence and impact of restless legs syndrome on patients with iron deficiency anemia," *American Journal of Hematology*, vol. 88, no. 4, pp. 261–264, 2013.

[46] S. Verma, R. Gupta, M. Kudesia, A. Mathur, G. Krishan, and S. Singh, "Coexisting iron deficiency anemia and Beta thalassemia trait: effect of iron therapy on red cell parameters and hemoglobin subtypes," *ISRN Hematology*, vol. 2014, Article ID 293216, 5 pages, 2014.

[47] D. Thio, V. Prasad, P. Anslow, and P. Lennox, "Marrow proliferation as a cause of hearing loss in beta-thalassaemia major," *The Journal of Laryngology and Otology*, vol. 122, no. 11, pp. 1253–1256, 2008.

[48] M. Manetti, L. Ibba-Manneschi, C. Fatini et al., "Association of a functional polymorphism in the matrix metalloproteinase-12 promoter region with systemic sclerosis in an Italian population," *Journal of Rheumatology*, vol. 37, no. 9, pp. 1852–1857, 2010.

[49] N. L. Starodubtseva, V. V. Sobolev, A. G. Soboleva, A. A. Nikolaev, and S. A. Bruskin, "Genes expression of metalloproteinases (MMP-1, MMP-2, MMP-9, and MMP-12) associated with psoriasis," *Russian Journal of Genetics*, vol. 47, no. 9, pp. 1117–1123, 2011.

[50] G. M. Hunninghake, M. H. Cho, Y. Tesfaigzi et al., "MMP12, lung function, and COPD in high-risk populations," *The New England Journal of Medicine*, vol. 361, no. 27, pp. 2599–2608, 2009.

[51] H. J. Durrington, S. N. Farrow, A. S. Loudon, and D. W. Ray, "The circadian clock and asthma," *Thorax*, vol. 69, no. 1, pp. 90–92, 2014.

[52] E. M. El-Omar, M. Carrington, W.-H. Chow et al., "Interleukin-1 polymorphisms associated with increased risk of gastric cancer," *Nature*, vol. 404, no. 6776, pp. 398–402, 2000.

[53] Y. Wang, N. Kato, Y. Hoshida et al., "Interleukin-1β gene polymorphisms associated with hepatocellular carcinoma in hepatitis C virus infection," *Hepatology*, vol. 37, no. 1, pp. 65–71, 2003.

[54] K.-S. Wu, X. Zhou, F. Zheng, X.-Q. Xu, Y.-H. Lin, and J. Yang, "Influence of interleukin-1 beta genetic polymorphism, smoking and alcohol drinking on the risk of non-small cell lung cancer," *Clinica Chimica Acta*, vol. 411, no. 19-20, pp. 1441–1446, 2010.

[55] D. N. Martínez-Carrillo, E. Garza-González, R. Betancourt-Linares et al., "Association of IL1B -511C/-31T haplotype and *Helicobacter pylori* vacA genotypes with gastric ulcer and chronic gastritis," *BMC Gastroenterology*, vol. 10, article 126, 2010.

[56] F. Hayashi, M. Watanabe, T. Nanba, N. Inoue, T. Akamizu, and Y. Iwatani, "Association of the -31C/T functional polymorphism in the interleukin-1β gene with the intractability of Graves' disease and the proportion of T helper type 17 cells," *Clinical and Experimental Immunology*, vol. 158, no. 3, pp. 281–286, 2009.

[57] L. Strandberg, D. Mellström, Ö. Ljunggren et al., "IL6 and IL1B polymorphisms are associated with fat mass in older men: The MrOS Study Sweden," *Obesity*, vol. 16, no. 3, pp. 710–713, 2008.

[58] P. Borkowska, K. Kucia, S. Rzezniczek et al., "Interleukin-1beta promoter (-31T/C and -511C/T) polymorphisms in major recurrent depression," *Journal of Molecular Neuroscience*, vol. 44, no. 1, pp. 12–16, 2011.

[59] C. Ávila Moraes, T. Cambras, A. Diez-Noguera et al., "A new chronobiological approach to discriminate between acute and chronic depression using peripheral temperature, rest-activity, and light exposure parameters," *BMC Psychiatry*, vol. 13, article 77, 2013.

[60] O. Pivovarova, K. Jürchott, N. Rudovich et al., "Changes of dietary fat and carbohydrate content alter central and peripheral clock in humans," *Journal of Clinical Endocrinology and Metabolism*, vol. 100, no. 6, pp. 2291–2302, 2015.

