



Article

Alteration of the Nucleotide Excision Repair (NER) Pathway in Soft Tissue Sarcoma

Adriano Pasqui ^{1,†}, Anna Boddi ^{2,†}, Domenico Andrea Campanacci ^{2,3}, Guido Scoccianti ², Andrea Bernini ^{4,*,‡}, Daniela Grasso ⁴, Elisabetta Gambale ⁵, Federico Scolari ², Iliaria Palchetti ⁶, Annarita Palomba ⁷, Sara Fancelli ⁸, Enrico Caliman ⁸, Lorenzo Antonuzzo ^{1,4,8,‡} and Serena Pillozzi ^{1,8}

- ¹ Medical Oncology Unit, Careggi University Hospital, 50134 Florence, Italy; apasqui.onclinica@gmail.com (A.P.); lorenzo.antonuzzo@unifi.it (L.A.); serena.pillozzi@unifi.it (S.P.)
- ² Orthopaedic Oncology Unit, Careggi University Hospital, 50134 Florence, Italy; anna.boddi@gmail.com (A.B.); domenicoandrea.campanacci@unifi.it (D.A.C.); scocciantig@aou-careggi.toscana.it (G.S.); scolari.fed@gmail.com (F.S.)
- ³ Orthopaedic Oncology Unit, Careggi University Hospital, Department of Health Sciences, University of Florence, 50134 Florence, Italy
- ⁴ Department of Biotechnology, Chemistry and Pharmacy, University of Siena, 53100 Siena, Italy; daniela.grasso@student.unisi.it
- ⁵ Clinical Oncology Unit, Careggi University Hospital, 50134 Florence, Italy; gambalee@aou-careggi.toscana.it
- ⁶ Department of Chemistry Ugo Schiff, University of Florence, 50019 Sesto Fiorentino, Italy; ilia.palchetti@unifi.it
- ⁷ Histopathology and Molecular Diagnostic Unit, Careggi University Hospital, 50134 Florence, Italy; annarita.palomba@aouc.unifi.it
- ⁸ Department of Experimental and Clinical Medicine, University of Florence, 50134 Florence, Italy; sara.fancelli@unifi.it (S.F.); enrico.caliman@unifi.it (E.C.)
- * Correspondence: andrea.bernini@unifi.it
- † These authors contributed equally to this work.
- ‡ These authors contributed equally to this work.



Citation: Pasqui, A.; Boddi, A.; Campanacci, D.A.; Scoccianti, G.; Bernini, A.; Grasso, D.; Gambale, E.; Scolari, F.; Palchetti, I.; Palomba, A.; et al. Alteration of the Nucleotide Excision Repair (NER) Pathway in Soft Tissue Sarcoma. *Int. J. Mol. Sci.* **2022**, *23*, 8360. <https://doi.org/10.3390/ijms23158360>

Academic Editor: Gianpaolo Papaccio

Received: 16 May 2022

Accepted: 25 July 2022

Published: 28 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Clinical responses to anticancer therapies in advanced soft tissue sarcoma (STS) are unluckily restricted to a small subgroup of patients. Much of the inter-individual variability in treatment efficacy is as result of polymorphisms in genes encoding proteins involved in drug pharmacokinetics and pharmacodynamics. The nucleotide excision repair (NER) system is the main defense mechanism for repairing DNA damage caused by carcinogens and chemotherapy drugs. Single nucleotide polymorphisms (SNPs) of NER pathway key genes, altering mRNA expression or protein activity, can be significantly associated with response to chemotherapy, toxicities, tumor relapse or risk of developing cancer. In the present study, in a cohort of STS patients, we performed DNA extraction and genotyping by SNP assay, RNA extraction and quantitative real-time reverse transcription PCR (qPCR), a molecular dynamics simulation in order to characterize the NER pathway in STS. We observed a severe deregulation of the NER pathway and we describe for the first time the effect of SNP rs1047768 in the ERCC5 structure, suggesting a role in modulating single-stranded DNA (ssDNA) binding. Our results evidenced, for the first time, the correlation between a specific genotype profile of ERCC genes and proficiency of the NER pathway in STS.

Keywords: nucleotide excision repair (NER); soft tissue sarcoma; ERCC; SNP; pharmacogenomic

1. Introduction

Soft tissue sarcomas (STS) are a heterogeneous group of rare malignancies with mesenchymal origin with an incidence of approximately 50–60 new cases per year for every million people, 20% with synchronous metastasis at diagnosis [1,2]. Approximately 30–40% of localized STS will finally develop metastasis. Generally, surgery with or without radiation, even if in selected cases of localized disease, represents the ideal treatment for

STS [3]. For locally advanced or metastatic STS, the therapeutic options are limited and clinical responses to chemotherapy (ChT) are uncommon, while the median of survival ranges between 18 and 22 months [3]. Anthracycline-based therapy is the backbone of first-line therapy for advanced disease and there are no formal data that show that the combination of ChT regimens is related to an improved overall survival (OS) than single-agent ChT with doxorubicin [4–6]. A higher response rate and longer progression-free survival (PFS) can be expected in a number of selected for probable better sensitivity histological subtypes in light of evidence from several randomized, clinical trials [7–9]. However, clinical responses to anticancer therapies continue to be limited to a small subset of STS patients. Thus, the detection of biomarkers that can predict drug response and toxicity can be useful to explain the basis for the differences in treatment efficacy and toxicity among patients.

In the last decade, pharmacogenetic and pharmacogenomic approaches have been increasingly applied to the genetic characterization of STS and have started to offer insights that may be considered to modulate and, hopefully, improve the effectiveness and safety of the currently used treatment regimens [10].

Most cancer therapy regimens damage DNA, killing cancer cells by inducing damage that interferes with DNA replication and transcription, finally activating cell death pathways. The efficient repair of DNA damage can render cells resistant to killing. The DNA repair pathways involved in acquired resistance mechanisms vary from drug to drug, and all of the major DNA repair pathways can alter the sensitivity of cancer cells to DNA damaging agents and thus confer chemoresistance. It has been suggested that deficiencies in DNA repair capacity could have a role in affecting the response to DNA damaging agents [11–13]. At least six major DNA repair pathways, as well as numerous sub-pathways, are involved in the repair of damage [11,14]. In particular, the nucleotide excision repair (NER) pathway (Figure 1) repairs bulky lesions and has been associated with tumor progression and response to chemotherapy [15,16]. Excision repair cross-complementing (ERCC) group 1 (*ERCC1/XPA*), group 2 (*ERCC2/XPD*), group 4 (*ERCC4/XPF*) and group 5 (*ERCC5/XPG*) are members of the NER pathway. It has been demonstrated that a deficiency of DNA repair capacity of NER genes can be related to the presence of single nucleotide polymorphisms (SNPs) that result in altered mRNA expression or protein activity [14]. A deficiency in the ability to repair DNA damages could therefore alter the response to treatment with a significant positive association between polymorphism *ERCC2* gene, Lys751Gln, but also *ERCC1*, Asn118Asn with the improved cisplatin response in osteosarcoma [17]. Moreover, emerging evidence associates the expression levels of genes involved in the NER process and some types of cancer. A study of 55 patients and a larger study of 744 patients with squamous cell carcinoma of head and neck revealed that patients had a reduced DNA repair capacity, reduced expression of *DDB1* and *ERCC3* genes, compared to healthy controls, resulting in an increased risk of developing cancer [18,19]. In patients with colorectal cancer *ERCC1*, *ERCC2*, *ERCC5* and *DDB2* genes are more expressed in tumor tissue than in matched normal tissue, while *ERCC4* is downregulated [20]. A bioinformatics analysis performed on the GSE28460 and GSE18497 expression datasets showed that, in pediatric acute lymphoblastic leukemia patients, NER pathway gene expression is significantly increased at relapse; moreover, at diagnosis, NER pathway gene expression is higher in early relapsing subpopulation than in the late relapsing group [21].

In the present study, in order to identify and characterize the role of the NER pathway in STS, we evaluate the expression of key genes (*ERCC1*, *ERCC2* and *ERCC5*) and a panel of SNPs within these genes in a cohort of STS patients. Finally, with a molecular dynamics simulation, we predicted the possible effect of selected SNPs on the secondary and tertiary structure of NER proteins.

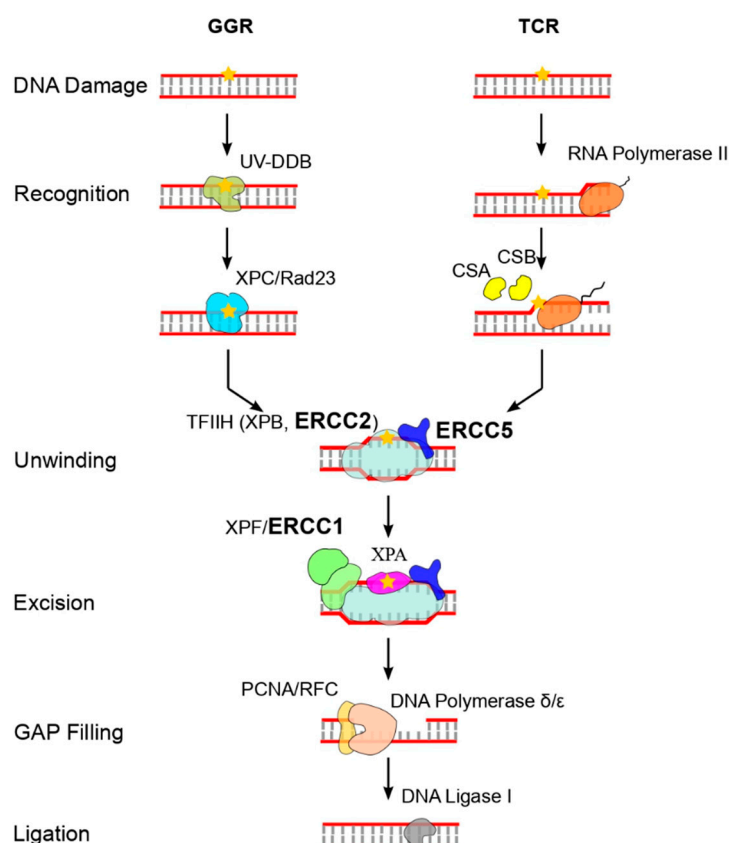


Figure 1. Nucleotide excision repair pathway (NER). GGR, global genomic NER; TCR, transcription coupled NER.

