UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Scienze Chirurgiche, Oncologiche e Gastroenterologiche

Ph.D. Course: "CLINICAL AND EXPERIMENTAL ONCOLOGY AND IMMUNOLOGY (XXXIII CYCLE)".

Curriculum "Prognostic risk stratification for patients with rare tumours: use of predictive-prognostic models in neoplasms such as melanoma and sarcomas".

"GERMLINE POLYMORPHISMS OF CANDIDATE GENES AS PREDICTORS OF RISK AND PROGNOSIS IN PATIENTS WITH SOFT TISSUE SARCOMAS"

Coordinator: Prof. Paola Zanovello

Supervisor: Prof. Carlo Riccardo Rossi

Co-Supervisor: Prof. Simone Mocellin

Ph.D. Student: Saveria Tropea

SUMMARY

Abstract	**pg. 3**
Background	**pg. 5**
Materials and Methods	**pg. 15**
Results	**pg. 22**
Discussion	**pg. 33**
Conclusions	**pg. 40**
References	**pg. 41**

ABSTRACT

Background: Genetic variants in the clock genes and in telomere-related genes are a potential risk factor for cancer development and progression. As for virtually all cancers, genetic variation of the host is believed to play an important role in the determinism of the risk of soft tissue sarcomas (STS). However, not much information is currently available on the relationship between circadian genes and telomerase related genes polymorphisms and susceptibility or prognosis of patients with STS. The present study aimed to explore this association.

Materials and methods: We analyzed a set of about 19 single nucleotide polymorphisms (SNPs) in the following candidate genes: A) Clock related genes: CLOCK (clock circadian regulator), PER1 (period circadian clock 1), PER2 (period circadian clock 2), NPAS2 (neuronal PAS domain protein 2), TIMELESS (timeless circadian clock), RORA (RAR related orphan receptor A) B) Telomere related genes: TERT (Telomerase Reverse Transcriptase). SNPs genotyping was performed by quantitative real time PCR. We collected peripheral blood and clinic-pathological data from 162 patients with liposarcoma and leiomyosarcoma and from 610 cancer-free individuals. The relationship between the selected SNPs and sarcoma susceptibility and prognosis was tested evaluating additive, recessive and dominant genetic models. Subgroup analysis based on sarcoma histotype was performed. Multivariate logistic regression and multivariate Cox proportional hazard regression analyses were used to evaluate the association between the selected SNPs with the risk of developing STS and prognosis of patients affected by STS. Pathway variation analysis was carried out using the Adaptive Rank Truncated Product model.

Results: We found that 4 SNPs are associated with a lower risk of developing soft tissue sarcomas and 2 SNPs with a higher risk of developing STS. In details: carriers of the minor allele of the *CLOCK rs1801260* had a reduced predisposition to sarcoma under an additive (OR 0.74; 95% CI 0.55-1; P=0.05) and a recessive (OR 0.40; 95% CI 0.18-0.88; P=0.02) genetic model and to liposarcoma at subgroup analysis; *PER2 rs934945* was

associated with a reduced predisposition to sarcoma under an additive (OR 0.65; 95% CI 0.45-0.94; P=0.02) and a dominant (OR 0.63; 95% CI 0.42-0.95; P=0.03) genetic model and to liposarcoma at subgroup analysis; *PER1 rs3027178* was associated with a reduced predisposition only for liposarcoma group (OR 0.68; 95% CI 0.47-0.98; P=0.04); *RORA rs339972* was associated with a decreased predisposition to sarcoma assuming an additive (OR 0.71; 95% CI 0.53-0.96; P=0.02) or a dominant (OR 0.64; 95% CI 0.44-0.93, P=0.02) genetic model and to leiomyosarcoma at subgroup analysis; *NPAS rs895520* was associated with an increased predisposition to sarcoma under an additive (OR 1.33; 95% CI 1.02-1.73; P=0.03) and a recessive (OR 1.70; 95% CI 1.08-2.68; P=0.02) and to leiomyosarcoma; *PER2 rs2304674* was associated with an increased predisposition to sarcoma under a recessive model (OR 2.51; 95% CI 1.26-5.00; P=0.009). About prognosis of patients, we found that: the minor allele of *rs 7602358 located upstream PER2* was associated with survival of patients with liposarcoma, assuming an additive model, (HR 1.98, 95% CI 1.02-3.58; P=0.04); NPAS2 rs2305160 appeared to be associated with liposarcoma survival assuming a recessive model (HR 2.78; 95% CI 1.12-6.91; P=0.02); *TERT rs 2736100* was associated with survival, assuming an additive (HR 1.46, 95% CI 1.03-2.07; P=0.03) and recessive model (HR 2.31; 95% CI 1.35-3.94, P=0.00) and with leiomyosarcoma survival under a recessive model. Germline genetic variation in the whole circadian pathway was associated with the risk of developing STS (P=0.035).

Conclusions: Genetic variation of circadian and telomere-related genes appears to play a role in the determinism of patient susceptibility and prognosis. These findings represent a starting point for further studies.

Keywords: Soft tissue sarcomas (STS), leiomyosarcoma, liposarcoma, Clock genes, telomere-related genes, Single-nucleotide polymorphisms (SNPs), Circadian pathway, risk, survival.

BACKGROUND

Soft tissue sarcomas (STS) are rare tumors of mesenchymal cell origin, originating from muscles, tendons, synovial membranes, fat and connective tissues [1-3]. Although STS are rare cancer, they include more than 50 different histological subtypes with a prevalence of fuso-cellular forms in adult age. STS account for about 2% of all incident malignancies with an estimated incidence averaging 3-5/100,000/year in Europe, and 8,700 new cases/year are estimated in the USA. The incidence rate of STS changes in relation to age with a first peak in pediatric age, followed by a plateau from the age of 20 until 60 years, after which it reaches its highest peak [1-4].

The anatomic site of STS represents an important variable that influences treatment and outcome: 60% of cases develop in the limbs (fig. 1, fig. 2), 10% in the trunk and 15% in the retroperitoneum [1-3].

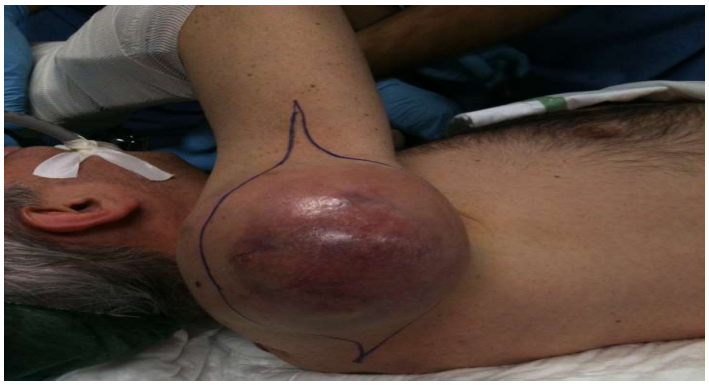

Fig.1: Sarcoma of the shoulder, Veneto Institute of Oncology

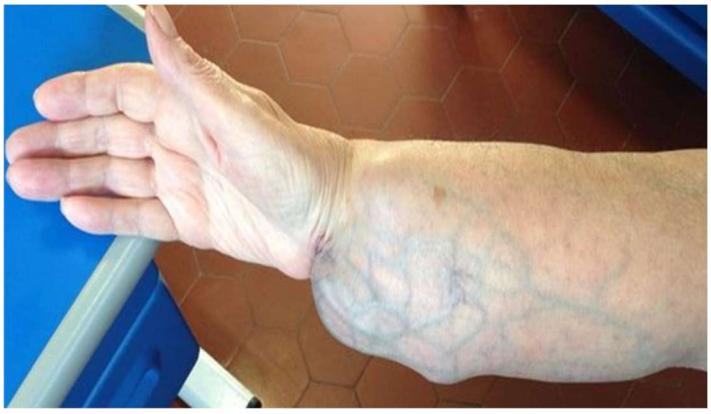

Fig. 2: Sarcoma of the forearm, Veneto Institute of Oncology

These tumors generally occur as considerable size solid masses and they may cause pain, soreness or other dysfunctions only when they increase in size, pressing against nearby nerves, muscle and viscera. The preferred and elective method of investigation, if a STS is suspected, is contrast-enhanced magnetic resonance imaging (MRI) (Fig. 3).

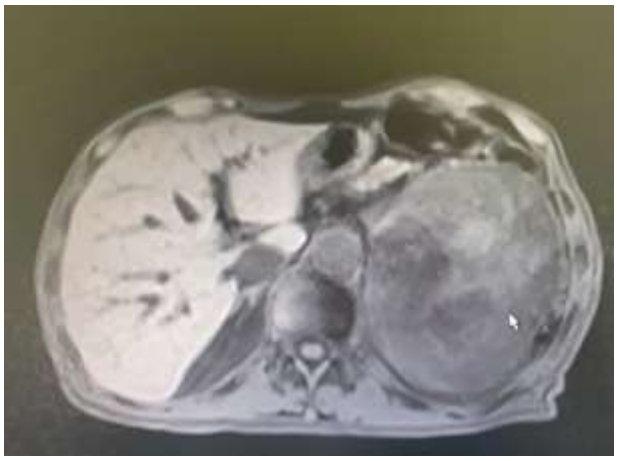

Fig. 3: MRI Image of left- sided retroperitoneal sarcoma.

After MRI, the standard approach to diagnosis of a suspicious mass consists of multiple core needle biopsies (FNAB), preferably image-guided and using >14-16 G needles. An incisional biopsy may be useful if FNAB results inadequate to make a histological diagnosis. Excisional biopsy may be performed, in selected cases, for <5 cm superficial lesion, but it is not recommended. The biopsy should be planned in such a way that the

biopsy pathway and the scar can be removed by definitive surgery to avoid the risk of tumor spread. Histological diagnosis should be performed according to the World Health Organization (WHO) classification and malignancy grade should be provided in all cases. In Europe, this is done according to the Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC) criteria. It distinguishes three malignancy grades based on differentiation, necrosis and mitotic rate [1-3,5-6] .

Some sarcomas are not considered gradable (as alveolar soft tissue sarcoma, epithelioid sarcoma, clear cell sarcoma), while for other histotypes, grading has no prognostic value (as malignant tumor of the peripheral nerve sheath). For cases characterized by a chromosomal translocation or for doubtful diagnoses or when the clinical pathological presentation is unusual, pathological diagnosis should be complemented with molecular techniques (fluorescent in situ hybridisation (FISH) or reverse transcription polymerase chain reaction -RT-PCR-) [1-3].

After histological diagnosis of STS, it is necessary to proceed with additional exams in order to stage the disease. Patients with STS should be staged with high resolution chest computed tomography (CT) to identify possible lung metastases and CT abdomen-pelvis, as well as an ultrasonography of regional lymph nodes may also be indicated. Positron emission tomography- CT (PET- CT) may be useful to clarify any doubts. The most widely-accepted STS staging classification is the American Joint Committee on Cancer (AJCC)/International Union against Cancer (UICC) stage classification system [1-3,5].

Approximately, half of all STS patients with intermediate and/or high-grade tumors develop metastatic disease, which requires systemic treatment with a 5-year overall survival (OS) of not more than 55% [3].