[61] C. J. Carter, "Multiple genes and factors associated with bipolar disorder converge on growth factor and stress activated kinase pathways controlling translation initiation: implications for oligodendrocyte viability," *Neurochemistry International*, vol. 50, no. 3, pp. 461–490, 2007.

[62] H. Yamazaki, M. Takeoka, M. Kitazawa et al., "ASC plays a role in the priming phase of the immune response to type II collagen in collagen-induced arthritis," *Rheumatology International*, vol. 32, no. 6, pp. 1625–1632, 2012.

[63] I. C. Chikanza, P. Petrou, G. Kingsley, G. Chrousos, and G. S. Panayi, "Defective hypothalamic response to immune and inflammatory stimuli in patients with rheumatoid arthritis," *Arthritis and Rheumatism*, vol. 35, no. 11, pp. 1281–1288, 1992.

[64] E. Arnaud, V. Barbalat, V. Nicaud et al., "Polymorphisms in the 5′ regulatory region of the tissue factor gene and the risk of myocardial infarction and venous thromboembolism: the ECTIM and PATHROS studies," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 20, no. 3, pp. 892–898, 2000.

[65] E. Haus, "Chronobiology of hemostasis and inferences for the chronotherapy of coagulation disorders and thrombosis prevention," *Advanced Drug Delivery Reviews*, vol. 59, no. 9-10, pp. 966–984, 2007.

[66] K. Oishi, S. Koyanagi, and N. Ohkura, "Circadian mRNA expression of coagulation and fibrinolytic factors is organ-dependently disrupted in aged mice," *Experimental Gerontology*, vol. 46, no. 12, pp. 994–999, 2011.

[67] A. Kavlie, L. Hiltunen, V. Rasi, and H. P. B. Prydz, "Two novel mutations in the human coagulation factor VII promoter," *Thrombosis and Haemostasis*, vol. 90, no. 2, pp. 194–205, 2003.

[68] I. Colognesi, V. Pasquali, A. Foà et al., "Temporal variations of coagulation factor VII activity in mice are influenced by lighting regime," *Chronobiology International*, vol. 24, no. 2, pp. 305–313, 2007.

[69] J. Carvalho De Sousa, E. Bruckert, P. Giral et al., "Coagulation factor VII and plasma triglycerides. Decreased catabolism as a possible mechanism of factor VII hyperactivity," *Haemostasis*, vol. 19, no. 3, pp. 125–130, 1989.

[70] P. Marckmann, B. Sandström, and J. Jespersen, "Dietary effects on circadian fluctuation in human blood coagulation factor VII and fibrinolysis," *Atherosclerosis*, vol. 101, no. 2, pp. 225–234, 1993.

[71] I. A. Clark, K. A. Rockett, and D. Burgner, "Genes, nitric oxide and malaria in African children," *Trends in Parasitology*, vol. 19, no. 8, pp. 335–337, 2003.

[72] J. A. González-Martínez, G. Möddel, Z. Ying, R. A. Prayson, W. E. Bingaman, and I. M. Najm, "Neuronal nitric oxide synthase expression in resected epileptic dysplastic neocortex," *Journal of Neurosurgery*, vol. 110, no. 2, pp. 343–349, 2009.

[73] W. A. Hofstra and A. W. de Weerd, "The circadian rhythm and its interaction with human epilepsy: a review of literature," *Sleep Medicine Reviews*, vol. 13, no. 6, pp. 413–420, 2009.

[74] B. Kaya, S. Ünal, A. B. Karabulut, and Y. Türköz, "Altered diurnal variation of nitric oxide production in patients with panic disorder," *Tohoku Journal of Experimental Medicine*, vol. 204, no. 2, pp. 147–154, 2004.

[75] F. Al-Shakfa, S. Dulucq, I. Brukner et al., "DNA variants in region for noncoding interfering transcript of Dihydrofolate reductase gene and outcome in childhood acute lymphoblastic leukemia," *Clinical Cancer Research*, vol. 15, no. 22, pp. 6931–6938, 2009.

[76] A. J. Casal, V. J. P. Sinclair, A. M. Capponi, J. Nicod, U. Huynh-Do, and P. Ferrari, "A novel mutation in the steroidogenic acute regulatory protein gene promoter leading to reduced promoter activity," *Journal of Molecular Endocrinology*, vol. 37, no. 1, pp. 71–80, 2006.