2. Materials and Methods

2.1. STS Patients and Samples

A total of 32 STS tissues, swab and plasma samples were obtained from patients enrolled at the Oncology and Reconstructive Orthopedics SOD, AOU Careggi (RESEARCH study) after written consent. The main clinical and pathological characteristics of the patients are summarized in Table 1. STS tissue samples were used for total DNA and RNA extraction. Swab samples were used for genomic DNA extraction.

Table 1. Clinicopathological features of the patients enrolled with STS.

Clinicopathological Features	Cohort (N = 32)
Gender	Male = 62.5% Female = 37.5%
Age at surgery (years)	Mean = 68.4 Max = 93 Min = 39
BMI (kg/m ²)	Mean = 25.5 Max = 44.3 Min = 19.2
Surgical margins	Large = 90.6% Marginal = 9.4%
Primary/Relapse	Primary = 81% Relapse = 19%
Grade	High = 78.0% Low = 22.0%

Table 1. *Cont.*

Clinicopathological Features	Cohort (N = 32)
Hystotype	Liposarcoma = 28.1%
	Pleomorphic sarcoma = 37.5%
	Leiomyosarcoma = 9.4%
	Myxofibrosarcoma = 12.5%
	Fibrosarcoma = 3.1%
	Fibromyxoid sarcoma = 3.1%
	Dermatofibrosarcoma protuberans = 3.1%
Anatomical site	Undifferentiated pleomorphic cell sarcoma = 3.1%
	Thigh = 62.5%
	Foot = 3.1%
	Gluteal-sacral region = 6.25%
	Scapular region = 3.1%
	Deltoid region = 3.1%
	Abdominal region = 3.1%
	Pectoral region = 6.25%
Upper limb = 12.5%	

Note. BMI, body mass index.

2.2. Genomic and Somatic DNA Extraction and Genotyping by SNP Assay

Genomic and somatic DNA was isolated from buccal swabs and tumor tissues using a QIAamp DNA Blood mini kit (Qiagen, Manchester, UK). Allelic discrimination was performed using TaqMan SNP Genotyping Assays (Applied Biosystems, Forster City, CA, USA) on the Rotor-gene 6000 instrument (Qiagen, Manchester, UK). The assay consisted of two allele-specific minor groove binding (MGB) probes, labeled with the fluorescent dyes VIC and FAM.

Real-time PCR was performed in 15 μ L reaction mixtures containing 7.5 μ L of Taq-path ProAmp MasterMix (Applied Biosystems, Forster City, CA, USA). TaqMan-MGB genotyping assay mixes were supplied at 40 \times concentration and 2 μ L of sample DNA. Thermocycler conditions were an initial hold step for 30 s at 60 $^{\circ}$ C and 5 min at 95 $^{\circ}$ C, followed by 40 cycles of 95 $^{\circ}$ C for 15 s and 60 $^{\circ}$ C for 60 s, finally a last post read step for 30 s at 60 $^{\circ}$ C. Genotypes were analyzed by measuring allele-specific fluorescence using the Rotor-gene Q Series software 2.3.5 for allelic discrimination (Qiagen, Manchester, UK).

2.3. RNA Extraction and Quantitative Real-Time Reverse Transcription PCR (qPCR)

Total RNA was extracted from STS tissue samples using the RNeasy Mini kit (Qiagen, Manchester, UK). The quantity and the quality of RNA were evaluated using a Nanodrop spectrophotometer version 3.5 (NanoDrop Technologies, Wilmington, DE, USA). A total of 50 ng of RNA from each sample was retro-transcribed with the PrimeScriptTM RT Reagent Kit (Takara Bio, Otsu, Shiga, Japan); the resulting cDNA was then used for qPCR analysis with "Powerup Sybr Green" (Applied Biosystems, Forster City, CA, USA) on a Rotor Gene Q (Qiagen, Manchester, UK). The relative quantification was performed using LinRegPCR software (version 11.0, <http://LinRegPCR.nl>) and the data were normalized to 18s rRNA. The primers used are listed in Table 2.

Table 2. Sequences of qPCR primers.

Gene Name	Primer Forward	Primer Reverse	Lenght (bp)
ERCC1	5' CTCAAGGAGCTGGCTAAGATGT 3'	5' CATAGGCCTTGTAGGCTCTCCAG 3'	90
ERCC2	5' CTGGAGGTGACCAAACTCATCTA 3'	5' CCTGCTTCTCATAGAAGTTGAGC 3'	100
ERCC5	5' CAGACACAGCTCCGAATTGA 3'	5' TTCTGGGTTTTTCGTTTTGC 3'	209
18S	5' AACCCGTTGAACCCCAT 3'	5' CCATCCAATCGGTAGTAGCG 3'	151

Note: bp, base pair.

2.4. Molecular Dynamics Simulations

Molecular dynamics simulations were carried out in the AMBER99sb force field employing the GROMACS version 2022.2 [22] software suite. For ERCC2, the cryo-EM structures 6NMI and 5OF4 were employed. The PDB entry 6TUX was taken as starting structure for the ERCC5-DNA complex; the model of activated ERCC5-DNA complex was built using the DNA in the T4 RNase H crystal structure (PDB 2IHN). Mutants were generated by PyMOL (Schrödinger Inc., New York, NY, USA) mutagenesis subroutine. Each structure was placed in a triclinic box of dimensions exceeding protein size by 1 nm each side after alignment along inertia axes; boxes were filled with TIP3P water molecules and counterions were added until neutrality. The simulation protocol involved an initial restrained minimization step, followed by a series of equilibration MD simulations in the NVT ensemble, during which temperature was gradually increased from 0 to 300 K, and restraints on protein heavy atoms were gradually revealed. Furthermore, unrestrained equilibration in the NPT ensemble was performed during 500 ps, prior to a 1000 ns production stage. The equations of motion were integrated with a time step of 2.0 fs. Periodic boundary conditions were used, in conjunction with the particle-mesh Ewald (PME) method. All X-H bonds were constrained using the LINCS algorithm. Snapshots were dumped to a trajectory file each picosecond. The trajectory was submitted for further analysis after the routine evaluation of correctness (i.e., RMSD and RMSF).

2.5. Statistical Analysis

Statistical analysis was performed using RStudio on relative expression values obtained from qPCR, genotyping data derived from real time PCR and clinical pathological characteristics of enrolled patients. χ^2 tests were performed to compare the distribution of selected SNPs between our patient cohort and the general population. Haplotype-base evaluation of SNPs in *ERCC1*, *ERCC2* and *ERCC5* genes was performed using the HaploView software version 4.2 (<http://www.broad.mit.edu/haploview>). The hierarchical clustering heatmaps were created with the pheatmap package (version 1.0.12) in RStudio using Euclidean correlation, while the correlation matrix was created with the GGally package (version 2.1.2) using Pearson's correlation. Levene's test significance on the dataset disproved the assumption of equality of variance; therefore, the relative expression values of the three *ERCC* genes in every patient was analyzed with Welch's *t*-test.

3. Results

3.1. Clinicopathological Features of OS Patients

The clinicopathological features of the 32 patients affected by STS (62.5% male and 37.5% female) enrolled in the study are illustrated/shown in Table 1. The median age of the patients is 68.4 years (range of 39–93), 20 patients (62.5%) had tumor localized in the thigh, 1 in foot (3.1%), 2 in gluteal-sacral region (6.25%), 1 in scapular region (3.1%), 1 in deltoid region (3.1%), 1 in abdominal wall (3.1%), 2 in pectoral region (6.25%) and 4 in upper limb (12.5%). Regarding the histotype, 9 (28.1%) patients had liposarcoma (4 pleomorphic and 5 myxoid), 12 patients (37.5%) had pleomorphic sarcoma, 3 patients leiomyosarcoma (9.4%), 4 patients myxofibrosarcoma (12.5%), 1 patient fibrosarcoma (3.1%), 1 fibromyxoid sarcoma (3.1%) and 1 dermatofibrosarcoma protuberans (3.1%), while 1 was undifferentiated pleomorphic cell sarcoma (3.1%); in addition, 25 (78%) patients were affected by high-grade STS and 7 patients were affected by low-grade STS (22%); 6 (19%) patients relapsed. The median value of Body Mass Index (BMI) was 25.5 kg/m² (range 19.2–44.3). Finally, large surgical margins were registered in 90.6% of cases.