A diverse range of treatments are carried out for the many types of sarcoma. Surgery is the mainstay of the treatment of STS, especially in primary cases. All non-metastastic adult type primary sarcomas are removed when possibile as part of front-line treatment. Surgery alone can cure more than half of adult-type sarcoma patients. Chemotherapy and radiotherapy may also be used before surgery (to devitalize tumours, to decrease tumor size and to prevent metastatic spread) and after (to reduce the risk of recurrence) depending on histology, grading and so the risk of relapse. For patients with

oligometastatic disease and local recurrence, surgery can also be an important part of treatment [1-3,5,6].

In conclusion, to date, surgery is the only potentially curative treatment in early cases, whereas metastatic disease is virtually incurable [1-3]. Therefore, early diagnosis (possibly after identification of high risk individuals) and pre/post-operative therapy (based on fine prognostic stratification) are of pivotal importance to increase the survival rates.

The cause of most STS is unknown. Risk factors for STS include age (about one third are diagnosed in people aged 65 and older, with a lower survival for this group), exposure to ionizing radiation or chemical agents and previous cancer but it is difficult to demonstrate a clear causal relationship [1-3,5].

Recent studies have demonstrated genetic associations with STS, such as the 10% lifetime risk of malignant peripheral nerve sheath tumor (MPNST) in individuals with familial neurofibromatosis by mutation in the NF1 gene or an increased risk of sarcomas in families with polyposis and Li-Fraumeni syndrome (related to mutation in the p53 tumor suppressor gene) or a genetic alteration of regulatory gene RB-1 (familial Retinoblastoma) [3,5,7].

Moreover, some STS patients have an excess of pathogenetic (and potentially aetiological) germline variants [5]. In details, Ballinger et al. investigate the genetic basis for bone and soft-tissue sarcoma seen in routine clinical practice. In this genetic study, they included 1162 patients with sarcoma who were older than 15 years without regard to family history. They found that about half of patients with sarcoma have putatively pathogenic monogenic and polygenic variation in known and novel cancer genes, with possible implications for risk management and treatment. Specifically, all pathogenic variants were associated with earlier age of cancer onset. In addition to tumor suppressor p53 (TP53) , Ataxia-Telangiectasia mutated (ATM), Ataxia-Telangiectasia and Rad3-related protein (ATR), and Breast Related Cancer Antigen-2 (BRCA2), an unexpected excess of functionally pathogenic variants was seen in Excision Repair Cross-Complementation group 2 (ERCC2). So, they demonstrated that probands were more likely than controls to have multiple pathogenic variants compared with the combined control cohort group [8].

As for virtually all cancers, genetic variation of the host is believed to play an important role in the determinism of the risk of STS [9,10-13]. However, much work is still needed before we can have the overall picture defining the molecular mechanisms underlying the predisposition to these tumors. The role of genetic variation in the aggressiveness of cancer has been only recently proposed, some results being already reported for STS [12,13].

The human genome encompasses approximately 50,000 genes and some millions of genetic variants, most of them being single nucleotide polymorphisms (SNPs) [9-10]. Therefore, the study of the role of genetic variation in cancer predisposition is highly challenging, especially considering the potential interactions across polymorphisms and between genetics and environmental factors [9-13].

Germline DNA variation has been long recognized as a key component of the individual risk to develop cancer and recently the discovery rate of susceptibility loci is being greatly accelerated by genome-wide association studies (GWAS) which can test up to one million SNPs in thousands of subjects at a time [12]. However, the proportion of genetic susceptibility to complex traits (such as cancer) explained by single locus analysis still remains small, whereas it is increasingly recognized that multiple locus analysis - such as gene-set (or pathway) analysis - is more powerful for dissecting the genetic architecture of complex diseases according to the principles of systems genetics. In fact, a single SNP can have an effect too small to be detected by the single locus approach, whereas gene/pathway analysis, which jointly tests multiple SNPs from the same gene/pathway, can more likely identify the association between the outcome and the basic functional unit involved in disease development [9,12].

Two main strategies are available to investigate the role of gene variants in disease mechanisms: hypothesis generating high-throughput studies (such as so called GWAS) and the so called candidate gene approach. We chose the latter and identified the two sets of candidate genes: clock genes and telomere related genes and we hypothesized that genetic variations in these genes could play an important role in the determinism of the risk and prognosis of STS.

Clock genes

Circadian rhythms are biological processes, present in almost all living organisms, arising from an ancient adaptation to the rotation of the earth, with a periodicity of about 24 hours. Many physiological process and behaviors in mammals are rhythmic. Sleep-wake cycles, cycling of body temperature, hormone secretion, heart rate, blood pressure, excretion and many other physiological parameters are all circadian functions. These circadian events are controlled by an endogenous molecular clock which is a self-sustained time-tracking rhythmic biological system that synchronizes with both environmental signals and changes (such as light, temperature, food availability) and with social cues and physiological functions (such as physical activity, basal metabolism, body temperature, blood pressure). Organisms modify their behavior and physiological functions, also anticipating environmental changes, thanks to this system. Since the eighties it has been hypothesized that disruption of circadian rhythms has been linked to the risk of different diseases such as diabetes, depression, sleep disorders, obesity, heart attack and cancer. Sleep deprivation, jet-lag, shiftwork involving nightshifts and unnatural light exposure are all potential causes of circadian disruption. In 2007, the International Agency for Research on Cancer (IARC) concluded that "shiftwork that involves circadian disruption is probably carcinogenic to humans" [9-11]. The circadian clock consists of two components: the central clock, located in the suprachiasmatic nucleus of the brain, and the peripheral clocks, which are present in virtually all body tissues. The two components communicate and synchronize with each other and, in particular, the central clock controls the peripheral clocks. Moreover, the central clock modulates the expression of the so-called clock-controlled genes (CCGs), many of them regulating cancer-related biological pathways such as cell proliferation, apoptosis, DNA damage and repair, carcinogen metabolism and/or detoxification [9-12].

The cogwheels of the circadian clock are proteins, whose production and degradation are controlled by interlocked feed-back loops [10]. At least 20% of all mammalian genes have been estimated to be clock-controlled, an indication of extensive circadian gene regulation [9,14] .

So, the biological clock is a mechanism able to maintain and synchronize circadian rhythms via transcription-translation feedback loop, constituted by core circadian genes.

They interact with each other in an intricate manner generating oscillations of gene expression. [9-12]. They can be divided in: positive activators as CLOCK (clock circadian regulator), NPAS2 (neuronal PAS domain protein 2), ARNTL (aryl hydrocarbon receptor nuclear translocator-like, also referred to as brain and muscle Arnt-like protein-1, BMAL1), RORA (RAR related orphan receptor A), and NR1D1 (nuclear receptor subfamily 1 group D member 1 also known as Rev-Erb alpha); negative effectors as CRY1 (cryptochrome circadian clock 1), CRY2 (cryptochrome circadian clock 2), PER1 (period circadian clock 1), PER2 (period circadian clock 2), PER3 (period circadian clock 3) and NR1D2 (nuclear receptor subfamily 1 group D member 2 also known as Rev-Erb beta) and modulators as CSNK1E (casein kinase I epsilon). An additional clock related gene which probably has the role of modulator is TIMELESS (timeless circadian clock) [12].

The principle of circadian clock is successive gene activation in the form of a cycle: the initial activation of a gene is regulated by the last one in the sequence, making up an auto-regulatory feedback loop for which one cycle takes about 24 hours. The functional interactions of their protein products are illustrated in the figure below (Fig. 4).

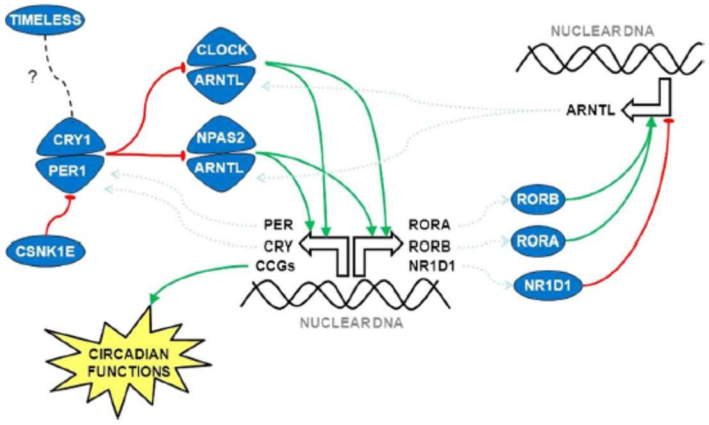

Fig. 4: Schematic view of the circadian pathway, from Mocellin et al. [9]

CLOCK and NPAS2 form heterodimers with ARNTL. These heterodimers function as transcription factors. They enter the nucleus where they bind to E-box sequences of target genes. Besides the CCGs, CLOCK and NPAS2 activate the transcription of other core circadian genes such as PER1, PER2, PER3 and CRY1, CRY2. PER and CRY proteins heterodimerize and activate a negative feedback loop acting directly on CLOCK and NPAS2. The activity of PER and CRY proteins is also regulated by additional proteins such as CSNK1E and CSNK1D (inhibition) and TIMELESS (unclear effect), respectively. CLOCK and NPAS2 also transactivate the expression of other pathway components such as NR1D1, NR1D2 and RORA, RORB and RORC. These proteins are able to inhibit or enhance ARNTL transcription, respectively, with a consequent further level of modulation of CLOCK/NPAS2 activity.

Many scientific evidences support the potential tumour suppressor role of the biological clock. In particular, single germline variations of circadian genes have been associated with the predisposition of some tumour types such as breast carcinoma, lung cancer, prostate carcinoma and non-Hodgkin lymphoma, although the evidence is not conclusive due to the scarcity of data in this recent field of research [9-12, 15-20].

An innovative approach was undertaken for the first time, in 2005, by Zhu et al. They analysed a structural variant in the circadian gene PER3, which was detected to be linked to an increased risk of develop breast cancer in young women [21]. Afterwards, various molecular epidemiological studies found that germline variations in clock genes were associated with different type of cancer susceptibility and, in some cases, with prognosis of cancer patients [9-12, 15-20].

Telomere related genes

Recent studies support also a key role of telomere lenght (TL) in tumorigenesis [22]. Human telomeres consist of repetitive *TTAGGG* DNA sequences that associate with a series of telomere binding proteins believed to provide genomic stability by protecting the linear chromosome ends from being recognized as DNA breaks to be repaired [23]. The inability of the DNA replication machinery to copy the extreme ends of chromosomes is consistent with the observation that cells can lose telomeric repeats without initially affecting cell function. However, a progressive telomere shortening occurs, until a subset of telomeres reach a critically shortened length and induce a DNA damage signal triggering a tumor protein p53 (TP53)–dependent G1/S cell cycle arrest [23,25]. With the inactivity of the gene encoding the enzyme telomerase reverse transcriptase (*TERT*), which synthesizes, in association with other proteins of the core telomerase complex, the *TTAGGG* DNA sequences, telomerase remains silent in the vast majority of human tissues. It is only expressed in a small number of normal cell types such as dividing male germline spermatocytes and a subset of proliferating somatic adult stem cells [23-25].