[77] W. Plengpanich, W. Le Goff, S. Poolsuk, Z. Julia, M. Guerin, and W. Khovidhunkit, "CETP deficiency due to a novel mutation in the CETP gene promoter and its effect on cholesterol efflux and selective uptake into hepatocytes," *Atherosclerosis*, vol. 216, no. 2, pp. 370–373, 2011.

[78] K. Oka, L. M. Belalcazar, C. Dieker et al., "Sustained phenotypic correction in a mouse model of hypoalphalipoproteinemia with a helper-dependent adenovirus vector," *Gene Therapy*, vol. 14, no. 3, pp. 191–202, 2007.

[79] S. Hirayama, S. Soda, Y. Ito et al., "Circadian change of serum concentration of small dense LDL-cholesterol in type 2 diabetic patients," *Clinica Chimica Acta*, vol. 411, no. 3-4, pp. 253–257, 2010.

[80] C. Gabás-Rivera, R. Martínez-Beamonte, J. L. Ríos et al., "Dietary oleanolic acid mediates circadian clock gene expression in liver independently of diet and animal model but requires apolipoprotein A1," *Journal of Nutritional Biochemistry*, vol. 24, no. 12, pp. 2100–2109, 2013.

[81] A. Matsunaga, J. Sasaki, H. Han et al., "Compound heterozygosity for an apolipoprotein A1 gene promoter mutation and a structural nonsense mutation with apolipoprotein A1 deficiency," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 19, no. 2, pp. 348–355, 1999.

[82] J. Zukunft, T. Lang, T. Richter et al., "A natural *CYP2B6* TATA box polymorphism (-82T→C) leading to enhanced transcription and relocation of the transcriptional start site," *Molecular Pharmacology*, vol. 67, no. 5, pp. 1772–1782, 2005.

[83] A. L. Mereness, Z. C. Murphy, and M. T. Sellix, "Developmental programming by androgen affects the circadian timing system in female mice1," *Biology of Reproduction*, vol. 92, no. 4, article 88, 2015.

[84] S. Bianchi, R. Bigazzi, R. Nenci, and V. M. Campese, "Hyperinsulinemia, circadian variation of blood pressure and end-organ damage in hypertension," *Journal of Nephrology*, vol. 10, no. 6, pp. 325–333, 1997.

[85] D. E. Blask, R. T. Dauchy, E. M. Dauchy et al., "Light exposure at night disrupts host/cancer circadian regulatory dynamics: impact on the Warburg effect, lipid signaling and tumor growth prevention," *PLoS ONE*, vol. 9, no. 8, Article ID e102776, 2014.

[86] G. A. Laughlin, C. E. Dominguez, and S. S. C. Yen, "Nutritional and endocrine-metabolic aberrations in women with functional hypothalamic amenorrhea," *Journal of Clinical Endocrinology and Metabolism*, vol. 83, no. 1, pp. 25–32, 1998.

[87] A. M. Sieuwerts, M. Ansems, M. P. Look et al., "Clinical significance of the nuclear receptor co-regulator DC-SCRIPT in breast cancer: an independent retrospective validation study," *Breast Cancer Research*, vol. 12, no. 6, article R103, 2010.

[88] S. Philips, A. Richter, S. Oesterreich et al., "Functional characterization of a genetic polymorphism in the promoter of the ESR2 gene," *Hormones and Cancer*, vol. 3, no. 1-2, pp. 37–43, 2012.

[89] L. Binkhorst, J. S. L. Kloth, A. S. de Wit et al., "Circadian variation in tamoxifen pharmacokinetics in mice and breast cancer patients," *Breast Cancer Research and Treatment*, vol. 152, no. 1, pp. 119–128, 2015.

[90] S. E. Royston, N. Yasui, A. G. Kondilis, S. V. Lord, J. A. Katzenellenbogen, and M. M. Mahoney, "ESR1 and ESR2 differentially regulate daily and circadian activity rhythms in female mice," *Endocrinology*, vol. 155, no. 7, pp. 2613–2623, 2014.

[91] L. K. Savinkova, I. A. Drachkova, T. V. Arshinova, P. Ponomarenko, M. Ponomarenko, and N. Kolchanov, "An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein," *PLoS ONE*, vol. 8, no. 2, Article ID e54626, 2013.