3.2. Evaluation of Germinal SNP Panel in NER Genes

Given the abundant evidence on the correlation between polymorphisms in NER genes and different types of cancer, we decided to evaluate the possible role of SNPs in the NER pathway in STS patients. In our cohort, we analyzed, at the germline level, a panel of SNPs belonging to relevant genes of the NER pathway: the synonymous variant

rs11615 (chr19:45420395) of *ERCC1*; the stop gained variant rs13181 (chr19:45351661) and the missense variant rs1799793 (chr19:45364001) for *ERCC2*; and the missense variant rs1047768 (chr13:102852167) and the upstream variant rs2296147 (chr13:102846025) for *ERCC5* (Table 3). One patient was excluded from the analysis as the amount of DNA extracted from his swab sample was not sufficient. We then compared the allele frequencies of our cohort to the healthy population (<https://www.ncbi.nlm.nih.gov/snp/>, accessed on 2 March 2022). The germline allele and genotype frequencies of the selected SNPs registered in our cohort of STS patients are summarized in Table 3. Regarding the SNP rs11615 (A>G) in the *ERCC1* gene, the homozygous reference condition (AA) is present in 19%, the heterozygous condition (AG) is present in 58% and the alternative homozygous condition (GG) in 23% of STS patients (A = 48 %; G = 52 %); the alternative allele has a higher frequency than the general population. Regarding the SNP rs13181 in the *ERCC2* gene (T>G; TT = 19%, TG = 45%, GG = 36%; T = 42%, G = 58%), the alternative allele has a higher frequency than the general population. Regarding the SNP rs1799793 in the *ERCC2* gene (C>T; CC = 19%, CT = 55%, TT = 26%; C = 47%, T = 53%), the alternative allele has a higher frequency than the general population. For the SNP rs1047768 in the *ERCC5* gene (T>C; TT = 19%, TC = 45%, CC = 36%; T = 42%, C = 58%), the alternative allele has the same frequency as the general population. Finally, regarding the SNP rs2296147 in the *ERCC5* gene (T>C; TT = 32%, TC = 49%, CC = 19%; T = 56%, C = 44%), the alternative allele has a lower frequency than the general population. For both SNPs in the *ERCC2* gene, this difference was statistically significant (Table 3) (rs13181 $\chi^2 = 9.72$ $p = 0.0018$; rs1799793 $\chi^2 = 9.02$ $p = 0.0027$). The presence or absence of the studied variants are summarized in Figure 2, which shows that the rs11615 variant in the *ERCC1* gene is present in 25 out of 31 patients, the rs13181 and rs1799793 variants in the *ERCC2* gene are present in 25 patients and finally the rs1047768 and rs2296147 variants in the *ERCC5* gene are present in 25 and 20 patients, respectively. In our patient cohort, 11 (35.5%) patients had all SNPs analyzed, while only 1 patient had no variations.

Table 3. Allelic and genotypic frequencies of SNPs in DNA extracted from buccal swabs.

Gene	ercc1	ercc2		ercc5		
SNP ID	rs11615	rs13181	rs1799793	rs1047768	rs2296147	
Assay ID	C_2532959_20	C_3145033_10	C_3145050_10	C_1891769_20	C_1891782_20	
Chromosome Location	chr19:45420395	chr19:45351661	chr19:45364001	chr13:102852167	chr13:102846025	
Variant	A>G	T>G	C>T	T>C	T>C	
SNP Type	Synonymous variant	Stop gained variant	Missense variant	Missense variant	Upstream variant	
Genotype	Homozygous Reference	AA = 6 (19%)	TT = 6 (29%)	CC = 6 (19%)	TT = 6 (19%)	TT = 10 (32%)
	Heterozygotes	AG = 18 (58%)	TG = 14 (45%)	CT = 17 (55%)	TC = 14 (45%)	TC = 15 (49%)
	Alternative Homozygotes	GG = 7 (23%)	GG = 10 (36%)	TT = 7 (26%)	CC = 11 (36%)	CC = 6 (19%)
STS cohort alleles frequencies	Reference	48%	42%	47%	42%	56%
	Alternative	52%	58%	53%	58%	44%
General population alleles frequencies *	Reference	59%	64%	68%	42%	54%
	Alternative	41%	36%	32%	58%	46%
χ^2 Test	χ^2	2.43	9.72	9.02	0.00	0.08
	p	0.1190	0.0018	0.0027	1.0000	0.7762

* <https://www.ncbi.nlm.nih.gov/snp/>; <https://www.snpedia.com/>; <https://www.pharmgkb.org/>, accessed on 2 March 2022; STS, soft tissue sarcoma; chr, chromosome.

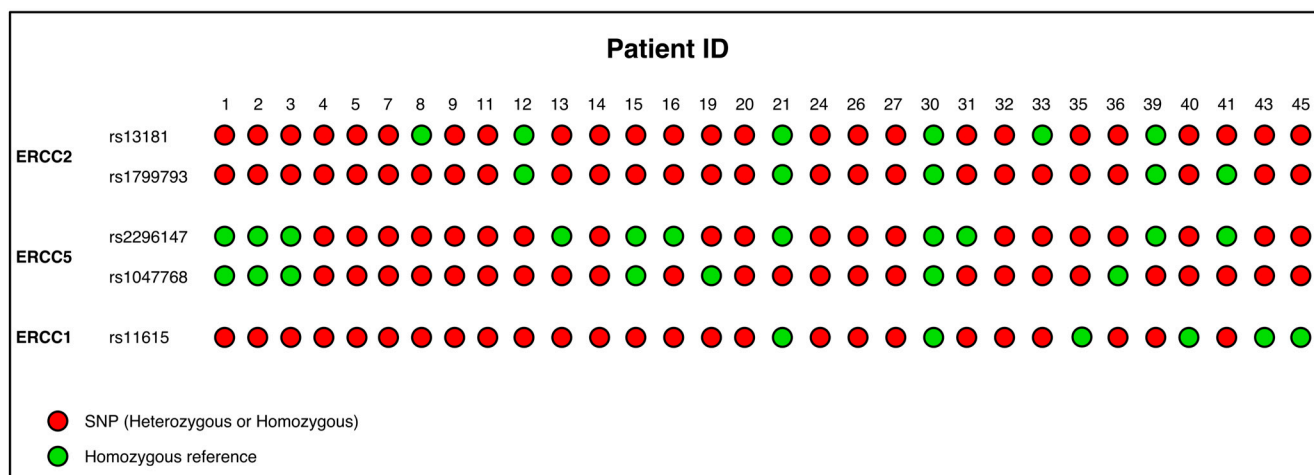


Figure 2. Presence/absence of the ERCC variants in the STS cohort. STS, soft tissue sarcoma.

3.3. Linkage Disequilibrium and Haplotypes

We calculated the linkage disequilibrium (LD) coefficients (D' and r^2) to evaluate the non-random association of alleles at different loci, and frequency of haplotypes using Haploview version 4.2 (Mark Daly’s Laboratory, Massachusetts Institute of Technology/Harvard Broad Institute, Cambridge, MA, USA) in ERCC genes. As shown in Figure 3A, the LD value between the variants of ERCC5 is 54, indicating a low value of linkage disequilibrium (white block) and the haplotype with the highest frequency is CC. For the variants concerning ERCC1 and ERCC2, as shown in Figure 3B, we found that the ERCC2 variants rs13181 and rs1799793 show a higher LD value, 69 (red block), instead of ERCC1 variants. The haplotype with the highest frequency for ERCC1-ERCC2 variants is TGG.

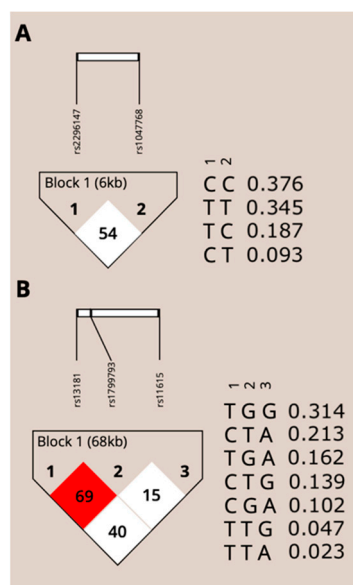


Figure 3. (A) Linkage disequilibrium (LD) block and haplotype frequencies for rs2296147 and rs1047768 in ERCC5 gene. (B) LD block and haplotype frequencies for rs13181, rs1799793 and rs11615 in ERCC1 and ERCC2 genes.

3.4. Evaluation of Somatic SNP Panel in NER Genes in STS Patients

In addition to germline SNPs, somatic variations in ERCC genes have also been shown to play a role in different tumor types, such as non-small cell lung cancer [23] and urothelial carcinoma [24]. In the cohort of patients described above, we investigated whether the

presence of a wild-type genotype recurs also in the somatic line, so we analyzed the same variants at the somatic level. For *ERCC1* rs11615 and *ERCC2* rs13181 and rs1799793, the homozygous reference condition was also maintained at the somatic level in all patients, whereas in *ERCC5* rs2296147 and rs1047768, we found that in a 12,5% and 20%, respectively, of homozygous reference patients there was a change in genotype to the mutated form. Frequencies revealed in the somatic line are summarized in Table 4.

Table 4. Genotypic frequency of SNPs in the somatic line.