In the early 1990s, investigators proposed a connection between telomeres, telomerase, aging and cancer [23,25,26]. The hypothesis was that most normal human cells lack telomerase activity and their telomeres shorten with each cell division, until they enter

replicative senescence. Cells that lose critical cell-cycle checkpoint functions escape replicative senescence and continue to divide and eventually enter a second growth arrest state (crisis) when many shortened chromosome ends fuse [23,25]. In human cells, replicative senescence and crisis represent two mechanisms to limit cell growth and they are potent anticancer protection mechanisms. Most human cells remain in this crisis period with cell growth being balanced by cell death until a rare cell acquires a mechanism, such as telomerase expression, that can maintain or lengthen telomeres.

Through this mechanism, cells become able to grow continuously and so immortal. This is generally believed to be a critical step in cancer progression [23-26].

Considering the evidence that telomerase activity is linked to the development of many tumor types, many investigators are designing a variety of methods to target telomerase as a novel therapeutic approach potentially useful in a range of cancers [27,28].

Moreover, other researchers tested the hypothesis that variability of the *TERT* gene sequence might be a general mechanism affecting individual cancer predisposition. In this respect, recent findings support the hypothesis that genetic variability in this genomic region can modulate cancer susceptibility in humans. [23]

For example, regarding cutaneous melanoma, the role of TL in the tumorigenesis is complex and either short or long telomeres may increase the individual risk of melanoma, depending on other predisposing/risk factors [29]. It has been demonstrated that germline polymorphisms underlying long telomeres increase cancer (including melanoma) risk by studies in which predisposing factors was not always evaluated. On the other hand, ultraviolet-light exposure, a known environmental risk factor for melanoma, was been demonstrated to induce telomere shortening [22,23,29].

Regarding soft tissue sarcoma, there are few studies demonstrating a positive association between TL and cancer risk. However, in mesenchymal cells, the telomere elongation is maintained by a telomerase-independent mechanism with not completely clarified molecular basis [23,30,31].

In conclusion, pioneering work in the field of molecular cancer epidemiology suggested that genetic variants in the clock genes and telomere-related genes are potential risk factors for developing cancer. However, no much information is currently available on

the association between variation of the clock and telomere-related genes and risk and prognosis of STS.

Our goal is to test two working hypotheses:

- is the genetic variation of telomere-related genes or clock genes associated with the risk of developing soft tissue sarcomas?

- is the genetic variation of telomere-related genes or clock genes associated with the prognosis of patients with soft tissue sarcomas?

By combining the findings of this study along with the results of similar studies, we accomplished:

- the definition of a panel of SNPs for the identification of individuals at high risk of soft tissue sarcomas;

- the definition of a panel of SNPs, for the stratification of the risk of disease recurrence, which would help us to identify patients who can most benefit from adjuvant or neoadjuvant therapy.

MATERIALS AND METHODS

Study design

This is a retrospective study where anthropometric, clinic-pathological, and genetic data were collected from 162 patients with soft tissue sarcomas treated at our institution (University Hospital of Padova- Veneto Institute of Oncology) between 1992 and 2019, using a prospectively maintained database linked to our institutional biobank (Clinica Chirurgica I–Veneto Insitute of Oncology).

Of these, 93 patients had a histological diagnosis of liposarcoma and 69 patients of leiomyosarcoma.

To be included in the study, each case had to meet the following requirements: (1) histologically confirmed diagnosis of liposarcoma or leiomyosarcoma; (2) pathology-based information on TNM stage; (3) follow up data (minimum follow up: 6 months); (4) availability of peripheral blood for genotyping purposes.

In addition, anthropometric and genetic data were collected from 610 cancer free control individuals.

Healthy controls selection was both population-based (n = 270 blood donors) and hospital-based (n = 340, healthy subjects who visited the Clinica Chirurgica I ambulatories for routine check-ups). All patients and all controls signed an informed consent form explaining the research purposes of the blood withdrawal.

The study consists of two parts: 1) a case-control study to investigate the role of germ-line polymorphisms in candidate genes in the determinism of predisposition to STS; 2) a prognostic biomarkers study to identify germ-line polymorphisms of candidate genes associated with the survival of patients with STS.

SNPs selection

For both studies, we analyzed a set of 19 SNPs in the following candidate genes:

- Core Clock genes: CLOCK (clock circadian regulator), PER1 (period circadian clock 1), PER2 (period circadian clock 2), NPAS2 (neuronal PAS domain

protein 2), RORA (RAR related orphan receptor A). The SNPs main features of core clock genes tested in this study are summarized in table 1.

Gene	SNP ID	Genotype	Ctrls	Cases	P-val.	Chr	Region	Residue
CLOCK	rs1801260	TT	323	92	0.21	4	3'UTR	
		TC	228	62				
		CC	56	8				
	rs3736544	GG	241	60	0.72	4	Exon	Asn>Asn
		GA	269	78				
		AA	95	24				
	rs3749474	CC	259	62	0.45	4	3'UTR	
		CT	266	80				
		TT	83	20				
	rs34897046	GG	566	149	0.60	4	Exon	Ser>Cys
		GC	40	13				
		CC	1	0				
NPAS2	rs895520	GG	211	49	0.20	2	Intron	
		GA	294	76				
		AA	103	37				
	rs2305160	GG	283	75	1.00	2	Exon	Thr>Ala
		GA	264	71				
		AA	59	16				
PER1	rs3027178	TT	281	84	0.29	17	Exon	Thr>Thr
		TG	253	64				
		GG	76	14				
PER2	rs2304674	AA	330	87	0,08	2	Intron	
		AG	246	58				
		GG	29	15				
	rs934945	CC	386	118	0.07	2	Exon	Gly>Glu
		CT	206	42				
		TT	17	2				
	rs7602358	TT	358	87	0.49	2	Intron	
		TG	213	64				
		GG	38	11				
RORA	rs339972	TT	312	97	0.15	15	Intron	
		TC	233	53				
		CC	62	12				
	rs10519097	CC	422	110	0.67	15	Intron	
		CT	173	50				
		TT	14	2				

Table 1. SNPs main features of core clock genes tested in this study

- Clock-related gene: TIMELESS (timeless circadian clock). The SNPs main features of the gene tested in this study are summarized in table 2.

Gene	SNP ID	Genotype	Ctrls	Cases	P-val.	Chr	Region	Residue
TIMELESS	rs774027	AA	157	44	0.63	12	Exon	Ile>Val
		AT	301	74				
		TT	148	44				
	rs3809125	CC	250	70	0.87	12	Upstream	
		CT	280	71				
		TT	74	20				
	rs7302060	TT	181	55	0.53	12	Intron	
		TC	304	74				
		CC	121	33				

Table 2. SNPs main features of TIMELESS tested in this study.

- Telomere related genes: TERT (Telomerase Reverse Transcriptase). The SNPs main characteristics of the telomere-related gene tested in this study are summarize in Table 3.

Gene	SNP ID	Genotype	Ctrls	Cases	P-val.	Chr	Region	Residue
TERT	rs2242652	GG	383	105	0,442	5	Intron	
		GA	188	51				
		AA	30	4				
	rs2736098	CC	347	85	0,612	5	Exon	Ala>Ala
		CT	225	62				
		TT	33	11				
	rs2736100	CC	169	47	0,917	5	Intron	
		CA	308	81				
		AA	129	32				
	rs2853676	TT	294	89	0,258	5	Intron	
		TC	250	59				
		CC	62	12				

Table 3: SNPs main features of telomere-related gene tested in this study

The CLOCK locus contains 28734 SNPs annotated by the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/) SNP data repository. The Genome Variation Server of the University of Washington (http://gvs.gs.washi ngton .edu/GVS/) and the TagSNP tool of the US National Institute of Environmental Health Sciences (https ://snpin fo.niehs .nih.gov/ snpin fo/snpta g.html) were interrogated to identify the SNPs that tag CLOCK SNPs with a minimum correlation coefficient (r2) of at least 0.80, minimal genotype data coverage of 50%, and minimal allele frequency of 5%, in the Caucasian population (CEU) genotyped by the HapMap Project. This process yielded an initial group of 134 Tag SNPs for genotyping, divided in 11 bins (groups of SNPs in high linkage disequilibrium, D' greater than 0.8). Excluding those SNPs tagging less than 10 SNPs, we selected from each bin the Tag SNPs with more functional/cancer related information available. They are the following: rs1801260, rs3736544, rs3749474. Moreover, we included rs34897046 which has a missense functional effect.

The TIMELESS locus contains 9200 SNPs annotated by the NCBI dbSNP. With the same procedure described, rs774027 was chosen as TagSNP. In order to enrich our analysis, rs3809125 (3′-UTR) and rs7302060 were added based on literature [32].

For NPAS2, PER1, PER2, RORA we decided to follow a different approach. Despite the elevated number of Tag SNPs for each gene (more than 200 in NPAS2 locus) none of them tags more than 9 SNPs and most of them tags 1 SNP. We relied on literature [10].

NPAS2 rs895520, PER2 rs7602358, RORA rs339972, and rs10519097 had a statistically significant association with cancer risk. Moreover, regarding the NPAS2 locus, we selected the most studied variant, based on number of datasets, with missense functional effect, rs2305160. PER1 rs3027178, PER2 rs934945 and PER 2 rs2304674 were selected based on literature [33-37].

TERT SNPs were chosen on the basis of literature. We selected some of the most studied variants: rs2242652, rs2736098, rs2736100, rs2853676 [38-43].

DNA extraction and genotyping

DNA extraction

Genomic DNA was isolated from peripheral whole blood using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) following the recommendations of the manufacturer and quantified by Nanodrop 1000 spectrophotometer (Thermofisher Scientific, Waltham, MAUSA).

Subjects' genotyping

10 to 20 ng of DNA of each patient were used for TaqMan SNP genotyping assays (comprising primers and fluorescent probes) according to the manufacturer's instructions (Thermofisher Scientific, Waltham, MA USA). Genotyping was performed by real–time PCR using either allelic discrimination in the 7300 RT–PCR System (Thermofisher Scientific, Waltham, MA USA), either using endpoint genotyping in an LightCycler 480II (Roche Molecular Diagnostics Pleasanton, CA, USA). The amplification procedure consisted of an initial denaturing step at 95 °C for 10 minutes followed by 40 cycles at 95 °C for 15 seconds and 60 °C for 1 minute. Post run data were analysed by 7300 SDS software (Thermofisher Scientific, Waltham, MA USA) and Automatic calls were assigned with approximately 99.8% quality. A call rate > 95% was considered the cutoff to consider genotyping. For Light- Cycler 480II PCR parameters involved an initial denaturation at 95 °C for 10 minutes followed by 40 cycles at 95 °C for 10 seconds, 60 °C for 1 minute, and 72 °C for 10 seconds. Post run data were analysed by LightCycler 480 Endpoint Genotyping software (Roche Molecular Diagnostics Pleasanton, CA, USA). Two blank (water) controls in each 96-well plate were used for the assay quality control.

Statistical analysis

We considered three genetic models:

1. Additive (or allelic; presence of 0 vs 1 vs 2 copies of the variant allele): the genotype for each SNP is considered as a continuous variable;

2. Recessive (presence of 2 copies of the variant allele vs 0 or 1 copies of the variant allele): the genotype for each SNP is considered as a categorical variable with 2

categories (homozygous for the common allele + heterozygous, homozygous for the variant allele);

3. Dominant (presence of at least one copy of the variant allele vs no copies of the variant allele): the genotype for each SNP is considered as a categorical variable with 2 categories (homozygous for the variant allele + heterozygous, homozygous for the common allele).