[92] K. Oishi, N. Ohkura, N. Amagai, and N. Ishida, "Involvement of circadian clock gene Clock in diabetes-induced circadian augmentation of plasminogen activator inhibitor-1 (PAI-1) expression in the mouse heart," *FEBS Letters*, vol. 579, no. 17, pp. 3555–3559, 2005.

[93] M. Elshazley, M. Sato, T. Hase et al., "The circadian clock gene BMAL1 is a novel therapeutic target for malignant pleural mesothelioma," *International Journal of Cancer*, vol. 131, no. 12, pp. 2820–2831, 2012.

[94] C.-X. He, N. Prevot, C. Boitard, P. Avner, and U. C. Rogner, "Inhibition of type 1 diabetes by upregulation of the circadian rhythm-related aryl hydrocarbon receptor nuclear translocator-like 2," *Immunogenetics*, vol. 62, no. 9, pp. 585–592, 2010.

[95] C.-X. He, P. Avner, C. Boitard, and U. C. Rogner, "Downregulation of the circadian rhythm related gene Arntl2 suppresses diabetes protection in Idd6 NOD.C3H congenic mice," *Clinical and Experimental Pharmacology and Physiology*, vol. 37, no. 12, pp. 1154–1158, 2010.

[96] A. Hashiramoto, T. Yamane, K. Tsumiyama et al., "Mammalian clock gene *Cryptochrome* regulates arthritis via proinflammatory cytokine TNF-α," *The Journal of Immunology*, vol. 184, no. 3, pp. 1560–1565, 2010.

[97] G. Savalli, W. Diao, S. Berger, M. Ronovsky, T. Partonen, and D. D. Pollak, "Anhedonic behavior in cryptochrome 2-deficient mice is paralleled by altered diurnal patterns of amygdala gene expression," *Amino Acids*, vol. 47, no. 7, pp. 1367–1377, 2015.

[98] L. Fang, Z. Yang, J. Zhou et al., "Circadian clock gene CRY2 degradation is involved in chemoresistance of colorectal cancer," *Molecular Cancer Therapeutics*, vol. 14, no. 6, pp. 1476–1487, 2015.

[99] H. Zhao, Z.-L. Zeng, J. Yang et al., "Prognostic relevance of Period1 (Per1) and Period2 (Per2) expression in human gastric cancer," *International Journal of Clinical and Experimental Pathology*, vol. 7, no. 2, pp. 619–630, 2014.

[100] Q. Cao, S. Gery, A. Dashti et al., "A role for the clock gene *Per1* in prostate cancer," *Cancer Research*, vol. 69, no. 19, pp. 7619–7625, 2009.

[101] S. Gery, N. Komatsu, L. Baldjyan, A. Yu, D. Koo, and H. P. Koeffler, "The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells," *Molecular Cell*, vol. 22, no. 3, pp. 375–382, 2006.

[102] A. Jilg, S. Lesny, N. Peruzki et al., "Temporal dynamics of mouse hippocampal clock gene expression support memory processing," *Hippocampus*, vol. 20, no. 3, pp. 377–388, 2010.

[103] T. Wang, P. Yang, Y. Zhan, L. Xia, Z. Hua, and J. Zhang, "Deletion of circadian gene Per1 alleviates acute ethanol-induced hepatotoxicity in mice," *Toxicology*, vol. 314, no. 2-3, pp. 193–201, 2013.

[104] V. Mehraj, J. Textoris, C. Capo, D. Raoult, M. Leone, and J.-L. Mege, "Overexpression of the per2 gene in male patients with acute Q fever," *Journal of Infectious Diseases*, vol. 206, no. 11, pp. 1768–1770, 2012.

[105] K. Miyazaki, M. Wakabayashi, Y. Hara, and N. Ishida, "Tumor growth suppression in vivo by overexpression of the circadian component, PER2," *Genes to Cells*, vol. 15, no. 4, pp. 351–358, 2010.

[106] Y. Shi, J. Cao, J. Gao et al., "Retinoic acid-related orphan receptor-α is induced in the setting of DNA damage and promotes pulmonary emphysema," *American Journal of Respiratory and Critical Care Medicine*, vol. 186, no. 5, pp. 412–419, 2012.