Gene	SNP ID	Homozygotes Reference (%)	Heterozygotes (%)	Homozygotes Alternative (%)
ERCC1	rs11615	100	0	0
ERCC5	rs1047768	80	20	0
ERCC5	rs2296147	87.5	12.5	0
ERCC2	rs13181	100	0	0
ERCC2	rs1799793	100	0	0

3.5. Expression Analysis of NER Genes in STS Patients by qPCR

Emerging evidence associates the expression levels of genes involved in the NER process and some types of cancer, while data on STS are limited, so we evaluated the gene expression of three key genes (*ERCC1*, *ERCC2* and *ERCC5*) in the NER pathway in 24 tissue samples from the 32 patients in our cohort. As illustrated in Figure 4, the gene *ERCC1* is overexpressed in 75% (18/24) of STS samples analyzed compared to healthy corresponding tissue with a median expression value of 0.432 (range of -2.32 ; 2.68). The *ERCC2* gene is overexpressed in 50% (12/24) STS with a median expression value of 0.497 (range -7.04 ; 3.04). Finally, the *ERCC5* gene is overexpressed in 42% (10/24) STS samples analyzed with a median expression value of -0.607 (range of -4.66 ; 2.22).

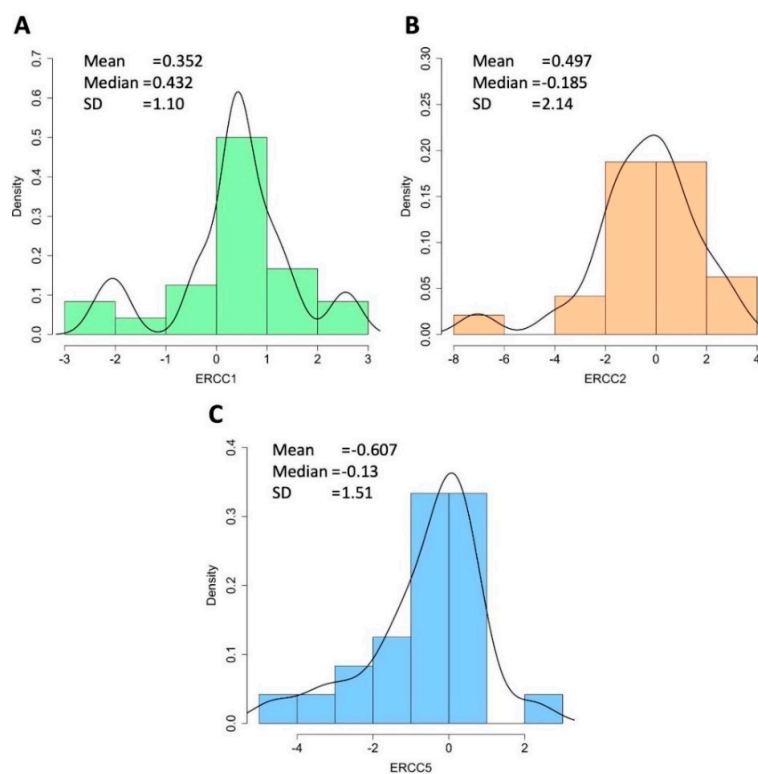


Figure 4. Expression plots showing trends of individual NER genes in STS samples. (A) *ERCC1*; (B) *ERCC2*; (C) *ERCC5*. SD, standard deviation; STS, soft tissue sarcoma.

We then investigated if a distinct genotype combination of *ERCC1*, *ERCC2* and *ERCC5* genes could influence the expression levels of these genes in the tumor tissue.

We divided the samples by whether they presented the wild-type allele in a SNP and compared the expression level of the resulting two groups (Figure 5). We found that patients with AA or AG genotypes in the rs11615 SNP have lower expression of the *ERCC1* gene than mutated homozygous GG (A); patients with TT or TG genotypes in the rs13181 SNP have lower expression of the *ERCC2* gene than mutated homozygous GG (B); patients with CC or CT genotypes in the rs1799793 SNP have significantly lower expression of the *ERCC2* gene than patients with both mutated alleles TT (C, $p = 0.0316$); and finally, patients with TT or CT genotype in the rs2296147 and rs1047768 have a slightly higher expression of the *ERCC5* gene than mutated homozygous CC (D,E). The expression values of *ERCC1*, *ERCC2* and *ERCC5* were compared with each other using R, and the results were represented as a heatmap as Pearson's correlation values (Figure 6B), showing the high level of correlation between the expression of *ERCC2* and *ERCC5*.

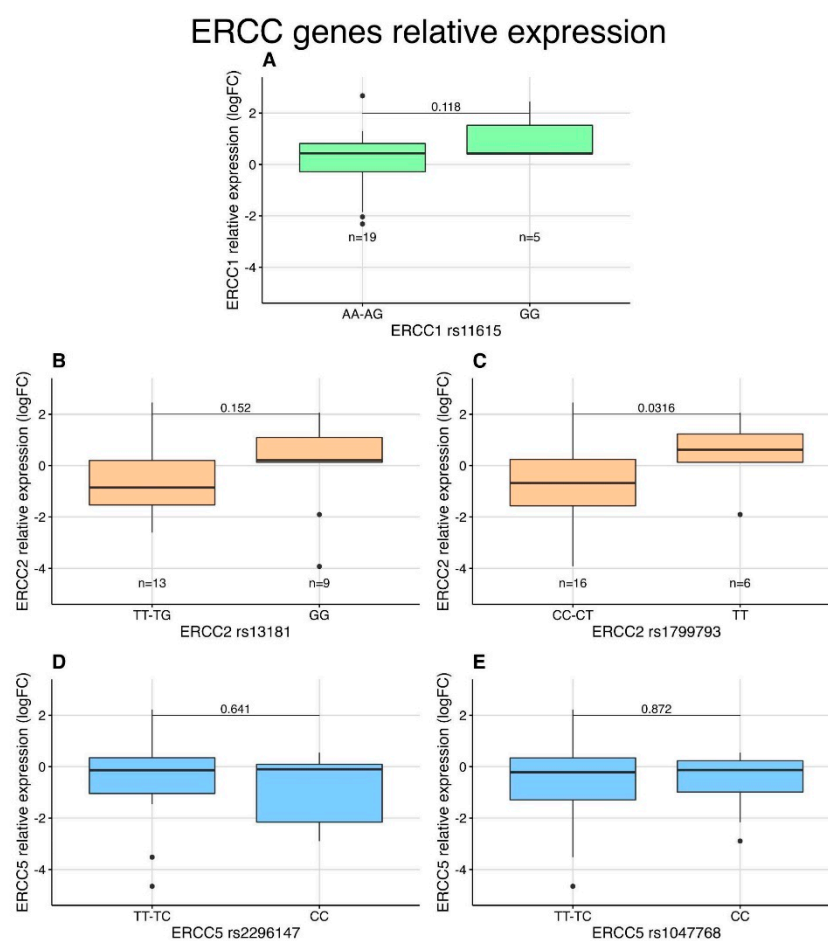


Figure 5. Correlation analysis between relative gene expression and genotype in 5 different SNPs (A) *ERCC1*_rs11615, (B) *ERCC2*_rs13181, (C) *ERCC2*_rs1799793, (D) *ERCC5*_rs2296147, (E) *ERCC*_rs1047768. The relative gene expression is presented as log2 Fold Change. In every plot, the samples were separated according to the presence (left side) or absence (right side) of at least one wild-type allele. The significance value at the top of every plot was obtained via Welch's *t*-test. FC, fold change.

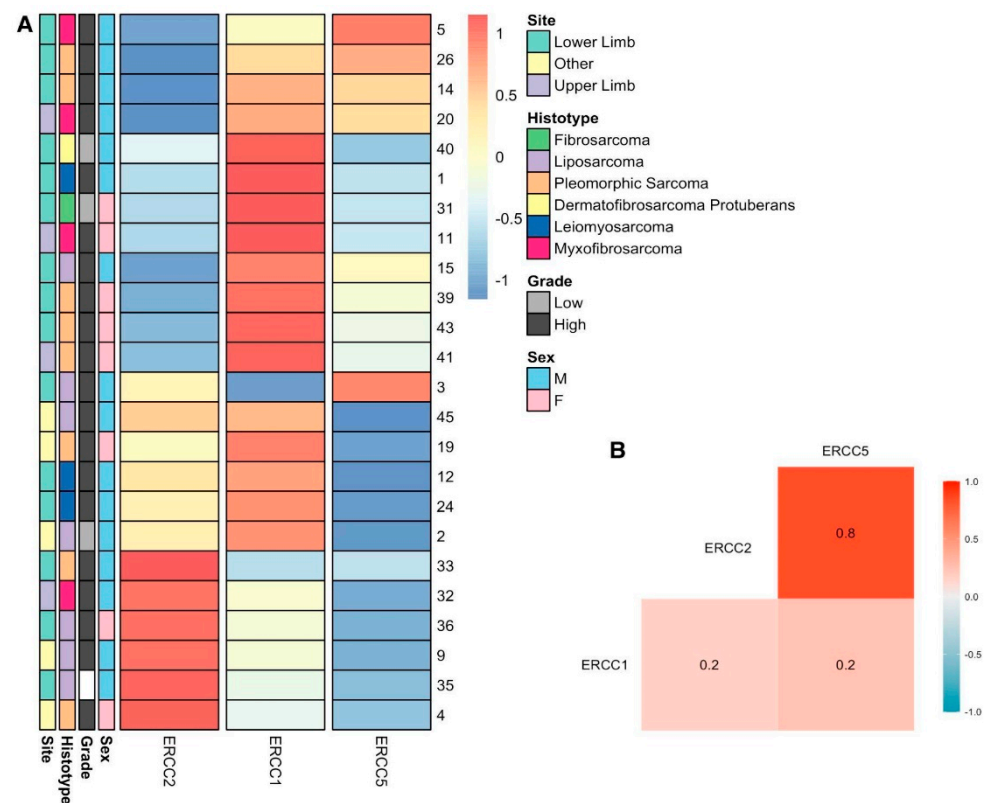


Figure 6. (A) Relative expression clusters of ERCC1, ERCC2 and ERCC5 per patient. The rows represent each sample and its respective ID, sex, histotype and grade. The columns represent the relative expression of ERCC1, ERCC2 and ERCC5 as log Fold Change values. (B) Correlation matrix of ERCC1, ERCC2 and ERCC5 relative gene expression with Pearson's correlation similarity. M, male; F, female.