For susceptibility of sarcoma assessment multivariate logistic regression was employed for each genetic model. Association between each SNP and risk of sarcoma was assessed by Odds ratios (OR) with 95% confidence intervals. In those multivariate models the evaluated outcome was the presence or absence of sarcoma, while the explanatory variables were the single SNPs adjusted for age and gender.

For prognosis of sarcoma assessment multivariate Cox proportional hazard regression was employed. Overall survival was defined as the time from the date of tumor diagnosis to the date of death by any cause or last follow up visit. Hazard ratios (HR) and 95% confidence intervals were used as a measure of association. In those multivariate models the evaluated event was the patient's survival, the time to event were the months of survival, and the explanatory variables were the single SNP adjusted for age, gender and sarcoma stage.

For discriminating whether sarcoma subtype interacts with SNP associations with susceptibility or prognosis, the above mentioned statistical analyses were repeated, for the liposarcoma subgroup and for the leiomyosarcoma subgroup, employing the additive, recessive and dominant models. Hardy–Weinberg equilibrium (HWE) was tested for both samples (patients and healthy controls) for each SNP employing OEGE—Online Encyclopedia for Genetic Epidemiology studies, http://www.oege.org/softw are/hwe-mr-calc.shtml . This tool is a HWE calculator for biallelic SNPs based on Chi-square statistic. Statistical power was calculated for each SNP employing the on-line tool "Power and Sample Size" of the University of Vanderbilt (http://biost at.mc.vande rbilt .edu/ wiki/Main/Power Sampl eSize)[44]. Power was defined as the probability of correctly rejecting the null hypothesis that the relative risk (OR) was equal to 1, given 162 case patients and 610 controls. The type I error probability α was set to 0.05 and ψ (OR

considered clinically relevant in our study) was set to 0.80. For all the analysis conducted in this study, a P value ≤ 0.05 was used as the cut-off for significance, not adjusted for multiple comparisons. Rcmdr: R Commander. R package version 2.4-2 was employed for the analyses.

Pathway variation analysis

We also investigated the associations of soft tissue sarcoma with genes, observed as combinations of SNPs, and with the circadian rhythm pathway as a whole, observed as combinations of 6 clock genes. This analysis was conducted employing the Adaptive Rank Truncated Product (ARTP) method. The ARTP method was developed to overcome the major limitations of other existing P-value-combining approaches which do not take into consideration the organization of the DNA into functional elements, ignore the linkage disequilibrium patterns between SNPs within the same gene/gene region, and arbitrarily specify a K rank truncation point so as to combine the K smallest P values as the summary statistics. Instead, ARTP takes into consideration the gene-based structure of biological pathways as well as the correlation between P values and selects the optimal rank truncation point among a set of candidates and then adjusts the generated P value for multiple testing using a permutation procedure [9,45].

RESULTS

The analysis was performed on a total of 162 cases of sarcoma and 610 healthy controls, all of European ancestry. The baseline features of the patients and the controls are provided in table 4.

The main features of the SNPs we investigated are shown in table 1, table 2 and table 3.

A set of 15 preselected SNPs in 6 circadian clock genes and a total of 4 preselected SNPs of in a telomere-related gene were successfully genotyped. The distribution of genotypes is consistent with the Hardy–Weinberg equilibrium for both cases and controls.

The *CLOCK* rs34897046 SNP showed minor allele frequencies < 5% across the whole study population (3% in controls and 4% in cases) and was not evaluated further statistically.

Characteristc	Controls (n=610)	Cases (n=162)	Liposarcoma (n=93, 57.4%)	Leiomyosarcoma (n=69, 42.6%)
Mean age, years (st. dev.)	48.6 (14.8)	60 (14)	59.8 (13.3)	60.2 (14.9)
Gender, n (%)				
Male	336 (55.2)	95 (58.6)	65 (69.9)	30 (43.5)
Female	273 (44.8)	67 (41.4)	28 (30.1)	39 (56.5)
Source of ctrls, n (%)				
Hospital	340 (55.7)			
Population	270 (44.3)			
Patient status, n (%)				
Alive		89 (54.9)	66 (71)	23 (33.3)
Deceased		73 (45.1)	27 (29)	46 (66.7)
Median Survival, months (min, max)		79 (1, 366)	91.5 (2, 366)	42 (1, 196)
Tumoral stage, n (%)				
I		37 (23)	23 (25)	14 (20.3)
II		54 (33.5)	36 (39.1)	18 (26.1)
III		55 (34.2)	29 (31.5)	26 (37.7)
IV		15 (9.3)	4 (4.3)	11 (15.9)

Table 4. Subjects' characteristics of sarcoma cases and healthy controls

Susceptibility assessment

We performed an univariate logistic regression analysis evaluating additive, recessive and dominant genetic models. We used odds ratios (ORs) and their corresponding 95% confidence intervals (95% CI) to evaluate the strength of association between each polymorphism and sarcoma predisposition. We found that 4 SNPs are associated with a lower risk of developing soft tissue sarcomas and 2 SNPs with a higher risk of developing STS**.** The results are summarized in table 5. We examined further subgroup SNP associations according to sarcoma histology, under the additive model: liposarcoma (n = 93) and leiomyosarcoma (n = 69). The results are reported in Table 6.

- *CLOCK* maps on chromosome 4 at 4q12. It encodes a member of the basic helix-loop-helix PAS family of transcription factors, known to play a central role in the control of diverse cellular events, and forms heterodimers with ARNTL (BMAL1) to enhance target gene expression. It appears that transcriptional activation is facilitated by the histone acetyl transferase (HAT) activity of the clock protein. Histone acetylation promotes transcription through the modification of histones and allows opening of the condensed chromatin. This provides access to the transcriptional machinery. CLOCK is also involved in growth arrest, DNA repair and apoptosis upon genotoxic stress caused by UV radiation, suggesting that this molecule may represent an important "caretaker" promoting cell cycle arrest upon DNA damage [9-12, 46]. **rs1801260** is located on 3′-UTR region of *CLOCK*. Carriers of the minor allele (C) rs1801260 had a reduced predisposition to sarcoma under an additive (per allele OR 0.74; 95% CI 0.55–1.00; P = 0.05) and a recessive (OR 0.40; 95% CI 0.18–0.88; P = 0.02) genetic model and to liposarcoma at subgroup analysis under an additive model (per allele OR 0.67; 95% CI 0.46–0.98; P = 0.04). No significant association was found with leiomyosarcoma.

- *NPAS2* is the largest human clock gene. It maps on chromosome 2 at 2q11.2 and, like its paralogue *CLOCK*, encodes for a member of the basic helix-loop-helix PAS class of transcription factors [9-12,47]. When dimerized with ARNTL, NPAS2 binds to E-box regulatory sequences in the promoter regions of target genes and contributes to the activation of their expression. Previous studies reported NPAS2 as a putative tumor suppressor [47]. **rs895520** is located on an intron of *NPAS2* locus and in this study it was associated with an increased predisposition to sarcoma under an additive (OR 1.33; 95% CI 1.02-1.73; P=0.03) and a recessive (OR 1.70; 95% CI 1.08-2.68; P=0.02) and to leiomyosarcoma (per allele OR 1.44; 95% CI 0.99–2.09; P=0.05).

- *PERs* code for PAS domain-containing key regulators of the circadian clock. *PER* genes control their own transcription by directly repressing ARNTL heterodimers, their activators [9-12]. Moreover, it has been suggested that PER1 and PER2 function as tumor suppressors. *PER2* gene can activate c-Myc signaling pathways leading to genomic instability and cell proliferation. PER2 dysfunction can also impair p53-mediated apoptosis and consequently result in genomic instability and

the accumulation of damaged cells. *PER1* maps on chromosome 17 at 17p13.1 and *PER2* on chromosome 2 at 2q37.3 [9-12, 48-54]. **rs934945 (C>T)** is located on the last exon of PER2 locus and has a missense functional effect, leading to the substitution Glycine–Glutamic Acid. Carriers of the minor allele PER2 rs934945 had a reduced predisposition to sarcoma under an additive (OR 0.65; 95% CI 0.45-0.94; P= 0.02) and a dominant (OR 0.63; 95% CI 0.42-0.95; P=0.03) genetic model and to liposarcoma at subgroup analysis (per allele OR 0.59; 95% CI 0.37–0.96; P = 0.03).

rs2304674 is an intronic SNP of PER 2 and was associated with an increased predisposition to sarcoma under a recessive model (OR 2,51; 95% CI 1.26-5.00; P=0.009).

Moreover **PER1 rs3027178**, a genetic variant with a synonymous functional effect, was associated with a reduced predisposition only for liposarcoma group (per allele OR 0.68; 95% CI 0.47–0.98; P = 0.04).

- *RORA* maps on chromosome 15 at 15q21-q22, spans a 730 kb large genomic region comprised of 15 exons. The protein encoded by this gene is a member of the retinoid orphan nuclear receptor subfamily of orphan receptors. It can bind as a monomer or as a homodimer to hormone response elements upstream of several genes to enhance the expression of those genes. It has been shown to aid in the transcriptional regulation of genes involved in circadian rhythm. The encoded protein has been also shown to interact with NM23-2, a nucleoside diphosphate kinase involved in organogenesis and differentiation, as well as with NM23-1, the product of a tumor metastasis suppressor candidate gene. ARNTL-CLOCK or ARNTL-NPAS2 heterodimers promote the transcription of *RORA*, which in turn activates the transcription of ***ARNTL*** [55-56]**. rs339972 C** allele was associated with a decreased predisposition to develop sarcoma assuming an additive (per allele OR 0.71;95% CI 0.53–0.96; P =0.02) or a dominant (OR 0.64; 95% CI 0.44–0.93; P=0.02) genetic model. This association is statistically significant in leiomyosarcoma subgroup (per allele OR 0.64; 95% CI 0.42–0.98; P=0.04).

No significant associations were found for the selected *TIMELESS* SNPs and the predisposition to develop sarcoma.

- The TERT gene is located in 5p15.33. TERT, the telomerase catalytic subunit, maintains telomere stability acting as a reverse transcriptase in the elongation of telomere length. It is reported that genetic variations in TERT gene is involved in many types of cancers [22-24, 57]. In this study we didn't find a statistically significant association between the selected TERT SNPs and the risk of developing soft tissue sarcomas.