[107] M. Doulazmi, F. Frédéric, F. Capone, M. Becker-André, N. Delhaye-Bouchaud, and J. Mariani, "A comparative study of Purkinje cells in two RORα gene mutant mice: staggerer and RORα−/−," *Developmental Brain Research*, vol. 127, no. 2, pp. 165–174, 2001.

[108] A. Hamzaoui, H. Maalmi, A. Berraies et al., "Transcriptional characteristics of CD4 T cells in young asthmatic children: RORC and FOXP3 axis," *Journal of Inflammation Research*, vol. 4, pp. 139–146, 2011.

[109] K. Hamzaoui, A. Borhani Haghighi, I. B. Ghorbel, and H. Houman, "RORC and Foxp3 axis in cerebrospinal fluid of patients with Neuro-Behçet's Disease," *Journal of Neuroimmunology*, vol. 233, no. 1-2, pp. 249–253, 2011.

[110] M. Toyoshima, H. L. Howie, M. Imakura et al., "Functional genomics identifies therapeutic targets for MYC-driven cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 24, pp. 9545–9550, 2012.

[111] N. Rodriguez, J. Yang, K. Hasselblatt et al., "Casein kinase I epsilon interacts with mitochondrial proteins for the growth and survival of human ovarian cancer cells," *EMBO Molecular Medicine*, vol. 4, no. 9, pp. 952–963, 2012.

[112] M. Flajolet, G. He, M. Heiman, A. Lin, A. C. Nairn, and P. Greengard, "Regulation of Alzheimer's disease amyloid-β formation by casein kinase I," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 10, pp. 4159–4164, 2007.

[113] C. D. Bryant, C. C. Parker, L. Zhou et al., "*Csnk1e* is a genetic regulator of sensitivity to psychostimulants and opioids," *Neuropsychopharmacology*, vol. 37, no. 4, pp. 1026–1035, 2012.

[114] M. C. Abba, H. Sun, K. A. Hawkins et al., "Breast cancer molecular signatures as determined by SAGE: correlation with lymph node status," *Molecular Cancer Research*, vol. 5, no. 9, pp. 881–890, 2007.

[115] J. E. Stajich, D. Block, K. Boulez et al., "The Bioperl toolkit: perl modules for the life sciences," *Genome Research*, vol. 12, no. 10, pp. 1611–1618, 2002.

[116] P. M. Ponomarenko, L. K. Savinkova, I. A. Drachkova et al., "A step-by-step model of TBP/TATA box binding allows predicting human hereditary diseases by single nucleotide polymorphism," *Doklady Biochemistry and Biophysics*, vol. 419, no. 1, pp. 88–92, 2008.

[117] V. V. Mironova, N. A. Omelyanchuk, P. M. Ponomarenko, M. P. Ponomarenko, and N. A. Kolchanov, "Specific/nonspecific binding of TBP to promoter DNA of the auxin response factor genes in plants correlated with ARFs function on gene transcription (activator/repressor)," *Doklady Biochemistry and Biophysics*, vol. 433, no. 1, pp. 191–196, 2010.

[118] S. Hahn, S. Buratowski, P. A. Sharp, and L. Guarente, "Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 15, pp. 5718–5722, 1989.

[119] M. P. Ponomarenko, J. V. Ponomarenko, A. S. Frolov et al., "Identification of sequence-dependent DNA features correlating to activity of DNA sites interacting with proteins," *Bioinformatics*, vol. 15, no. 7-8, pp. 687–703, 1999.

[120] P. Bucher, "Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences," *Journal of Molecular Biology*, vol. 212, no. 4, pp. 563–578, 1990.

[121] D. Flatters and R. Lavery, "Sequence-dependent dynamics of TATA-box binding sites," *Biophysical Journal*, vol. 75, no. 1, pp. 372–381, 1998.

[122] M. P. Ponomarenko, L. K. Savinkova, Y. V. Ponomarenko, A. E. Kel', I. I. Titov, and N. A. Kolchanov, "Simulation of TATA box sequences in eukaryotes," *Molecular Biology*, vol. 31, no. 4, pp. 616–622, 1997.

[123] IUPAC-IUB Commission on Biochemical Nomenclature (CBN), "Abbreviations and symbols for nucleic acids, polynucleotides and their constituents," *Journal of Molecular Biology*, vol. 55, no. 3, pp. 299–310, 1971.

[124] A. J. Waardenberg, S. D. Basset, R. Bouveret, and R. P. Harvey, "CompGO: an R package for comparing and visualizing Gene Ontology enrichment differences between DNA binding experiments," *BMC Bioinformatics*, vol. 16, no. 1, article 275, 2015.