We compared the expression level of each gene in each sample using the R package pheatmap. We observed that the expression level does not generate a cluster of samples in their respective histotype, sex or grade (Figure 6A).

A statistical analysis was then performed to see how the characteristics of the patients are correlated to the expression levels of the NER genes, and we observed that, in our cohort, female patients had significantly higher expression levels of *ERCC1* than male patients ($p = 0.02$).

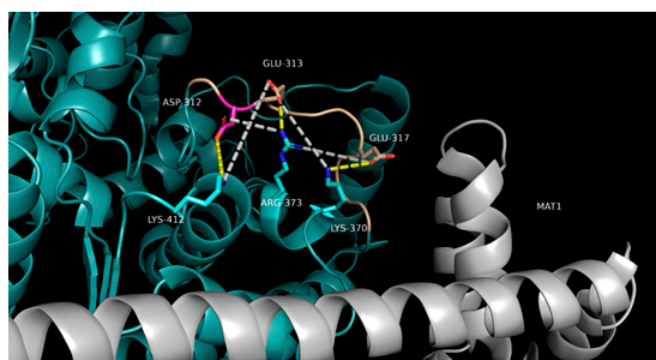
3.6. Correlation between *TP53* and NER Genes

It has been reported that allelic variation at SNPs on the *TP53* gene affects *ERCC5* transcript and this could result in variation in NER function in normal bronchial epithelial cells [25]; in particular, T allele at SNP rs2296147 upregulates and C allele at SNP rs17655 downregulates *ERCC5* expression levels. The common SNP sites rs751402 and rs2296147 reside in the *ERCC5* 5'-untranslated region (UTR). Variation at rs2296147 is predicted to alter the binding of the *TP53* transcription factor [26]. Notably, rs2296147 T allele plays part in the formation of an in silico predicted *TP53* transcription factor-binding site and that site is predicted to be lost when the C allele is present. Based on these premises, we evaluated the *TP53* gene for the SNP rs1042522 a missense variant (G>C) in our cohort of STS patients; the homozygous reference condition was present in 5%, the heterozygous condition was present in 50% and the alternative homozygous condition in 45% of STS patients. Then, we evaluated the correlation between *TP53* genotyping and expression level of *ERCC5* and other *ERCC* genes, finding no relevant differences.

3.7. Effect of SNPs on the Tertiary Structure of NER Proteins

The SNPs rs13181 and rs1799793 (*ERCC2*) and rs1047768 (*ERCC5*) were evaluated for their effect on protein structure folding and stability, and also in view of linkage disequilibrium. In *ERCC2*, Asp312 is part of a markedly negatively charged loop spanning residues Pro311-Pro320 (Figure 7A), with three carboxylic sidechains driving the conformation thanks to the apparent ionic pairs: Asp312-Lys412, Glu313-Arg373 and Glu317-Lys370. Moreover, more cross-interactions between opposite charges are established within the 12 Å limit of the Coulomb cutoff threshold, forming a surface electrostatic interaction network stabilising the local structure. Indeed, the molecular dynamics simulation of the native *ERCC2* showed an unusual stability level for such a loop, comparable to that of a secondary structure (Figure 8). Consequently, the simulation of the Asp312Asn mutant (rs1799793) shows such segment to be destabilized by the loss of negative charge and the consequent disruption of the electrostatic network. The asparagine residue cannot offer a replacement interaction such as, e.g., hydrogen bonding because the distance to neighbouring sidechains exceeds even the largely allowed 3.4 acceptor–donor distance. The other variant of *ERCC2* (rs13181) is of the stop gain type and causes the loss of the C-terminal segment (Figure 7B) starting at Lys751; half of the α -helix and the final β -strand Ala756-Q758 and relative pairing with Ala479-Phe481 are lost. Molecular dynamics simulation showed such loss of secondary structures does not disrupt the overall folding, but affects local stability significantly (Figure 7B). Although both SNPs seem not to disrupt the overall folding, a cooperative effect in lowering the stability of tertiary structure, with a possible cascade effect on Transcription factor IIIH (TFIIH) complex formation, can be suggested to explain the linkage disequilibrium effect.

A



B

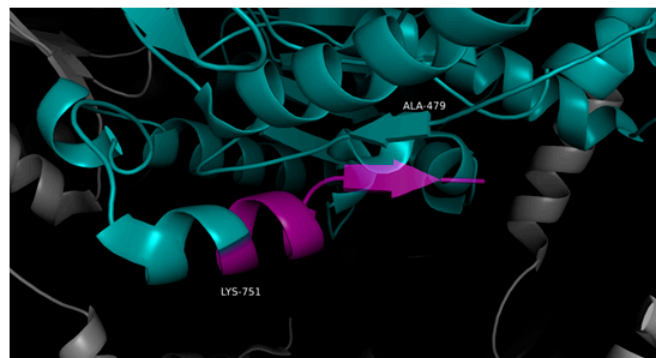


Figure 7. (A) Model of the *ERCC2* electrostatic interaction network between the negative loop Pro311-Pro320 and positive charges; mutant Asp312Asn (position in magenta) destabilise the interaction and loop. (B) Loss of C-terminal α -helix and β -sheet (transparent segment in magenta) following mutation Lys751*.

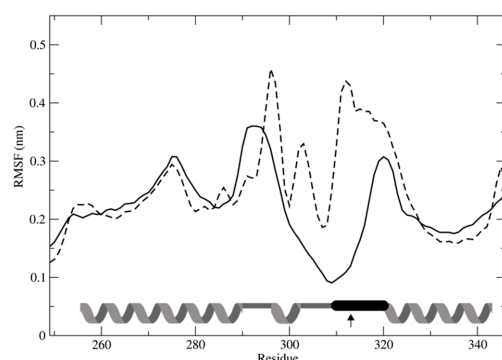


Figure 8. Protein local flexibility is represented with the root mean square fluctuation profile (RMSF, nm) for wild type (solid line) and Asp312Asn mutant (dashed line) as calculated from the molecular dynamics simulations. The secondary structure is reported at the bottom (helix for α -helices, line for loops). Loop segment Pro310-Pro320 (black thick line) shows flexibility lower than helices thanks to the extended network of electrostatic interactions. Point mutation Asp312Asn (indicated by arrow) causes electrostatic network disruption and segment flexibility to increase up to 4-fold. The RMSF of the remaining structure is not affected, suggesting only local conformational change occurs without hampering the overall folding.

The aminoacidic position His46 of ERCC5 is central to the hydrophobic wedge (residues 31–67), including helix $\alpha 2$ and helix $\alpha 3$ and their connecting loop, which is mostly disordered. A positively charged canyon defined by the hydrophobic wedge and β -pin motifs was proposed to provide a DNA-binding surface to the well-known helical arch formed by α -helices $\alpha 4$ (residues 82–97) and $\alpha 6$ (residues 750–779) [27] and H2TH motif (residues 848–880) [28]. To test the possible role of wedge His46 in binding DNA, ERCC5 in complex with the 3' side of the repair bubble was modelled, using the structure of the FEN1 homolog T4 RNase H, where the path of both ssDNA strands beyond the junction was observed [29]. The model was refined by molecular dynamics simulation, resulting in a clash-free structure showing both DNA strands fit in protein crevices. The undamaged ssDNA fits in the crevice formed by the hydrophobic wedge and the β -pin motifs. Such crevice was already reported to accommodate the phosphate backbone by interaction with basic amino acids H60 and R61 in the wedge helix $\alpha 3$, K828 in the β -pin, and K972 in $\alpha 15$. Our model showed the His46 to contribute to the binding with a similar interaction but with the solvent-exposed phosphate of the first decoupled nucleobase, stabilizing the interaction at ss/dsDNA junction level (Figure 9). The model of the His46Gln mutant showed a shift of the wedge dynamics back to the disordered state possibly affecting the wedge effectiveness in opening the double strand.

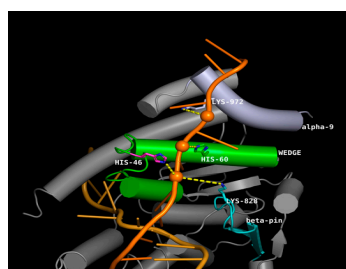


Figure 9. Model of activated ERCC5 using the DNA in the T4 RNase H crystal structure (PDB 2IHN). Close-up views of the gateway and the canyon from the model shown in panel B. Residues mutated in silico are colored in magenta. The undamaged ssDNA fits in the crevice formed by the hydrophobic wedge (green), the β -pin motif (cyan) and α -helix 9 (light blue). The crevice accommodates the phosphate backbone and is stabilised by charge interactions with bottom Lys828, His80 and Lys972; the MD simulation also showed wedge flexible loop to participate in binding by His46 driven closure onto the phosphate backbone. The His46Gln mutant largely lost such interaction.