Gene	SNP ID	Additive		Recessive		Dominant	
		OR (95%CI)	P-val	OR (95%CI)	P-val	OR (95%CI)	P-val
CLOCK	**rs1801260**	**0.74 (0.55-1.00)**	**0.05**	**0.40 (0.18-0.88)**	**0.02**	0.79 (0.54-1.14)	0.20
	rs3736544	1.15 (0.89-1.50)	0.29	1.10 (0.66-1.83)	0.72	1.27 (0.87-1.86)	0.22
	rs3749474	1.06 (0.81-1.38)	0.68	0.94 (0.54-1.62)	0.82	1.15 (0.79-1.67)	0.47
NPAS2	**rs895520**	**1.33 (1.02-1.73)**	**0.03**	**1.70 (1.08-2.68)**	**0.02**	1.30 (0.87-1.93)	0.20
	rs2305160	0.95 (0.72-1.26)	0.71	0.97 (0.52-1.79)	0.92	0.92 (0.64-1.34)	0.68
PER1	rs3027178	0.80 (0.61-1.06)	0.12	0.72 (0.39-1.35)	0.31	0.76 (0.53-1.10)	0.15
PER2	**rs2304674**	1.20 (0.89-1.62)	0.23	**2.51 (1.26-5.00)**	**0.00**	1.04 (0.72-1.50)	0.83
	rs934945	**0.65 (0.45-0.94)**	**0.02**	0.48 (0.10-2.25)	0.35	**0.63 (0.42-0.95)**	**0.03**
	rs7602358	1.21 (0.90-1.62)	0.21	1.18 (0.56-2.48)	0.66	1.29 (0.89-1.86)	0.18
RORA	**rs339972**	**0.71 (0.53-0.96)**	**0.02**	0.69 (0.35-1.36)	0.28	**0.64 (0.44-0.93)**	**0.02**
	rs10519097	0.98 (0.68-1.40)	0.89	0.55 (0.12-2.53)	0.44	1.02 (0.69-1.51)	0.93
TIMELESS	rs774027	0.98 (0.76-1.26)	0.85	1.02 (0.67-1.55)	0.92	0.92 (0.61-1.39	0.69
	rs3809125	0.92 (0.70-1.20)	0.52	0.89 (0.51-1.55)	0.67	0.89 (0.62-1.30)	0.55
	rs7302060	0.98 (0.76-1.27)	0.89	1.03 (0.65-1.62)	0.91	0.94 (0.63-1.39)	0.75
TERT	rs2242652	0.89 (0.64-1.25)	0.50	0.50 (0.17-1.52)	0.22	0.95 (0.64-1.40)	0.78
	rs2736098	1.11 (0.82-1.50)	0.49	1.11 (0.52-2.37)	0.78	1.14 (0.79-1.66)	0.79
	rs2736100	0.90 (0.69-1.17)	0.43	0.78 (0.49-1.24)	0.28	0.95 (0.63-1.42)	0.28
	rs2853676	0.80 (0.60-1.07)	0.13	0.74 (0.37-1.46)	0.38	0.76 (0.52-1.10)	0.14

Table 5. Associations of circadian pathway genes and TERT gene with predisposition to sarcoma under 3 models of inheritance: additive, recessive, dominant.

Gene	SNP ID	Liposarcoma		Leiomyosarcoma	
		OR (95%CI)	**P-val**	**OR (95%CI)**	P-val
CLOCK	**rs1801260**	**0.67 (0.46-0.98)**	**0.04**	0.84 (0.56-1.25)	0.39
	rs3736544	1.23 (0.88-1.71)	0.22	1.09 (0.76-1.57)	0.64
	rs3749474	1.09 (0.78-1.53)	0.60	1.00 (0.69-1.46)	0.98
NPAS2	**rs895520**	1.26 (0.90-1.57)	0.18	**1.44 (0.99-2.09)**	**0.05**
	rs2305160	1.01 (0.72-1.43)	0.95	0.88 (0.59-1.32)	0.54
PER1	**rs3027178**	**0.68 (0.47-0.98)**	**0.04**	0.96 (0.66-1.41)	0.85
PER2	rs2304674	1.21 (0.83-1.78)	0.31	1.20 (0.78-1.85)	0.39
	rs934945	**0.59 (0.37-0.96)**	**0.03**	0.71 (0.42-1.20)	0.20
	rs7602358	1.23 (0.85-1.77)	0.28	1.16 (0.77-1.75)	0.49
RORA	**rs339972**	0.75 (0.52-1.07)	0.12	**0.64 (0.42-0.98)**	**0.04**
	rs10519097	0.82 (0.51-1.31)	0.40	1.28 (0.80-2.06)	0.31
TIMELESS	rs774027	0.98 (0.71-1.35)	0.90	0.95 (0.66-1.35)	0.76
	rs3809125	0.99 (0.71-1.38)	0.95	0.79 (0.53-1.16)	0.23
	rs7302060	1.00 (0.72-1.37)	0.97	1.01 (0.70-1.45)	0.97
TERT	rs2242652	1.04 (0.70-1.55)	0.83	0.72 (0.44-1.18)	0.19
	rs2736098	1.09 (0.75-1.60)	0.63	1.17 (0.77-1.78)	0.46
	rs2736100	0.86 (0.62-1.19)	0.35	0.90 (0.63-1.29)	0.56
	rs2853676	0.76 (0.53-1.11)	0.15	0.84 (0.56-1.25)	0.38

Table 6. Association of circadian pathway genes and TERT gene with predisposition to liposarcoma and leiomyosarcoma in subgroups analysis under the additive genetic model.

Prognosis assessment

rs7602358 located upstream PER2 was significantly associated with liposarcoma survival (HR 1.98; 95% CI 1.02– 3.85; P=0.04) employing an additive genetic model with an increased risk of mortality of 98**%.**

rs2305160 is an exonic SNP of **NPAS2,** leading to the substitution threonine–alanine. Assuming a recessive model it appears to be associated with liposarcoma survival (HR 2.78; 95% CI 1.12-6.91; P=0.02)

The rs2736100 polymorphism in the second intron of the TERT gene is associated with survival in STS patients, assuming an additive model (HR 1.46, 95% CI 1.03-2.07; P =0.03) and a recessive genetic model (HR 2.31, 95% CI 1.35-3.94; P=0.00).

Moreover at subgroup analysis, under a recessive model, it was associated with an increased risk only for leiomyosarcoma group (OR 2.43; 95% CI 1.24–4.75; P=0.009).

No further statistically significant associations with prognosis were found, neither in primary analysis nor in subgroup analysis.

The results of the associations of the CLOCK genes and TERT SNPs with prognosis of sarcoma under additive, recessive and dominant models of inheritance are reported in table 7. The results of subgroup analysis under additive, recessive and dominant genetic models are summarized in table 8, 9 and 10 respectively.

		Additive		***Recessive***		***Dominant***	
Gene	**SNP ID**	**HR (95% CI)**	**P-val.**	**HR (95% CI)**	**P-val.**	**HR (95% CI)**	**P-val.**
CLOCK	rs1801260	1.06 (0.71-1.60)	0.77	1.39 (0.50-3.92)	0.53	1.02 (0.62-1.66)	0.94
	rs3736544	0.95 (0.66-1.37)	0.76	1.41 (0.72-2.73)	0.31	0.75 (0.45-1.23)	0.25
	rs3749474	1.01 (0.69-1.48)	0.97	1.77 (0.88-3.58)	0.11	0.81 (0.50-1.31)	0.40
NPAS2	rs895520	1.01 (0.73-1.39)	0.95	0.87 (0.48-1.55)	0.63	1.15 (0.68-1.94)	0.59
	rs2305160	0.90 (0.63-1.29)	0.56	0.98 (0.48-2.01)	0.97	0.83 (0.51-1.34)	0.44
PER1	rs3027178	0.98 (0.66-1.44)	0.91	2.09 (0.94-4.63)	0.07	0.79 (0.49-1.26)	0.33
PER2	rs934945	0.70 (0.41-1.20)	0.19			0.71 (0.41-1.22)	0.21
	rs2304674	0.80 (0.55-1.18)	0.26	0.81 (0.34-1.91)	0.68	0.73 (0.45-1.20)	0.22
	rs7602358	1.20 (0.82-1.75)	0.34	1.58 (0.70-3.57)	0.27	1.16 (0.72-1.87)	0.55
RORA	rs339972	1.03 (0.70-1.51)	0.88	1.66 (0.74-3.72)	0.22	0.91 (0.56-1.47)	0.69
	rs10519097	0.72 (0.43-1.20)	0.21	1.49 (0.20-11.25)	0.70	0.68 (0.40-1.16)	0.16
TIMELESS	rs774027	0.86 (0.63-1.18)	0.36	0.70 (0.40-1.25)	0.23	0.91 (0.55-1.53)	0.73
	rs3809125	0.88 (0.63-1.23)	0.45	0.57 (0.27-1.21)	0.14	1.00 (0.62-1.63)	0.99
	rs7302060	1.10 (0.81-1.51)	0.54	1.08 (0.62-1.87)	0.78	1.21 (0.73-2.01)	0.46
TERT	rs2242652			0.75 (0.18-3.13)	0.70	0.70 (0.49-1.17)	0.17
	rs2736098	1.21 (0.80-1.83)	0.36			1.15 (0.68-1.91)	0.59
	rs2736100	**1.46 (1.03-2.07)**	**0.03**	**2.31 (1.35-3.94)**	**0.00**	1.18 (0.70-2.00)	0.52
	rs2853676	0.74 (0.50-1.09)	0.13	0.48 (0.18-1.24)	0.13	0.77 (0.46-1.27)	0.31

Table 7. Associations of circadian pathway genes and TERT gene with prognosis of sarcoma under additive, recessive and dominant models of inheritance.

Additive Model		*Liposarcoma*		*Leiomyosarcoma*	
Gene	**SNP ID**	**HR (95% CI)**	**P-val.**	**HR (95% CI)**	**P-val.**
CLOCK	rs1801260	1.04 (0.51-2.10)	0.92	0.81 (0.49-1.35)	0.42
	rs3736544	0.86 (0.47-1.58)	0.64	1.04 (0.65-1.64)	0.88
	rs3749474	1.16 (0.59-2.28)	0.66	1.15 (0.74-1.79)	0.54
NPAS2	rs895520	0.96 (0.55-1.69)	0.89	0.97 (0.65-1.45)	0.87
	rs2305160	0.95 (0.53-1.70)	0.85	0.87 (0.55-1.37)	0.55
PER1	rs3027178	0.99 (0.47-2.05)	0.97	1.00 (0.60-1.67)	0.99
PER2	rs934945	0.50 (0.17-1.45)	0.20	0.66 (0.34-1.28)	0.22
	rs2304674	0.65 (0.34-1.24)	0.19	0.83 (0.48-1.42)	0.50
	rs7602358	**1.98 (1.02-3.85)**	**0.04**	0.95 (0.58-1.54)	0.83
RORA	rs339972	1.01 (0.55-1.85)	0.98	1.14 (0.66-1.99)	0.63
	rs10519097	0.44 (0.15-1.30)	0.14	0.91 (0.50-1.66)	0.75
TIMELESS	rs774027	0.63 (0.35-1.11)	0.11	0.77 (0.51-1.15)	0.20
	rs3809125	0.88 (0.48-1.61)	0.68	0.87 (0.56-1.33)	0.52
	rs7302060	1.65 (0.93-2.90)	0.08	1.20 (0.78-1.82)	0.41
TERT	rs2242652	0.90 (0.42-1.93)	0.79	0.83 (0.47-1.48)	0.54
	rs2736098	1.10 (0.56-2.17)	0.76	1.30 (0.74-2.29)	0.34
	rs2736100	0.93 (0.48-1.80)	0.84	1.51 (0.97-2.33)	0.06
	rs2853676	0.63 (0.30-1.34)	0.23	0.83 (0.52-1.32)	0.44

Table 8. Association of circadian pathway genes and TERT gene with prognosis of liposarcoma and leiomyosarcoma in subgroups analysis under the additive genetic model.

