SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

3465.

BOZSÓDI ÁRPÁD

A támasztó és mozgató szervrendszer működésének fiziológiája

című program

Programvezető: Dr. Szőke György, egyetemi tanár

Témavezető: Dr. Lazáry Áron, PhD

The Role of Genetics in the Musculoskeletal System: Muscle Performance and Tumorigenesis Influenced by Altered Gene Expression and Gene Polymorphisms

PhD thesis

Árpád Bozsódi

Semmelweis University Doctoral School

Surgical Medicine Division

Supervisor: Áron Lazáry MD Ph.D

Official reviewers: Viktor Szabó MD Ph.D
János Kósa MD Ph.D

Head of the Complex Examination Committee:
Prof.Péter András Lakatos MD Ph.D

Members of the Complex Examination Committee:
Krisztián Szalay MD PhD
Miklós Tunyogi-Csapó MD PhD

Budapest
2026

Table of Contents

List of Abbreviations

Abbreviation	Full Term
ACE	Angiotensin-converting enzyme
BMD	Bone mineral density
CH	Chordoma
DNA	Deoxyribonucleic acid
HGS	Hand grip strength
OA	Osteoarthritis
OPG	Osteoprotegerin genes
PDGF	Platelet-derived growth factor
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SNP	Single nucleotide polymorphism
VDR	Vitamin D receptor

1. Introduction

During the short history of genetics, the most important factor was understanding the nature of the hereditary material. The gradual development of our knowledge in the field has allowed us to better understand the properties of this material. The integration of genetics into clinical research in recent years has resulted in an explosive increase in our knowledge, due to which new questions to be answered have arisen. Knowing the properties of the genetic material and mapping the errors has led to the understanding of the development of certain diseases. Correcting and preventing these errors can help prevent and treat certain diseases.

1.1. Deoxyribonucleic Acid

Deoxyribonucleic acid (DNA) consists of two nucleotide chains forming a double helix. Both consist of a deoxyribose sugar-phosphate (phosphodiester bonds) basic unit, and the bases are attached to the opposite chain by complementary bases[1]. Eukaryotic cells contain several types of DNA, mostly found in the nucleus. DNA has three cellular functions: replication (the two DNA strands separate and each serves as a template to build a new complementary strand), hereditary information (each base pair and nucleotide sequence is needed to build and maintain the organism), and cell division (through mRNA expression)[2].

1.2. Ribonucleic Acid

Ribonucleic acid (RNA) is a giant polymeric molecule similar to DNA, which is made up of repeating units. Its units are ribonucleotides[3]. The number of ribonucleotides within an RNA molecule can range from 75 to several thousand. Each ribonucleotide consists of a ribose sugar molecule, a nitrogen-containing organic base, and a phosphate group [4]. The individual units are connected to each other by the phosphate group, a so-called phosphodiester bond. Organic bases in RNA can be adenine (A), cytosine (C), guanine (G) and uracil (U) (thymine is found instead of uracil in DNA)[3, 4].

All organisms use RNA molecules to synthesize proteins. The genetic material of some simple organisms (such as viruses) is RNA. Some RNA molecules have catalytic properties, so they also fulfill an enzyme function (ribozyme enzymes)[5].

1.3. Nucleotides

Nucleotides are the structural units of RNA and DNA. The human genome sequence contains 2.85 billion nucleotides, which are interrupted by only 341 gaps. The total sequence covers 99% of the euchromatic genome and is accurate to one error per 100,000 bases[6].

Human chromosomes, composed of DNA and RNA, are found in the nucleus of every cell: 44 autosomes, which determine somatic characteristics, and two allosomes, which are responsible for sexual characteristics[2].

The Human Genome Project, started in 1990, outlined a complete genome sequence in 2003 and described that there are only 20,000-25,000 protein-coding genes [1].

1.4. Non-Coding RNA

Non-coding RNAs are transcripts that do not code information regarding proteins, but play a role in the coding of many other biological factors at the cellular and tissue level. Dysregulation of non-coding RNAs can contribute to the development of certain diseases[3].

The sequencing of the entire human genome supported the hypothesis that only a small part of the genome is responsible for coding proteins, which are ribosomal and transfer RNAs. Comparative studies of different organisms have shown that the transcriptional activity size of non-coding RNAs correlates better with evolutionary complexity than protein-coding regions[7].

Initially, the majority of the entire human genome sequence had no known assigned function. However, it was later proven that these regions can also be transcribed. Practically 98% of the human genome appears as RNA transcripts in certain cell types[5].