[125] I. Mogno, F. Vallania, R. D. Mitra, and B. A. Cohen, "TATA is a modular component of synthetic promoters," *Genome Research*, vol. 20, no. 10, pp. 1391–1397, 2010.

[126] I. Missala, U. Kassner, and E. Steinhagen-Thiessen, "A systematic literature review of the association of lipoprotein(a) and autoimmune diseases and atherosclerosis," *International Journal of Rheumatology*, vol. 2012, Article ID 480784, 10 pages, 2012.

[127] G. R. Abecasis, A. Auton, L. D. Brooks et al., "An integrated map of genetic variation from 1,092 human genomes," *Nature*, vol. 491, no. 7422, pp. 56–65, 2012.

[128] M. Kasowski, F. Grubert, C. Heffelfinger et al., "Variation in transcription factor binding among humans," *Science*, vol. 328, no. 5975, pp. 232–235, 2010.

[129] I. Ioshikhes, E. N. Trifonov, and M. Q. Zhang, "Periodical distribution of transcription factor sites in promoter regions and connection with chromatin structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 6, pp. 2891–2895, 1999.

[130] C.-Y. Chen, I.-S. Chang, C. A. Hsiung, and W. W. Wasserman, "On the identification of potential regulatory variants within genome wide association candidate SNP sets," *BMC Medical Genomics*, vol. 7, article 34, 2014.

[131] M. Barenboim and T. Manke, "ChroMoS: an integrated web tool for SNP classification, prioritization and functional interpretation," *Bioinformatics*, vol. 29, no. 17, pp. 2197–2198, 2013.

[132] A. Riva, "Large-scale computational identification of regulatory SNPs with rSNP-MAPPER," *BMC genomics*, vol. 13, supplement 4, article S7, 2012.

[133] J. V. Ponomarenko, G. V. Orlova, A. S. Frolov, M. S. Gelfand, and M. P. Ponomarenko, "SELEX_DB: a database on *in vitro* selected oligomers adapted for recognizing natural sites and for analyzing both SNPs and site-directed mutagenesis data," *Nucleic Acids Research*, vol. 30, no. 1, pp. 195–199, 2002.

[134] I. V. Deyneko, Y. M. Kalybaeva, A. E. Kel, and H. Blöcker, "Human-chimpanzee promoter comparisons: property-conserved evolution?" *Genomics*, vol. 96, no. 3, pp. 129–133, 2010.

[135] M. C. Andersen, P. G. Engstrom, S. Lithwick et al., "In silico detection of sequence variations modifying transcriptional regulation," *PLoS Computational Biology*, vol. 4, no. 1, article e5, 2008.

[136] C.-C. Chen, S. Xiao, D. Xie et al., "Understanding variation in transcription factor binding by modeling transcription factor genome-epigenome interactions," *PLoS Computational Biology*, vol. 9, no. 12, Article ID e1003367, 2013.

[137] G. Macintyre, J. Bailey, I. Haviv, and A. Kowalczyk, "is-rSNP: a novel technique for in silico regulatory SNP detection," *Bioinformatics*, vol. 26, no. 18, pp. i524–i530, 2010.

[138] A. P. Boyle, E. L. Hong, M. Hariharan et al., "Annotation of functional variation in personal genomes using RegulomeDB," *Genome Research*, vol. 22, no. 9, pp. 1790–1797, 2012.

[139] J. V. Ponomarenko, D. P. Furman, A. S. Frolov et al., "ACTIVITY: a database on DNA/RNA sites activity adapted to apply sequence-activity relationships from one system to another," *Nucleic Acids Research*, vol. 29, no. 1, pp. 284–287, 2001.

[140] L. O. Bryzgalov, E. V. Antontseva, M. Y. Matveeva et al., "Detection of regulatory SNPs in human genome using ChIP-seq ENCODE data," *PLoS ONE*, vol. 8, no. 10, Article ID e78833, 2013.

[141] N. L. Podkolodnyy, D. A. Afonnikov, Y. Y. Vaskin et al., "Program complex SNP-MED for analysis of single-nucleotide polymorphism (SNP) effects on the function of genes associated with socially significant diseases," *Russian Journal of Genetics: Applied Research*, vol. 4, no. 3, pp. 159–167, 2014.