Recessive Model		***Liposarcoma***		***Leiomyosarcoma***	
Gene	**SNP ID**	**HR (95% CI)**	**P-val.**	**HR (95% CI)**	**P-val.**
CLOCK	rs1801260	1.17 (0.15-8.67)	0.87	0.52 (0.16-1.72)	0.29
	rs3736544	0.43 (0.10-1.86)	0.26	1.57 (0.77-3.17)	0.20
	rs3749474	1.37 (0.40-4.63)	0.60	1.22 (0.54-2.75)	0.62
NPAS2	rs895520	0.95 (0.35-2.53)	0.92	0.88 (0.44-1.74)	0.72
	rs2305160	**2.78 (1.12-6.91)**	**0.02**	0.50 (0.15-1.64)	0.25
PER1	rs3027178	0.98 (0.23-4.18)	0.98	1.35 (0.57-3.22)	0.48
PER2	rs934945	0.00 0	0.99	0.00 0	0.99
	rs2304674	0.56 (0.12-2.45)	0.44	1.43 (0.49-4.16)	0.50
	rs7602358	1.75 (0.52-5.84)	0.36	1.04 (0.37-2.92)	0.93
RORA	rs339972	1.37 (0.47-3.96)	0.56	1.66 (0.51-5.4)	0.39
	rs10519097	0.00 0	0.99	1.63 (0.22-11.96)	0.63
TIMELESS	rs774027	0.64 (0.24-1.69)	0.37	0.88 (0.44-1.74)	0.72
	rs3809125	0.74 (0.22-2.48)	0.63	0.95 (0.37-2.42)	0.91
	rs7302060	1.44 (0.63-3.31)	0.38	1.05 (0.50-2.17)	0.89
TERT	rs2242652	1.03 (0.12-8.93)	0.97	0.60 (0.08-4.52)	0.62
	rs2736098	1.91 (0.41-8.90)	0.40	1.75 (0.49-6.20)	0.38
	rs2736100	1.06 (0.38-2.93)	0.90	**2.43 (1.24-4.75)**	**0.009**
	rs2853676	0 .00 0	0.99	0.44 (0.16-1.20)	0.11

Table 9 Association of circadian pathway genes and TERT gene with prognosis of liposarcoma and leiomyosarcoma in subgroups analysis under the recessive genetic model.

Dominant Model		*Liposarcoma*		*Leiomyosarcoma*	
	SNP ID	**HR (95% CI)**	**P-val.**	**HR (95% CI)**	**P-val.**
CLOCK	rs1801260	0.99 (0.45-2.13)	0.97	0.77 (0.43-1.39)	0.39
	rs3736544	1.24 (0.52-2.96)	0.61	1.12 (0.62-2.02)	0.69
	rs3749474	1.02 (0.46-2.24)	0.95	0.96 (0.53-1.73)	0.91
NPAS2	rs895520	0.82 (0.37-1.77)	0.61	1.79 (0.88-3.64)	0.10
	rs2305160	1.08 (0.50-2.34)	0.83	0.87 (0.49-1.56)	0.65
PER1	rs3027178	0.73 (0.33-1.59)	0.43	0.83 (0.46-1.49)	0.54
PER2	rs934945	0.89 (0.35- 2.21)	0.80	0.58 (0.29-1.14)	0.11
	rs2304674	0.56 (0.24-1.23)	0.19	0.66 (0.34-1.30)	0.23
	rs7602358	1.09 (0.51- 2.32)	0.82	1.33 (0.74-2.38)	0.32
RORA	rs339972	0.73 (0.33-1.63)	0.45	0.89 (0.49-1.62)	0.71
	rs10519097	0.53 (0.20-1.42)	0.20	0.97 (0.53-1.76)	0.92
TIMELESS	rs774027	0.93 (0.40-2.13)	0.87	0.88 (0.46-1.69)	0.71
	rs3809125	1.31 (0.60-2.87)	0.49	1.26 (0.69-2.28)	0.43
	rs7302060	1.40 (0.59- 3.32)	0.44	0.89 (0.48-1.66)	0.72
TERT	rs2242652	0.87 (0.37-2.05)	0.76	0.84 (0.44-1.61)	0.60
	rs2736098	0.99 (0.43-2.30)	0.99	1.27 (0.65-2.49)	0.47
	rs2736100	0.81 (0.30-2.16)	0.68	1.23 (0.59-2.56)	0.56
	rs2853676	0.68 (0.30-1.55)	0.37	1.06 (0.54-2.10)	0.84

Table 10. Association of circadian pathway genes and TERT gene with prognosis of liposarcoma and leiomyosarcoma in subgroups analysis under the dominant genetic model.

Pathway variation analysis

We found a significant association between circadian pathway variation and risk of developing STS (circadian pathway *P* value 0.035). This result was based on the data regarding 12 SNPs located in six genes. The results of the pathway variation analysis are reported in Table 11.

The top genes were *PER2* (2 SNPs, circadian gene *P* value 0.036) and RORA (2 SNPs, circadian gene *P* value 0.050).

Gene	Chr N	SNP	P-value
PER2	2	2	0.036
RORA	15	2	0.050
NPAS2	2	2	0.068
PER1	17	1	0.120
CLOCK	4	2	0.504
TIMELESS	12	3	0.814
PATHWAY		12	0.035

Table 11. Pathway variation analysis

DISCUSSION

This study investigated the role of the genetic variation of clock genes and telomere associated genes in the determinism of predisposition to STS and prognosis of patients affected with STS.

We performed an univariate logistic regression analysis evaluating additive, recessive and dominant genetic models and we found that 4 SNPs are associated with a lower risk of developing soft tissue sarcomas and 2 SNPs with a higher risk of developing STS.

About prognosis of patients affected by STS, we performed a multivariate Cox regression analysis, testing additive, recessive and dominant models and we found that 3 SNPs were associated with survival.

Susceptibility

Previous studies [9,10] suggested that germline genetic variation in the circadian pathway is associated with the risk of developing breast, prostate and lung carcinoma. The implication of most circadian genes in all three tumour types indicates that variation of this pathway could actually be involved in the predisposition to cancer in general, which still requires more investigation to be demonstrated in patients affected with other malignancies. However, it appears that the germline variation of some genes is shared by all three tumour types, whereas the polymorphisms of other genes might be more specific to one or two malignancies. This result support the link between the genetic control of circadian pathway and cancer susceptibility but suggest also that some circadian genes might be more relevant than others in terms of cancer predisposition. Results of a previous meta-analysis on sarcoma genetic variations [13] demonstrated that susceptibility, defined as odds ratio, associated with single variants ranged between 1.35 and 1.48, which are values higher than those usually observed for other malignancies such as breast, colorectal, and gastric carcinomas, which generally include odds ratios between 1.10 and 1.30. Considering also that our results were in line with this observation (mean approximately 1.40), it is possible that germline DNA variation is especially important in the determinism of the predisposition to STS.

So, as suggested for many different types of cancer as breast, prostate, colorectal, ovarian, pancreatic, lung, glioma and non-Hodgkin lymphoma [9-12; 15-20,57], the present study seems to confirm the association between genetic variation of clock genes and susceptibility to STS.

In fact, a statistically significant association with the risk of developing STS was found for 6 SNP of the 19 analyzed. None of the TERT SNPs investigated in the present study resulted statistically significantly associated with sarcoma susceptibility.

In this study carriers of the minor allele of the ***CLOCK SNP rs1801260*** had a reduced predisposition to sarcoma under an additive and a recessive genetic model and to liposarcoma.

CLOCK rs1801260, was studied by many Authors in relation to breast cancer [36, 58, 59], colorectal cancer [34, 60], esophageal carcinoma [61, 62], and gastric cancer [16]. Karantanos and Colleagues [34] investigated the association between polymorphisms in the CLOCK1, PER2, and PER3 genes with the colorectal cancer (CRC) susceptibility and they concluded that *CLOCK* rs1801260 gene significantly increases the risk for CRC development while it does not affect the outcome of CRC patients. A previous meta-analysis on genetic variation of clock genes and cancer risk [11] failed to reveal an association with cancer risk.

PER2 **rs934945** has a missense functional effect causing glycine to glutamic acid substitution in PER2 protein and is probably related to a decreased PER2 activity. In this study, it was associated with a reduced predisposition to sarcoma under an additive and a dominant genetic model and to liposarcoma.

RORA **rs339972** is located on an intron of *RORA* locus. In a previous meta-analysis this SNP was found to be associated to the risk of cancer (summary OR 1.08; 95% CI 1.01–1.15; P=0.02) [10]. Primary meta-analysis relied on 2 datasets of breast cancer patients and one of pancreatic cancer. In the current study *RORA rs339972 SNP* was associated with a decreased predisposition to sarcoma assuming an additive or a dominant genetic model and to leiomyosarcoma.

NPAS2 **rs895520** is an intronic A>G SNP. A previous meta-analysis of four datasets (including 19,865 subjects) revealed a highly significant association with an intermediate level of evidence with cancer risk (summary OR 1.08; 95% CI 1.03–1.13; P=0.001) [10].

This result is also congruent with findings from a NCI genomewide association study on prostate cancer (CGEMS project), not included in the above-mentioned meta-analysis, that showed some significant associations between circadian genes SNPs and risk of prostate cancer. A total of 104 SNPs in circadian genes were included. Eight of these, located in four genes, NPAS2, CSNK1E, CRY1 and CRY2, were significantly associated with prostate cancer susceptibility; *NPAS2 rs895520* was associated with $P \leq 0.01$ [63]. In the present analysis, *NPAS rs895520 SNP* resulted associated with an increased predisposition to sarcoma under an additive and a recessive genetic model and to leiomyosarcoma;

rs2304674 is an intronic SNP of **PER2**. Previous study suggested an association between this SNP with susceptibility to reumathoid arthritis [64]. We found that *PER2 rs2304674* was associated with an increased predisposition to sarcoma under a recessive model and this result represents the first findings in literature about the association between this SNP and cancer susceptibility.

Susceptibility in liposarcoma and leiomyosarcoma subgroups

Subgroup analysis based on histotype showed also potentially interesting results. Assuming an additive genetic model, ***CLOCK*** **rs1801260** and ***PER2*** **rs934945** were statistically significantly associated with liposarcoma, while ***NPAS2*** **rs895520** and ***RORA*** **rs339972** were statistically significantly associated with leiomyosarcoma. Moreover, ***PER1*** **rs3027178** was found to be associated only with liposarcoma susceptibility. It can be argued that different circadian clock genes influence specifically a particular sarcoma subtype. CLOCK and NPAS are cardinal components of the circadian system and their corresponding protein product belongs to the basic helix-loop-helix PAS family of transcription factors and forms heterodimers with ARNTL (BMAL1) to promote target gene expression [9,10,12]. However, while *CLOCK* is mainly expressed in the ‘‘central pacemaker’’ of the circadian system, the suprachiasmatic nucleus of the hypothalamus, *NPAS2* is expressed mainly in the forebrain [12,65]. This may suggest that while these 2 genes are functionally analogous, they might be involved in different circadian-controlled processes. The circadian clock directs nearly all aspects of diurnal physiology, including lipid metabolism and fat cell differentiation [12, 66–68] Recently,

CLOCK and *PER2* polymorphisms have been associated to obesity and the metabolic syndrome [12, 69, 70]. Since it has been proposed that liposarcoma could arise from the dedifferentiation of fat cells [12, 71], it is possible to hypothesize a specific role of *CLOCK* and *PER2* genes in this process.