Non-coding RNAs can be divided into three large groups:

I. RNAs that have independent transcriptional and posttranscriptional regulation mechanisms and specific post-transcriptional regulation. The groups of these RNAs are microRNAs (miRNAs), RNAs associated with PIWI proteins (piRNAs) and small nuclear RNAs (snRNAs, snoRNAs), long non-coding RNAs (lncRNAs), as well as ribosomal and transfer RNAs[3, 5].

II. RNAs, which were created together with the transcriptional and post-transcriptional processing of actively transcribed genes. These include different regulatory regions (promoter region), short RNAs that are produced at the 5' or 3' end region of different genes, as well as linear and circular RNAs that are produced during the cleavage of different transcripts and pseudogenes[7].

III. RNAs that are produced during the transcription of intergenic regions. These include repetitive sequences such as LINE, SINE, LTR sequences, which make up a large percentage of the human genome in the centromere and telomere sections[4, 7].

1.5. The Role of Genetics in the Function of the Musculoskeletal System

Cell biology and genetics are rapidly developing basic sciences. Their significance is increasing in understanding the background of musculoskeletal diseases. Many research works focus on the treatment of musculoskeletal diseases, while the role of the study of genomics and its development in personalized medicine is becoming increasingly prominent.[8].

At the beginning of this millennium, the concept of orthogenomics was born and entered the public consciousness. The importance of genomics in future orthopedic practice has been defined, but its implementation is a slow process [9]. Different strategies have been proposed to identify the genetic basis of diseases such as diseases with a significant genetic component (osteoarthritis), questioning the effectiveness of surgical or conservative treatments (disc degeneration), and diseases affecting large populations [10].

The genetic role in the etiology and progression of common orthopedic diseases is known. Several mutations are associated with osteoporosis, for example in the osteoprotegerin (OPG) genes, the vitamin D receptor genes (VDR), or the low-density lipoprotein receptor-related protein 5 (LRP5)[11]. These genes play a role in inhibiting osteoclast production and wingless-related integration site (Wnt) signaling, in reducing bone mineral density (BMD), thereby causing osteoporosis.[12]. OPG and Lrp5-related SNPs (single nucleotide polymorphisms) increase the risk of osteoporotic fractures, independent of reduced BMD [13]. The genetic background plays a decisive role in the appearance of osteoarthritis (OA) in 65% of the knee, 60% in the hip, and 39% in the case of the hand. Association studies identified two loci: growth differentiation factor-5

(GDF5), which is associated with bone and cartilage development, and a component of the oligomeric Golgi complex-5 (COG5)[14]. There are 56 possible SNPs in the 50 genes and gene loci that can be associated with OA or 25 OA subtypes. These genes influence Wnt-associated bone mass, bone changes due to compression, cartilage turnover, chondrogenic processes mediated by TGF-β1 (transforming growth factor β1)[15]. However, the power of these loci is low and several factors are needed to diagnose clinical OA. Some OA SNPs have been shown to be risk factors such as calmodulin-1 (CALM1) and asporin (ASPN) SNPs, or frizzled-related protein-2 (FRZB2) and collagen type II alpha-1 (COL2A1)[10].

A study described the use of SNP analysis in sports injuries and discussed interactions between polymorphic collagen genes and Achilles tendinitis or Ehlers-Danlos syndrome [9]. Most of the existing studies are not published in orthopedic journals. For example, information on bone-related cancers focuses on pathological identification and chemotherapy rather than surgical treatment.[16].

Pediatric osteosarcomas and Ewing sarcomas express platelet-derived growth factor (PDGF) ligand and receptor and/or KIT (tyrosine kinase) kinase [17]. However, drugs designed to target PDGF or KIT kinase (e.g., imatinib mesylate) have only been shown to be effective against gastrointestinal stromal tumors and chronic myeloid lymphomas. [15]. Genotype determines the risk of osteosarcoma, Paget's disease, and chondrosarcoma, as well as the prognosis after diagnosis. The occurrence of several mutations (e.g., p53 gene) is characteristic only in advanced diseases. SNPs in osteosarcoma-related genes are involved in several biological processes[18].