[142] A. D. Johnson, R. E. Handsaker, S. L. Pulit, M. M. Nizzari, C. J. O'Donnell, and P. I. W. De Bakker, "SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap," *Bioinformatics*, vol. 24, no. 24, pp. 2938–2939, 2008.

[143] I. V. Deyneko, B. Bredohl, D. Wesely et al., "FeatureScan: revealing property-dependent similarity of nucleotide sequences," *Nucleic Acids Research*, vol. 34, pp. W591–W595, 2006.

[144] S. F. Saccone, R. Bolze, P. Thomas et al., "SPOT: a web-based tool for using biological databases to prioritize SNPs after a genome-wide association study," *Nucleic Acids Research*, vol. 38, no. 2, pp. W201–W209, 2010.

[145] Y. Fu, Z. Liu, S. Lou et al., "FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer," *Genome Biology*, vol. 15, no. 10, article 480, 2014.

[146] S. G. Coetzee, S. K. Rhie, B. P. Berman, G. A. Coetzee, and H. Noushmehr, "FunciSNP: an R/bioconductor tool integrating functional non-coding data sets with genetic association studies to identify candidate regulatory SNPs," *Nucleic Acids Research*, vol. 40, no. 18, article e139, 2012.

[147] D. A. Rasskazov, E. V. Antontseva, L. O. Bryzgalov et al., "rSNP_Guide-based evaluation of SNPs in promoters of the human APC and MLH1 genes associated with colon cancer," *Russian Journal of Genetics: Applied Research*, vol. 4, no. 4, pp. 245–253, 2014.

[148] J. Ponomarenko, T. Merkulova, G. Orlova, O. Fokin, E. Gorshkov, and M. Ponomarenko, "Mining DNA sequences to predict sites which mutations cause genetic diseases," *Knowledge-Based Systems*, vol. 15, no. 4, pp. 225–233, 2002.

[149] J. Ponomarenko, G. Orlova, T. Merkulova, G. Vasiliev, and M. Ponomarenko, "Mining genome variation to associate genetic disease with mutation alterations and ortho/paralogous polimorphysms in transcription factor binding site," *International Journal on Artificial Intelligence Tools*, vol. 14, no. 4, pp. 599–619, 2005.

[150] Y. Ni, A. W. Hall, A. Battenhouse, and V. R. Iyer, "Simultaneous SNP identification and assessment of allele-specific bias from ChIP-seq data," *BMC Genetics*, vol. 13, article 46, 2012.

[151] S. Leschner, I. V. Deyneko, S. Lienenklaus et al., "Identification of tumor-specific *Salmonella* Typhimurium promoters and their regulatory logic," *Nucleic Acids Research*, vol. 40, no. 7, pp. 2984–2994, 2012.

[152] J. Hu, J. W. Locasale, J. H. Bielas et al., "Heterogeneity of tumor-induced gene expression changes in the human metabolic network," *Nature Biotechnology*, vol. 31, no. 6, pp. 522–529, 2013.

[153] M. Hein and S. Graver, "Tumor cell response to bevacizumab single agent therapy in vitro," *Cancer Cell International*, vol. 13, no. 1, article 94, 2013.

[154] M. P. Ponomarenko, J. V. Ponomarenko, A. S. Frolov et al., "Oligonucleotide frequency matrices addressed to recognizing functional DNA sites," *Bioinformatics*, vol. 15, no. 7-8, pp. 631–643, 1999.

[155] M. G. Dozmorov, L. R. Cara, C. B. Giles, and J. D. Wren, "GenomeRunner web server: regulatory similarity and differences define the functional impact of SNP sets," *Bioinformatics*, vol. 32, 2016.

[156] C. Liu, B. Ho, C. Chen et al., "ePIANNO: ePIgenomics ANNOtation tool," *PLOS ONE*, vol. 11, no. 2, article e0148321, 2016.

[157] J. Bendl, M. Musil, J. Štourač et al., "PredictSNP2: a unified platform for accurately evaluating SNP effects by exploiting the different characteristics of variants in distinct genomic regions," *PLoS Computational Biology*, vol. 12, no. 5, Article ID e1004962, 2016.

[158] S. S. Yoo, C. Jin, D. K. Jung et al., "Putative functional variants of XRCC1 identified by RegulomeDB were not associated with lung cancer risk in a Korean population," *Cancer Genetics*, vol. 208, no. 1-2, pp. 19–24, 2015.