4. Discussion

The scientific community, still aims to improve and optimize the management of STSs, because of their heterogeneity and rarity. The pathogenetic and histopathological knowledge of STS has increased in recent decades, but unfortunately without the desired impact on patient survival rates. However, a higher response rate and longer PFS can be expected in a number of selected for probable better sensitivity histological subtypes in light of evidence from several randomised clinical trials. For example, doxorubicin plus dacarbazine is an optional multi-agent first-line regimen for leiomyosarcoma, in which the activity of ifosfamide is debated [6]. Gemcitabine represents a therapeutic option, alone or in combination with docetaxel [30]. Notably, a phase III study compared single-agent doxorubicin with the combination of gemcitabine–docetaxel as first-line regimen in advanced STS patients of all types. The combination did not show improvement in PFS and objective response rate (ORR) and is not recommended as upfront therapy for advanced STS patients [8], including uterine LMS. Taxanes can be an effective treatment options for angiosarcoma, which has shown high sensitivity to this ChT [9]. Neurotrophic tyrosine receptor kinase (NTRK) inhibitor larotrectinib represent the standard treatment of those rare patients with locally advanced or metastatic NTRK rearranged sarcomas [31]. Second-line treatments include pazopanib and chemotherapeutic agents such as trabectedin, eribulin, and gemcitabine combinations. The median PFS in first-line therapy for STS ranges between 4.5 and 6.0 months, whereas the median PFS for most of the second-line drugs is below 5 months [32–34]. Therefore, there is an urgent need to increase the knowledge behind these diseases and to develop more effective diagnostic and prognostic strategies.

The analysis of data in the literature and the importance of DNA repair mechanisms in tumor progression led us to the characterization of the NER pathway, the DNA repair pathway often implicated in chemoresistance processes [35–37], in a cohort of STS patients enrolled at the Orthopaedic Oncology Unit of Careggi University Hospital. It has been demonstrated that efficiency changes of DNA repair capacity of NER genes can be related to the presence of single nucleotide polymorphisms (SNPs) that result in altered mRNA expression or protein activity [14]. A deficiency in the ability to repair DNA damages could therefore alter the response to treatment with a significant positive association between the polymorphism *ERCC2* gene, Lys751Gln, but also *ERCC1*, Asn118Asn with the improved cisplatin response in osteosarcoma [17]. Additionally, variant alleles in NER SNPs *ERCC2* Lys751Gln, *ERCC1* 8092 C/A and *ERCC1* 118C/T were individually associated with esophageal adenocarcinoma risk [38]. *ERCC2* rs50871 G/T, *ERCC6* rs1917799 G/T and *DDB2* rs3781619 A/G polymorphisms were significantly associated with shorter OS, while *ERCC1* rs3212961 A/C, *ERCC5* rs2094258 A/G and *DDB2* rs300083 C/G could predict favorable OS of gastric cancer patients [39]. The *ERCC2* rs3810366 and rs238406 polymorphisms were shown to significantly enhance Wilms tumor susceptibility [40]. There is also evidence in the literature on the association between SNPs in NER genes and STS, e.g., *ERCC5* and *ERCC1* status represented a potential DNA repair signature that could be used for the prediction of clinical response to trabectedin in STS [41] and to irifolven in urothelial cancer [42]. In addition, the genotyping of the SNPs rs17655 (Lys751Gln) in *ERCC5* and rs13181 (Asp1104His) in *ERCC2* revealed the potential implication of the Asp1104His polymorphism in the occurrence of chromosomal translocations associated with specific sarcoma subtypes [43].

These data led us to evaluate a panel of SNPs in germline *ERCC* genes in a cohort of STS, considering that the alternative allele for SNP rs11615 in the *ERCC1* gene and the alternative allele for SNPs (rs13181 and rs1799793) in the *ERCC2* gene has a higher frequency in STS patients versus the distribution in the general healthy population and NER pathway is deeply affected since 97% (30 out of 31) of STS patients have at least one variation in *ERCC* genes. In addition, rs1799793 and rs13181 variants on *ERCC2* shown a higher linkage value, as also reported in previous studies conducted with different populations [44,45].

We then evaluated the presence of somatic mutations on *ERCC* genes in those patients presenting a homozygous reference genotype, observing that 20% (rs1047768) and 12.5%

(rs2296147) of patients have somatic variant on *ERCC5*; therefore, both germline and somatic line are involved in our cohort of patients.

The altered expression of *ERCC* genes has been associated with increased risk of colorectal cancer [46], where *ERCC1*, *ERCC2* and *ERCC5* were overexpressed, and invasive cervical carcinoma [47], where *ERCC1* and *ERCC2* were underexpressed. In addition, the higher expression levels of the *ERCC1* gene have been linked with poor response of FOLFOX-based chemoradiation in colorectal cancer [48] and with the aggressiveness of esophageal cancer [49]. Furthermore, *ERCC1* expression has been correlated with trabectedin sensitivity in STS patients [50]. Data available in the literature on this type of rare tumor are extremely limited, and the most relevant dataset to date is TCGA-SARC (Supplementary Materials Table S1 and Figures S1–S3) [51]. In this study, it was found that *ERCC1*, *ERCC2* and *ERCC5* were expressed in all of the STS tissue samples selected within our cohort. Subsequent genotyping of *ERCC* genes was performed to determine how differences at the germline level could be associated with altered expression levels and the results show that the *ERCC2* gene has significantly lower expression values in patients presenting the wild-type C allele in the rs1799793 variant.

A subset of the observed SNPs (namely, rs13181, rs1799793 and rs1047768) on NER pathway genes, being of the missense or stop gained type, were suitable for the structural modeling of the induced three-dimensional structures alteration of NER proteins. The correlation with protein folding stability, also in view of linkage disequilibrium, is here discussed. The SNPs rs13181 and rs1799793 code for Asp312Asn and Lys751* mutants, located in the Arch domain and in the C-terminal CTE domain, respectively. The Arch domain is known to be mechanistically important for transcription and DNA repair via its binding to MAT1 [52]; dynamics alteration induced by Asp312Asn affects the loop closely interacting with MAT1 (Figure 7). The CTE domain has been suggested to form an interaction surface with the p44 protein [53], and modeling of Lys751* mutant showed a large disruption of such surface. Taking everything into consideration, Asp312Asn and Lys751* modeled as double mutant showed no effect on overall protein folding stability, but the co-operative local alterations of shape and dynamics affecting the interaction network also explain linkage disequilibrium. Coming to *ERCC5*, we reported how His46 contributes to the binding of ssDNA with the solvent-exposed phosphate of the first decoupled nucleobase, stabilizing the interaction at ss/dsDNA junction level (Figure 9). Molecular dynamics simulation the His46Gln mutation (rs1047768) showed a shift of the wedge dynamics back to the disordered state possibly affecting the wedge effectiveness in opening the double strand. Although the critical role of many basic residues in the wedge area has been described [26], to our knowledge, this is the first structural rationale proposed for His46 mutation.

In conclusion, we registered a severe deregulation of NER pathway in STS patients, with alterations in genotyping and expression of *ERCC* genes, i.e., the correlation between *ERCC2* expression and its genotype. Finally, the effect of SNP rs1047768 in the *ERCC5* structure was hypothesized for the first time in this study suggesting a role in modulating ssDNA binding, a result that needs to be further validated, improving knowledge and optimizing the management of STSs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23158360/s1>.

Author Contributions: Conceptualization, S.P.; writing—original draft preparation, A.P. (Adriano Pasqui), G.S. and A.B. (Anna Boddi); writing—review and editing, A.B. (Andrea Bernini), L.A. and D.A.C.; funding acquisition, E.G., D.A.C. and L.A.; supervision, S.P.; formal analysis, F.S.; data curation, E.G., G.S., A.P. (Annarita Palomba), S.F., E.G. and E.C.; investigation, A.B. (Anna Boddi), A.P. (Adriano Pasqui), A.B. (Andrea Bernini), I.P. and D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Regione Toscana Bando Salute 2018, grant number D78D20000870002.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki. The protocol was approved by the Comitato Etico Regionale for clinical experimentation of Toscana region (Italy), Area Vasta Centro section, number: 16933_bio (9 July 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available upon request from the corresponding author.