Prognosis

Genetic variation in the circadian system and in TERT gene has been associated with tumor aggressiveness or patient survival in many different cancers [11, 37, 41]. Our results corroborate the potential association between the genetic variation of clock genes and telomere related genes and prognosis of patient with STS, in details:

- **PER2 rs7602358** SNP was significantly associated with liposarcoma survival, employing an additive genetic model.
- **NPAS2rs 2305160** SNP was associated with liposarcoma survival, assuming a recessive model.
- **TERT rs2736100** SNP was associated with survival in STS patients, assuming an additive model and a recessive genetic model and with leyomiosarcoma survival under a recessive model.

PER2 rs7602358 was previously considered by four research groups [36, 72, 73, 74] evaluating the associations with prostate cancer, breast cancer, and glioma risk. Zhu and Colleagues [72] found a difference in risk association of rs7602358 with prostate cancer between less aggressive and more aggressive prostate cancer subgroup, leading to suppose a role of *PER2* in malignant cells aggressiveness.

About **NPAS2 rs 2305160**, Yuan et al. [75] found that two SNPs, rs1053096 and rs2305160, in the *NPAS2* gene showed significant associations with overall death risk in HCC patients in the recessive model (HR=1.48; 95% CI 1.13–1.94; *P*=0.004) and in the dominant model (HR = 1.63; 95% CI 1.29–2.07; $P < 0.001$), respectively.

The role of **TERT rs2736100 T > G** polymorphism in cancer susceptibility was widely studied. However, to date, there are very little results about the role of this SNP in cancer prognosis.

In 2012, Zou, *et al.* observed a significant association between this polymorphism and overall cancer risk, suggesting that the TERT rs2736100 polymorphism may be a risk factor for cancer [76].

Several meta-analyses published in 2014 and in 2015 associated the *TERT* rs2736100 polymorphism with increased glioma and lung cancer susceptibility [77-79]. Thus, Li and Colleagues [80] performed an updated meta-analysis to more precisely assess the *TERT* rs2736100 polymorphism-cancer association, including 72 studies derived from 61 articles with 269,720 total subjects and their results confirm that the *TERT* rs2736100 polymorphism confers increased overall cancer risk.

About prognosis, Ma et al. [81] found that TERT rs2736100-CC/CA variants represents an independent predictor of a poor prognosis in renal cell carcinoma and so, it may serve as a potential marker for risk stratification.

Pathway variation analysis

With the aim to clarify the molecular mechanisms underlyng disease susceptibility, we also investigate the overall effect of circadian clock gene germline variations on STS risk. This approach permit to detect the combined effects of genetic polymorphisms that are weakly associated with the disease but may not be detected in single-SNP analyses. We found a significant association between circadian pathway variation and risk of developing STS. This association was mostly driven by PER2 and RORA. To date, this is the first time that ARTP-based gene and pathway analysis has been applied to the relationship between circadian genes' germline variation and STS susceptibility [12].

Strength, limitations and future perspectives

To the best of our knowledge, this is the first study investigating the associations of germline circadian clock genes and TERT gene polymorphisms in relation to risk and prognosis of soft tissue sarcoma. Despite the rarity of sarcoma, we evaluated a relatively large cohort of patients with available germline DNA, drawing from our biobank. We used several strategies to select the potentially interesting variations: by using tag SNPs, we were able to efficiently interrogate multiple regions of the *CLOCK* and *TIMELESS* genes.

Our study is also limited by a number of weaknesses. Due to sample size considerations and to our resources, we limited the number of genes selected and SNPs evaluated. We considered three genetic models of inheritance without correcting for multiple testing. Any multiple testing correction would probably nullify statistical significance, nevertheless the strength of those association is noticeably higher than that usually observed for other tumor types. However, we considered this work a good basis for further studies and we did not know a priori the best fitted model. The power of most of our comparisons is around 30–40% and a larger sample would be needed to reach the commonly used statistical power of 80%. We also performed a subgroup analysis for hystotype, considering that, given the complexity of this pathology, it was possible that different variants in CLOCK and TERT genes may influence specifically a particular sarcoma subtype. In fact we found that different histotypes interact with different SNPs.

Finally, the source of controls is mixed, population and hospital based, nevertheless to avoid selection biases for the hospital based fraction we chose patients from different conditions.

The present study represents a starting point for further investigation. On the basis of these results we have the following goals:

-To collect new cases of STS with the aim to increase the sample size. This could permit to obtain more statistically significant results and to validate the present findings.

-To collect new cases of liposarcoma to explore the speculation of the role of CLOCK and PER2 genes in the dedifferentiation of fat cell. This hypothesis need further investigation and could open new perspectives of study.

-To extend the study to other histotypes of STS to explore the hypothesis that different SNPs of different Clock and telomere-related genes may influence specifically a particular sarcoma subtype. This is very important especially considering the heterogeneity of this family of tumors in which more than 50 histotypes were identified and every one presents a different biological behavior.

- To explore the role of a greater number of Clock and especially telomere-related genes SNPs in sarcoma predisposition and prognosis. We obtained significant results about the role of TERT rs2736100 SNP in STS prognosis but we didn't reach any results about the role of this SNP in STS susceptibility. So, we have the intention to explore the role of

other telomere-related genes SNPs. In fact, the determination of SNPs involved in STS predisposition could explain better the mechanicism of disease, considering also the potential interactions across polymorphisms and between genetics and environmental factors. The identification of other SNPs involved in STS prognosis could be the first step for stratification of the risk of disease recurrence which would help us to identify patients who can most benefit from neoadjuvant or/and adjuvant therapy.

So, on the basis of these and future results we have the goal to build a nomogram predicting risk and prognosis of patients affected by STS.

CONCLUSIONS

STS constitute a heterogeneous group of rare tumors. To date, surgery is the only potentially curative treatment in early cases, whereas metastatic disease is virtually incurable. Therefore, early diagnosis, after identification of high risk patients and pre/post-operative therapy, based on fine prognostic stratification, are of crucial importance to increase the survival rates.

The cause of most STS is unknown. Several risk factors for STS have been hypothesized but it is difficult to demonstrate a clear causal relationship. So, in respect of the different biological characteristics of these rare tumors and the absence of definite causes and risk factors, the research about the molecular mechanisms underlying the predisposition to these tumors and their prognosis become almost mandatory.

The present study is the first one to explore the relationship between CLOCK and telomere-related genes SNPs and STS predisposition and prognosis. Our results suggest that genetic variation of CLOCK and telomere-related genes appears to play a role in STS susceptibility and prognosis. The study of the role of genetic variation in cancer predisposition is highly challenging, especially considering the potential interactions between genetics and environmental factors. Obviously other studies are needed to clarify the molecular mechanisms underlying this association. These findings represent a good starting point for further investigation.

REFERENCES

1) Soft Tissue Sarcoma, Version 2.2018, NCCN Clinical Practice Guidelines in Oncology. von Mehren M, Randall RL, Benjamin RS et al. J Natl Compr Canc Netw. 2018 May;16(5):536-563.

2) Soft tissue and visceral sarcomas: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. Casali PG, Abecassis N, Aro HT, at al. ESMO Guidelines Committee and EURACAN. Ann Oncol. 2018 Oct 1;29(Suppl 4):iv51-iv67.

3) Soft tissue limb and trunk sarcomas: diagnosis, treatment and follow-up. Rastrelli M, Tropea S, Basso U et al. Anticancer Res. 2014 Oct;34(10):5251-62

4) Cancer statistics, 2016. Siegel RL, Miller KD, Jemal A et al. J Clin. 2016 Jan-Feb;66(1):7-30

5) ECCO Essential Requirements for Quality Cancer Care: Soft Tissue Sarcoma in Adults and Bone Sarcoma. A critical review. Andritsch E, Beishon M, Bielack S , et al. Crit Rev Oncol Hematol. 2017 Feb;110:94-105.

6) Treatment of soft tissue sarcoma: a focus on earlier stages. Gronchi A, Maki RG, Jones RL. Future Oncol. 2017 Jan;13(1s):13-21.

7) Comprehensive mapping of p53 pathway alterations reveals an apparent role for both SNP309 and MDM2 amplification in sarcomagenesis. Ito M, Barys L, O'Reilly T et al. J. Clin Cancer Res. 2011 Feb 1;17(3):416-26

8) Monogenic and polygenic determinants of sarcoma risk: an international genetic study. Ballinger ML, Goode DL, Ray-Coquard I et al. Lancet Oncol. 2016 Sep;17(9):1261-71.

9) Circadian pathway genetic variation and cancer risk: evidence from genome-wide association studies. Mocellin S, Tropea S, Benna C, et al. BMC Med. 2018 Feb 19;16(1):20.

10) Genetic variation of clock genes and cancer risk: a field synopsis and meta-analysis. Benna C, Helfrich-Förster C, Rajendran S et al. Oncotarget. 2017 Apr 4;8(14):23978-23995.

11) Germline variation of circadian pathway genes and prognosis of gastric cancer patients. Rajendran S, Benna C, Monticelli H et al. Gut. 2018 Apr;67(4):779-780.

12) Associations of clock genes polymorphisms with soft tissue sarcoma susceptibility and prognosis. Benna C, Rajendran S, Spiro G et al. J Transl Med. 2018 Dec 5;16(1):338.

13) Genetic susceptibility to bone and soft tissue sarcomas: afield synopsis and meta-analysis. Benna C, Simioni A, Pasquali S et al. Oncotarget. 2018; 9:18607-18626.

14) Transcriptional architecture of the mammalian circadian clock. Takahashi JS. Nat Rev Genet. 2017;18(3):164–79.

15) Circadian rhythmicity and the influence of 'clock' genes on prostate cancer. Kiss Z, Ghosh PM. Endocrine-related cancer. 2016.

16) Cancer Clocks Out for Lunch: Disruption of Circadian Rhythm and Metabolic Oscillation in Cancer. Altman BJ. Frontiers in cell and developmental biology. 2016; 4: 62.

17) Circadian clocks and breast cancer. Blakeman V, Williams JL, Meng QJ et al. Breast Cancer Res. 2016; 18: 89- 016-0743-z.

18) Circadian Rhythm Disruption Promotes Lung Tumorigenesis. Papagiannakopoulos T, Bauer MR, Davidson SM et al. Cell metabolism. 2016; 24: 324-331.

19) Shift work, circadian gene variants and risk of breast cancer. Grundy A, Schuetz JM, Lai AS et al. Cancer Epidemiol. 2013;37(5):606–12.

20) Common genetic variation in circadian rhythm genes and risk of epithelial ovarian cancer (EOC). Jim HS, Lin HY, Tyrer JP et al. J Genet Genome Res. 2015;2(2):017.

21) Period3 structural variation: a circadian biomarker associated with breast cancer in young women. Zhu Y, Brown HN, Zhang Y et al. Cancer Epidemiol Biomarkers Prev . 2005 Jan;14 (1):268-70.

22) Telomerase and the search for the end of cancer. Mocellin S, Pooley KA, Nitti D. Trends Mol Med. 2013 Feb;19(2):125-33.

23) Telomerase reverse transcriptase locus polymorphisms and cancer risk: a field synopsis and meta-analysis. Mocellin S, Verdi D, Pooley KA et al. J Natl Cancer Inst. 2012 Jun 6;104(11):840-54.