Acknowledgments: The authors are grateful to Irida Gyshi for help in revising the English language.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stiller, C.A.; Trama, A.; Serraino, D.; Rossi, S.; Navarro, C.; Chirilaque, M.D.; Casali, P.G.; RARECARE Working Group. Descriptive epidemiology of sarcomas in Europe: Report from the RARECARE project. *Eur. J. Cancer* **2013**, *49*, 684–695. [[CrossRef](#)]
2. Sbaraglia, M.; Bellan, E.; Tos, A.P.D. The 2020 WHO Classification of Soft Tissue Tumours: News and perspectives. *Pathologica* **2020**, *113*, 70–84. [[CrossRef](#)] [[PubMed](#)]
3. Casali, P.G.; Abecassis, N.; Bauer, S.; Biagini, R.; Bielack, S.; Bonvalot, S.; Boukovinas, I.; Bovee, J.V.M.G.; Brodowicz, T.; Broto, J.; et al. ESMO Guidelines Committee and EURACAN. Soft tissue and visceral sarcomas: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2018**, *29*, iv51–iv67. [[CrossRef](#)] [[PubMed](#)]
4. Antman, K.; Crowley, J.; Balcerzak, S.P.; E Rivkin, S.; Weiss, G.R.; Elias, A.; Natale, R.B.; Cooper, R.M.; Barlogie, B.; Trump, D.L. An intergroup phase III randomized study of doxorubicin and dacarbazine with or without ifosfamide and mesna in advanced soft tissue and bone sarcomas. *J. Clin. Oncol.* **1993**, *11*, 1276–1285. [[CrossRef](#)] [[PubMed](#)]
5. Judson, I.; Verweij, J.; Gelderblom, H.; Hartmann, J.T.; Schöffski, P.; Blay, J.-Y.; Kerst, J.M.; Sufliarsky, J.; Whelan, J.; Hohenberger, P.; et al. Doxorubicin alone versus intensified doxorubicin plus ifosfamide for first-line treatment of advanced or metastatic soft tissue sarcoma: A randomised controlled phase 3 trial. *Lancet Oncol.* **2014**, *15*, 415–423. [[CrossRef](#)]
6. D’Ambrosio, L.; Touati, N.; Blay, J.; Grignani, G.; Flippot, R.; Czarnecka, A.M.; Piperno-Neumann, S.; Martin-Broto, J.; Sanfilippo, R.; Katz, D.; et al. European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group. Doxorubicin plus dacarbazine, doxorubicin plus ifosfamide, or doxorubicin alone as a first-line treatment for advanced leiomyosarcoma: A propensity score matching analysis from the European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group. *Cancer* **2020**, *126*, 2637–2647. [[CrossRef](#)] [[PubMed](#)]
7. Stacchiotti, S.; Palassini, E.; Sanfilippo, R.; Vincenzi, B.; Arena, M.; Bochicchio, A.M.; De Rosa, P.; Nuzzo, A.; Turano, S.; Morosi, C.; et al. Gemcitabine in advanced angiosarcoma: A retrospective case series analysis from the Italian Rare Cancer Network. *Ann. Oncol.* **2012**, *23*, 501–508. [[CrossRef](#)] [[PubMed](#)]
8. Seddon, B.; Strauss, S.J.; Whelan, J.; Leahy, M.; Woll, P.J.; Cowie, F.; Rothermundt, C.; Wood, Z.; Benson, C.; Ali, N.; et al. Gemcitabine and docetaxel versus doxorubicin as first-line treatment in previously untreated advanced unresectable or metastatic soft-tissue sarcomas (GeDDiS): A randomised controlled phase 3 trial. *Lancet Oncol.* **2017**, *18*, 1397–1410. [[CrossRef](#)]
9. Penel, N.; Bui, B.N.; Bay, J.-O.; Cupissol, D.; Ray-Coquard, I.; Piperno-Neumann, S.; Kerbrat, P.; Fournier, C.; Taieb, S.; Jimenez, M.; et al. Phase II trial of weekly paclitaxel for unresectable angiosarcoma: The ANGIOTAX Study. *J. Clin. Oncol.* **2008**, *26*, 5269–5274. [[CrossRef](#)] [[PubMed](#)]
10. Gambale, E.; Boddi, A.; Pasqui, A.; Campanacci, D.A.; Scoccianti, G.; Palchetti, I.; Bernini, A.; Antonuzzo, L.; Pillozzi, S. Pharmacogenomics of soft tissue sarcomas: New horizons to understand efficacy and toxicity. *Cancer Treat. Res. Commun.* **2022**, *31*, 100528. [[CrossRef](#)] [[PubMed](#)]
11. Nagel, Z.D.; Kitange, G.J.; Gupta, S.K.; Joughin, B.A.; Chaim, I.A.; Mazzucato, P.; Lauffenburger, D.A.; Sarkaria, J.N.; Samson, L.D. DNA Repair Capacity in Multiple Pathways Predicts Chemoresistance in Glioblastoma Multiforme. *Cancer Res.* **2017**, *77*, 198–206. [[CrossRef](#)]
12. Lindahl, T.; Wood, R.D. Quality control by DNA repair. *Science* **1999**, *286*, 1897–1905. [[CrossRef](#)] [[PubMed](#)]
13. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **2001**, *411*, 366–374. [[CrossRef](#)] [[PubMed](#)]
14. Moura, D.S.; Peña-Chilet, M.; Varela, J.A.C.; Alvarez-Alegret, R.; Agra-Pujol, C.; Izquierdo, F.; Ramos, R.; Ortega-Medina, L.; Martin-Davila, F.; Castilla-Ramirez, C.; et al. A DNA damage repair gene-associated signature predicts responses of patients with advanced soft-tissue sarcoma to treatment with trabectedin. *Mol. Oncol.* **2021**, *15*, 3691–3705. [[CrossRef](#)] [[PubMed](#)]
15. Rao, D.; Mallick, A.B.; Augustine, T.; Daroqui, C.; Jiffry, J.; Merla, A.; Chaudhary, I.; Seetharam, R.; Sood, A.; Gajavelli, S.; et al. Excision repair cross-complementing group-1 (ERCC1) induction kinetics and polymorphism are markers of inferior outcome in patients with colorectal cancer treated with oxaliplatin. *Oncotarget* **2019**, *10*, 5510–5522. [[CrossRef](#)] [[PubMed](#)]
16. Suzuki, T.; Sirimangalakitti, N.; Baba, A.; Toyoshima-Nagasaki, R.; Enomoto, Y.; Saito, N.; Ogasawara, Y. Characterization of the nucleotide excision repair pathway and evaluation of compounds for overcoming the cisplatin resistance of non-small cell lung cancer cell lines. *Oncol. Rep.* **2022**, *47*, 70. [[CrossRef](#)]
17. Biondo, P.; Hattinger, C.; Innocenti, F.; Talamini, R.; Alberghini, M.; Scotlandi, K.; Zanusso, C.; Serra, M.; Toffoli, G. Nucleotide excision repair gene variants and association with survival in osteosarcoma patients treated with neoadjuvant chemotherapy. *Pharmacogenomics J.* **2011**, *12*, 476–483. [[CrossRef](#)]

18. Han, P.; Gao, F.; Liu, H.; Liu, Z.; Shi, Q.; Troy, J.D.; Owzar, K.; Lee, W.; Zevallos, J.P.; Sturgis, E.M.; et al. Reduced mRNA expression of nucleotide excision repair genes in lymphocytes and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis* **2017**, *38*, 504–510. [[CrossRef](#)]
19. Ren, P.; Niu, X.; Liu, C.; Liu, J.; Li, H.; Zhao, Q.; Xing, J.; Bai, Y.; Liang, Y.; Han, P. Associations between expression levels of nine core nucleotide excision repair genes in lymphocytes and risk of head and neck squamous cell carcinomas in a Chinese population. *Int. J. Clin. Oncol.* **2019**, *25*, 660–669. [[CrossRef](#)]
20. Zhang, Y.; Wu, S.; Zhou, X.; Huang, F.; Chen, R.; Wang, Y.; Wu, J. Association between nucleotide excision repair gene polymorphism and colorectal cancer risk. *J. Clin. Lab. Anal.* **2019**, *33*, e22956. [[CrossRef](#)]
21. Ibrahim, O.M.; Sobeai, H.M.A.; Grant, S.; Latimer, J.J. Nucleotide excision repair is a predictor of early relapse in pediatric acute lymphoblastic leukemia. *BMC Med Genom.* **2018**, *11*, 95. [[CrossRef](#)] [[PubMed](#)]
22. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. [[CrossRef](#)] [[PubMed](#)]
23. Aiello, M.M.; Solinas, C.; Santoni, M.; Battelli, N.; Restuccia, N.; Latteri, F.; Paratore, S.; Verderame, F.; Albanese, G.V.; Bruzzi, P.; et al. Excision Repair Cross Complementation Group 1 Single Nucleotide Polymorphisms and Nivolumab in Advanced Non-Small Cell Lung Cancer. *Front. Oncol.* **2020**, *10*, 1167. [[CrossRef](#)]
24. Van Allen, E.M.; Mouw, K.W.; Kim, P.; Iyer, G.; Wagle, N.; Al-Ahmadie, H.; Zhu, C.; Ostrovnaya, I.; Kryukov, G.V.; O'Connor, K.W.; et al. Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. *Cancer Discov.* **2014**, *4*, 1140–1153. [[CrossRef](#)]
25. Zhang, X.; Crawford, E.L.; Blomquist, T.M.; Khuder, S.A.; Yeo, J.; Levin, A.M.; Willey, J.C. Haplotype and diplotype analyses of variation in ERCC5 transcription cis-regulation in normal bronchial epithelial cells. *Physiol. Genom.* **2016**, *48*, 537–543. [[CrossRef](#)] [[PubMed](#)]
26. Marinescu, V.D.; Kohane, I.S.; Riva, A. The MAPPER database: A multi-genome catalog of putative transcription factor binding sites. *Nucleic Acids Res.* **2004**, *33*, D91–D97. [[CrossRef](#)]
27. Orans, J.; McSweeney, E.A.; Iyer, R.R.; Hast, M.A.; Hellinga, H.W.; Modrich, P.; Beese, L.S. Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell* **2011**, *145*, 212–223. [[CrossRef](#)] [[PubMed](#)]
28. González-Corrochano, R.; Ruiz, F.M.; Taylor, N.M.I.; Huecas, S.; Drakulic, S.; Spínola-Amilibia, M.; Fernández-Tornero, C. The crystal structure of human XPG, the xeroderma pigmentosum group G endonuclease, provides insight into nucleotide excision DNA repair. *Nucleic Acids Res.* **2020**, *48*, 9943–9958. [[CrossRef](#)]
29. Devos, J.M.; Tomanicek, S.J.; Jones, C.E.; Nossal, N.G.; Mueser, T.C. Crystal structure of bacteriophage T4 5' nuclease in complex with a branched DNA reveals how flap endonuclease-1 family nucleases bind their substrates. *J. Biol. Chem.* **2007**, *282*, 31713–31724. [[CrossRef](#)] [[PubMed](#)]
30. Maki, R.G.; Wathen, J.K.; Patel, S.R.; Priebat, D.A.; Okuno, S.H.; Samuels, B.; Fanucchi, M.; Harmon, D.C.; Schuetze, S.M.; Reinke, D.; et al. Randomized phase II study of gemcitabine and docetaxel compared with gemcitabine alone in patients with metastatic soft tissue sarcomas: Results of sarcoma alliance for research through collaboration study 002. *J. Clin. Oncol.* **2007**, *25*, 2755–2763. [[CrossRef](#)] [[PubMed](#)]
31. Drilon, A.; Laetsch, T.W.; Kummar, S.; Dubois, S.G.; Lassen, U.N.; Demetri, G.D.; Nathenson, M.; Doebele, R.C.; Farago, A.F.; Pappo, A.S.; et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N. Engl. J. Med.* **2018**, *378*, 731–739. [[CrossRef](#)]
32. van der Graaf, W.T.; Blay, J.-Y.; Chawla, S.P.; Kim, D.-W.; Bui-Nguyen, B.; Casali, P.G.; Schöffski, P.; Aglietta, M.; Staddon, A.P.; Beppu, Y.; et al. Pazopanib for metastatic soft-tissue sarcoma (PALETTE): A randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* **2012**, *379*, 1879–1886. [[CrossRef](#)]
33. Schöffski, P.; Chawla, S.; Maki, R.G.; Italiano, A.; Gelderblom, H.; Choy, E.; Grignani, G.; Camargo, V.; Bauer, S.; Rha, S.Y.; et al. Eribulin versus dacarbazine in previously treated patients with advanced liposarcoma or leiomyosarcoma: A randomised, open-label, multicentre, phase 3 trial. *Lancet* **2016**, *387*, 1629–1637. [[CrossRef](#)]
34. Frezza, A.M.; Jones, R.L.; Vullo, S.L.; Asano, N.; Lucibello, F.; Ben-Ami, E.; Ratan, R.; Teterycz, P.; Boye, K.; Brahmi, M.; et al. Anthracycline, Gemcitabine, and Pazopanib in Epithelioid Sarcoma: A Multi-institutional Case Series. *JAMA Oncol.* **2018**, *4*, e180219. [[CrossRef](#)] [[PubMed](#)]
35. Gentile, F.; Tuszyński, J.A.; Barakat, K.H. New design of nucleotide excision repair (NER) inhibitors for combination cancer therapy. *J. Mol. Graph. Model.* **2016**, *65*, 71–82. [[CrossRef](#)] [[PubMed](#)]
36. Bowden, N.A. Nucleotide excision repair: Why is it not used to predict response to platinum-based chemotherapy? *Cancer Lett.* **2014**, *346*, 163–171. [[CrossRef](#)]
37. Duan, M.; Ulibarri, J.; Liu, K.J.; Mao, P. Role of Nucleotide Excision Repair in Cisplatin Resistance. *Int. J. Mol. Sci.* **2020**, *21*, 9248. [[CrossRef](#)]
38. Tse, D.; Zhai, R.; Zhou, W.; Heist, R.S.; Asomaning, K.; Su, L.; Lynch, T.J.; Wain, J.C.; Christiani, D.C.; Liu, G. Polymorphisms of the NER pathway genes, ERCC1 and XPD are associated with esophageal adenocarcinoma risk. *Cancer Causes Control* **2008**, *19*, 1077–1083. [[CrossRef](#)] [[PubMed](#)]
39. Liu, J.; Deng, N.; Xu, Q.; Sun, L.; Tu, H.; Wang, Z.; Xing, C.; Yuan, Y. Polymorphisms of multiple genes involved in NER pathway predict prognosis of gastric cancer. *Oncotarget* **2016**, *7*, 48130–48142. [[CrossRef](#)] [[PubMed](#)]