24) How telomeres solve the end-protection problem. De Lange T. Science. 2009; 326 (5955) : 948 – 952 .

25) Telomere diseases. Calado RT , Young NS. N Engl J Med. 2009 ; 361 (24) : 2353 – 2365 .

26) Telomeres and human disease: ageing, cancer and beyond . Blasco MA. Nat Rev Genet. 2005 ; 6 (8) : 611 – 622 .

27) Telomerase and cancer therapeutics. Harley CB. Nat Rev Cancer. 2008; 8(3) :167-179 .

28) Telomerase in cancer immunotherapy. Liu JP , Chen W , Schwarer AP , Li H. Biochim Biophys Acta. 2010; 1805 (1): 35-42 .

29) Differences in telomere length between sporadic and familial cutaneous melanoma. Menin C, Bojnik E, Del Bianco P et al. Br J Dermatol. 2016 Nov;175(5):937-943.

30) Targetable Alterations in Adult Patients With Soft-Tissue Sarcomas: Insights for Personalized Therapy. Lucchesi C, Khalifa E, Laizet Y et al. JAMA Oncol. 2018 Oct 1;4(10):1398-1404

31) Human Sarcomas Are Mosaic for Telomerase-Dependent and Telomerase-Independent Telomere Maintenance Mechanisms: Implications for Telomere-Based Therapies. April R.S. Gocha, Gerard Nuovo, Obiajulu H. Iwenofu et al. Am J Pathol. 2013 Jan; 182(1): 41–48.

32) Genetic and epigenetic associations of circadian gene TIMELESS and breast cancer risk. Fu A, Leaderer D, Zheng T et al. MolCarcinog. 2012;51(12):923–9.

33) The role of polymorphisms in circadian pathway genes in breast tumorigenesis. Breast. Dai H, Zhang L, Cao M, Song F, et al. Cancer Res Treat. 2011;127(2):531–40.

34) Association of the clock genes polymorphisms with colorectal cancer susceptibility. Karantanos T, Theodoropoulos G, Gazouli M et al. J SurgOncol. 2013;108(8):563–7.

35) A combined analysis of genome-wide association studies in breast cancer. Li J, Humphreys K, Heikkinen T, Aittomaki K et al. Breast Cancer Res Treat. 2011;126(3):717–27.

36) Breast cancer risk, nightwork, and circadian clock gene polymorphisms. Truong T, Liquet B, Menegaux F et al. EndocrRelat Cancer. 2014;21(4):629–38.

37) A functional polymorphism in PER3 gene is associated with prognosis in hepatocellular carcinoma. Zhao B, Lu J, Yin J et al. Liver Int. 2012;32(9):1451–9.

38) Prognostic relevance of telomere length and telomerase reverse transcriptase variant (rs2242652) on the multiple myeloma patients. Aref S, Al Saeed A, El Menshawy N et al. J Clin Lab Anal. 2020 Apr; 34(4): e23133.

39) Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Stig E Bojesen 1 , Karen A Pooley, Sharon E Johnatty et al. Nat Genet. 2013;45(4):371-84, 384e1-2.

40) The TERT promoter mutation incidence is modified by germline TERT rs2736098 and rs2736100 polymorphisms in hepatocellular carcinoma. Xiaotian Yuan, Guanghui Cheng, Jingya Yu et al. Oncotarget. 2017; 8(14): 23120–23129.

41) *TERT* rs2853676 polymorphisms correlate with glioma prognosis in Chinese population. Xue He, Yahui Wei, Zhengshuai Chen et al. Oncotarget. 2016; 7(45): 73781–73791.

42) Association between the *TERT* Genetic Polymorphism rs2853676 and Cancer Risk: Meta-Analysis of 76 108 Cases and 134 215 Controls. Jin-Lin Cao, Ping Yuan, Abudumailamu Abuduwufuer et al. PLoS One. 2015; 10(6): e0128829.

43) Functional dissection of breast cancer risk-associated *TERT* promoter variants. Helbig S, Wockner L, Bouendeu A et al. Oncotarget. 2017 Sep 15; 8(40): 67203–67217.

44) Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. Rodriguez S, Gaunt TR, Day IN. Am J Epidemiol. 2009;169(4):505–14.

45) Pathway analysis by adaptive combination of P-values. Yu K, Li Q, Bergen AW, Pfeiffer RM et al. Genet Epidemiol 2009;33(8):700–9.

46) Mutations in the circadian gene CLOCK in colorectal cancer. Alhopuro P, Bjorklund M, Sammalkorpi H, et al. Mol Cancer Res. 2010;8(7):952–60.

47) The circadian gene NPAS2, a putative tumor suppressor, is involved in DNA damage response. Hoffman AE, Zheng T, Ba Y, Zhu Y. Mol Cancer Res. 2008;6(9):1461–8.

48) A role for the clock gene per1 in prostate cancer. Cao Q, Gery S, Dashti A et al. Cancer Res. 2009;69(19):7619–25.

49) Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers. Chen ST, Choo KB, Hou MF. Carcinogenesis. 2005;26(7):1241–6.

50) The analysis of deregulated expression and methylation of the PER2 genes in gliomas. Fan W, Chen X, Li C, Yongluo, et al. J Cancer Res Ther. 2014;10(3):636–40.

51) Expression of core clock genes in colorectal tumour cells compared with normal mucosa: a systematic review of clinical trials. Fonnes S, Donatsky AM, Gogenur I. Colorectal Dis. 2015;17(4):290–7.

52) The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. Fu L, Pelicano H, Liu J et al. Cell. 2002;111(1):41–50.

53) Correlated downregulation of estrogen receptor beta and the circadian clock gene Per1 in human colorectal cancer. Mostafaie N, Kallay E, Sauerzapf E et al. MolCarcinog. 2009;48(7):642–7.

54) The circadian clock: pacemaker and tumour suppressor. Fu L, Lee CC. Nat Rev Cancer. 2003;3(5):350–61.

55) RORalpha suppresses breast tumor invasion by inducing SEMA3F expression. Xiong G, Wang C, Evers BM et al. Cancer Res. 2012;72(7):1728–39.

56) RORA, a large common fragile site gene, is involved in cellular stress response. Zhu Y, McAvoy S, Kuhn R et al. Oncogene. 2006;25(20):2901–8.

57) A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Petersen GM, Amundadottir L, Fuchs CS et al. Nat Genet. 2010;42(3):224–8.

58) CLOCK in breast tumorigenesis: genetic, epigenetic, and transcriptional profiling analyses. Hoffman AE, Yi CH, Zheng T et al. Cancer Res. 2010;70(4):1459–68.

59) Circadian genes and breast cancer susceptibility in rotating shift workers. Monsees GM, Kraft P, Hankinson SE et al. Int J Cancer. 2012;131(11):2547–52.

60) Functional polymorphisms of circadian positive feedback regulation genes and clinical outcome of Chinese patients with resected colorectal cancer. Zhou F, He X, Liu H et al. Cancer. 2012;118(4):937–46.

61) Circadian variability of pharmacokinetics of 5-fluorouracil and CLOCK T3111C genetic polymorphism in patients with esophageal carcinoma. Miki I, Tamura T, Nakamura T et al. Ther Drug Monit. 2005;27(3):369–74.

62) Favorable genetic polymorphisms predictive of clinical outcome of chemoradiotherapy for stage II/III esophageal squamous cell carcinoma in Japanese. Okuno T, Tamura T, Yamamori M et al.Am J ClinOncol. 2007;30(3):252–7.

63) Genomewide association study of prostate cancer identifies a second risk locus at 8q24. Yeager M, Orr N, Hayes RB, et al. Nat Genet. 2007;39(5):645–9.

64) PER2 is downregulated by the LPS-induced inflammatory response in synoviocytes in rheumatoid arthritis and is implicated in disease susceptibility. Lee H, Nah SS, Chang SH et al. Mol Med Rep. 2017;16(1):422-428.

65) NPAS2: an analog of clock operative in the mammalian forebrain. Reick M, Garcia JA, Dudley C et al. Science. 2001;293(5529):506–9.

66) Circadian regulation of glucose, lipid, and energy metabolism in humans. Poggiogalle E, Jamshed H, Peterson CM Metabolism. 2018;84:11–27.

67) Circadian metabolism: from mechanisms to metabolomics and medicine. Brown SA. Trends EndocrinolMetab. 2016;27(6):415–26.

68) A transcriptional circuit filters oscillating circadian hormonal inputs to regulate fat cell differentiation. Bahrami-Nejad Z, Zhao ML, Tholen S et al. Cell Metab. 2018;27(4):854–868.

69) PERIOD2 variants are associated with abdominal obesity, psychobehavioral factors, and attrition in the dietary treatment of obesity. Garaulet M, Corbalan-Tutau MD, Madrid JA et al. J Am Diet Assoc. 2010;110(6):917–21.

70) Association between polymorphisms in the Clock gene, obesity and the metabolic syndrome in man. Scott EM, Carter AM, Grant PJ. Int J Obes(Lond). 2008;32(4):658–62.

71) Liposarcoma: molecular targets and therapeutic implications. Scott EM, Casadei L, Prudner BC et al. Cell Mol Life Sci. 2016;73(19):3711–8.

72) Testing the circadian gene hypothesis in prostate cancer: a population-based case-control study. Zhu Y, Stevens RG, Hoffman AE et al. Cancer Res. 2009;69(24):9315–22.

73) Circadian clock genes and risk of fatal prostate cancer. Markt SC, Valdimarsdottir UA, Shui IM et al. Cancer Causes Control. 2015;26(1):25–33.

74) Circadian pathway genes in relation to glioma risk and outcome. Madden MH, Anic GM, Thompson RC et al. Cancer Causes Control. 2014;25(1):25–32.

75) Functional polymorphisms in the NPAS2 gene are associated with overall survival in transcatheter arterial chemoembolization-treated hepatocellular carcinoma patients. Yuan P, Wang S, Zhou F et al. Cancer Sci. 2014;105(7):825–32

76) The TERT rs2736100 polymorphism and cancer risk: a meta-analysis based on 25 case-control studies. Zou P, Gu A, Ji G et al.BMC Cancer. 2012; 12: 7.

77) Association between TERT rs2736100 polymorphism and lung cancer susceptibility: evidence from 22 case-control studies. Yuan Y, Lu C, Xue L et al. Tumour Biol. 2014 May;35(5):4435-42.

78) Increased lung cancer risk associated with the TERT rs2736100 polymorphism: an updated meta-analysis. Yang J, Jiao S, Yang J et al. Tumour Biol. 2014 Jun;35(6):5763-9.

79) Telomerase reverse transcriptase (TERT) rs2736100 polymorphism contributes to increased risk of glioma: evidence from a meta-analysis. Peng Z, Tian D, Chen Q et al. Int J Clin Exp Med. 2015 Jan 15;8(1):422-30.

80) The TERT rs2736100 polymorphism increases cancer risk: A meta-analysis. Li H, Xu Y, Mei H et al. Oncotarget. 2017 Jun 13;8(24):38693-38705.

81) The TERT locus genotypes of rs2736100-CC/CA and rs2736098-AA predict shorter survival in renal cell carcinoma. Ma R, Liu C, Lu M et al. Urol Oncol. 2019 May;37(5):301.e1-301.e10.