40. Zhu, J.; Fu, W.; Jia, W.; Xia, H.M.; Liu, G.-C.; He, J. Association between NER Pathway Gene Polymorphisms and Wilms Tumor Risk. *Mol. Ther. Nucleic Acids* **2018**, *12*, 854–860. [[CrossRef](#)] [[PubMed](#)]
41. Italiano, A.; Laurand, A.; Laroche, A.; Casali, P.G.; Sanfilippo, R.; Le Cesne, A.; Judson, I.; Blay, J.-Y.; Ray-Coquard, I.; Bui, B.; et al. ERCC5/XPG, ERCC1, and BRCA1 gene status and clinical benefit of trabectedin in patients with soft tissue sarcoma. *Cancer* **2011**, *117*, 3445–3456. [[CrossRef](#)] [[PubMed](#)]
42. Börcsök, J.; Sztupinszki, Z.; Bekele, R.; Gao, S.P.; Diossy, M.; Samant, A.S.; Dillon, K.M.; Tisza, V.; Spisák, S.; Ruzs, O.; et al. Identification of a Synthetic Lethal Relationship between Nucleotide Excision Repair Deficiency and Irofulven Sensitivity in Urothelial Cancer. *Clin. Cancer Res.* **2020**, *27*, 2011–2022. [[CrossRef](#)] [[PubMed](#)]
43. Le Morvan, V.; Longy, M.; Bonaïti-Pellié, C.; Bui, B.; Houédé, N.; Coindre, J.M.; Robert, J.; Pourquier, P. Genetic polymorphisms of the XPG and XPD nucleotide excision repair genes in sarcoma patients. *Int. J. Cancer* **2006**, *119*, 1732–1735. [[CrossRef](#)] [[PubMed](#)]
44. Lu, X.; Liu, Y.; Yu, T.; Xiao, S.; Bao, X.; Pan, L.; Zhu, G.; Cai, Y.; Liu, Q.; Jin, C.; et al. ERCC1 and ERCC2 haplotype modulates induced BPDE-DNA adducts in primary cultured lymphocytes. *PLoS ONE* **2013**, *8*, e60006. [[CrossRef](#)] [[PubMed](#)]
45. Khrunin, A.V.; Firsov, S.Y.; Limborska, S.A. Polymorphisms of DNA repair genes ERCC2 and XRCC1 in populations of Russia. *Russ. J. Genet.* **2011**, *47*, 1391–1393. [[CrossRef](#)]
46. Liu, J.; Li, H.; Sun, L.; Feng, X.; Wang, Z.; Yuan, Y.; Xing, C. The Differential Expression of Core Genes in Nucleotide Excision Repair Pathway Indicates Colorectal Carcinogenesis and Prognosis. *BioMed Res. Int.* **2018**, *2018*, 9651320. [[CrossRef](#)]
47. Bajpai, D.; Banerjee, A.; Pathak, S.; Jain, S.K.; Singh, N. Decreased expression of DNA repair genes (XRCC1, ERCC1, ERCC2, and ERCC4) in squamous intraepithelial lesion and invasive squamous cell carcinoma of the cervix. *Mol. Cell. Biochem.* **2013**, *377*, 45–53. [[CrossRef](#)]
48. Huang, M.-Y.; Lee, H.-H.; Huang, C.-W.; Huang, C.-M.; Ma, C.-J.; Yin, T.-C.; Tsai, H.-L.; Chai, C.-Y.; Chen, Y.-T.; Wang, J.-Y. ERCC overexpression associated with a poor response of cT4b colorectal cancer with FOLFOX-based neoadjuvant concurrent chemoradiation. *Oncol. Lett.* **2020**, *20*, 212. [[CrossRef](#)] [[PubMed](#)]
49. Kirkilevsky, S.I.; Krakhmalev, P.S.; Malyshok, N.V.; Zadvornyi, T.V.; Borikun, T.V.; Yalovenko, T.M. Prognostic significance of microRNA-200b and ERCC1 expression in tumor cells of patients with esophageal cancer. *Exp. Oncol.* **2020**, *42*, 167–171. [[CrossRef](#)]
50. Moura, D.S.; Sanchez-Bustos, P.; Fernandez-Serra, A.; Lopez-Alvarez, M.; Mondaza-Hernandez, J.L.; Blanco-Alcaina, E.; Gavilan-Naranjo, A.; Martinez-Delgado, P.; Lacerenza, S.; Santos-Fernandez, P.; et al. *CUL4A*, *ERCC5*, and *ERCC1* as Predictive Factors for Trabectedin Efficacy in Advanced Soft Tissue Sarcomas (STS): A Spanish Group for Sarcoma Research (GEIS) Study. *Cancers* **2020**, *12*, 1128. [[CrossRef](#)]
51. Cancer Genome Atlas Research Network. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell* **2017**, *171*, 950–965e28. [[CrossRef](#)] [[PubMed](#)]
52. Peissert, S.; Sauer, F.; Grabarczyk, D.B.; Braun, C.; Sander, G.; Poterszman, A.; Egly, J.-M.; Kuper, J.; Kisker, C. In TFIIH the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat. Commun.* **2020**, *11*, 1667. [[CrossRef](#)] [[PubMed](#)]
53. Rump, A.; Benet-Pages, A.; Schubert, S.; Kuhlmann, J.D.; Janavičius, R.; Machackova, E.; Foretova, L.; Kleibl, Z.; Lhota, F.; Zemánková, P.; et al. Identification and Functional Testing of ERCC2 Mutations in a Multi-national Cohort of Patients with Familial Breast- and Ovarian Cancer. *PLoS Genet.* **2016**, *12*, e1006248. [[CrossRef](#)] [[PubMed](#)]