1.6. Vitamin D Receptor

The vitamin D receptor (VDR) is a widely studied candidate gene for several musculoskeletal phenotypes. Research studies have shown that VDR polymorphisms are associated with bone density and risk of bone fracture [19-21]. Osteoporotic fractures are related to the frequency of falls, which are also related to VDR genotypes [22, 23]. Falls are indirectly influenced by the VDR, through the effect of its polymorphisms on muscle function [22]. For the above reasons, VDR candidate polymorphisms were also studied in connection with different muscle strength-related phenotypes. However, the results of these studies are significantly inconsistent and

hardly comparable due to multiple measures of muscle strength and stratified study populations.[22, 24-29].

1.7. Chordoma

The appearance and progression of a tumor is the result of a series of random genetic mutations that occur during the clinical and pathological phases of cell development. [30]. This genetic rearrangement is the first step in oncogenesis. Mutations can usually be associated with a triggering event, which can occur at any age. The most accepted theory as a triggering event is a defect in DNA repair mechanisms [31, 32]. Chordoma is a rare primary malignant bone tumor, localised to the spine, but the development of the tumor is still not clearly defined. The failure of classical and novel targeted chemotherapy agents against chordoma may be related to its complex and heterogeneous oncogenesis, which is characterized by a modest somatic mutation rate but high genome instability affecting only certain chromosomal arms or resulting in chromothripsis[32, 33]. Importantly, these changes can affect not only protein-coding genes, but also non-coding RNA genes, which regulate several molecular networks at the transcriptional or post-transcriptional level. MicroRNAs (miRNAs) and long non-coding RNAs (lncRNAs) are critical regulators of normal cell and tissue homeostasis [34], and since miRNAs have been shown to play an important role in tumorigenesis, they are widely studied as new targets for anticancer drug development. [35-37]. In recent years, increasing evidence supports the possible role of altered miRNA expression in the development and progression of chordoma. [38-50] However, the published studies are characterized by a wide variety of biological samples and research methods used. Recent developments in whole transcriptome profiling techniques, such as next-generation sequencing (NGS) and related bioinformatics tools, provide the opportunity to perform comparative, quantitative determination of miRNA expression patterns in different tumors in parallel with their genes as expression profiling [51, 52]. The technique can be used to identify complex RNA-based regulatory networks and molecular pathways that reveal potential new molecular targets.

In recent years, more and more emphasis has been placed on research into the genetic influence on the functioning of the moving and supporting organ system. With the development of molecular genetic tests and the spread of new-generation sequencing

methods, the opportunity to learn about lesser-known areas has increased. The genetic influence on the physiological functioning of the musculoskeletal system, as well as the exact genetic background of the development of cancer, still contain several unclear points. The purpose of this study was to investigate the genetics of the functioning of the musculoskeletal system and the genetic background of bone tumors (chordoma).

2. Objectives

2.1. Vitamin D receptor gene

In addition to its many other effects, the VDR affects muscle performance, but the exact mechanism of the genetic influence on muscle function is not known.

The aim of our study was to assess the possible effect of VDR gene polymorphisms on several parameters related to muscle performance. In our study, single nucleotide polymorphisms (SNPs) were genotyped. We studied:

- Whether VDR polymorphisms are associated with muscle performance
- The role of the selected single nucleotide polymorphisms in hand grip strength of the dominant and non-dominant hand

2.2. Chordoma

Although more and more information is emerging regarding the genetic background of the development of chordoma, the oncogenesis of the tumor has not yet been deciphered. Among the currently used treatment strategies, surgical solution remains the only therapy with acceptable long-term results. However, due to the tumor's increased tendency to relapse, serial surgical solutions significantly worsen the patients' quality of life [111], which necessitates the development of newer, more effective treatment options.

Whole-transcriptome RNA-sequencing is a new and powerful method for the complete mapping of gene expression changes and genetic mutations [112].

Our goal was

- Determination of a new protocol to isolate RNA from fresh frozen chordoma samples
- Make a whole-transcriptom RNA-sequening of chordoma samples
- Determination of altered miRNA expression profile of miRNA expression profile of sacral chordoma compared to nucleus pulposus cells and prediction their effect on chordoma miRNA-mRNA networks.
- Determination of possible molecular targets in oncogenesis of chordoma

3. Materials and methods

3.1. Vitamin D Receptor Polymorphisms are Associated with Hand Grip Strength

Vitamin D (cholecalciferol) is a prohormone with lipophilic properties, which can partly enter the body from dietary sources and partly can be synthesized in the skin under the influence of UV light. [21] Cholecalciferol undergoes hydroxylation in the liver and kidney to form 25-hydroxycholecalciferol (D2) and then 1,25-dihydroxycholecalciferol (D3), which is the active form of vitamin D [53].

Vitamin D exerts its effect by binding to the vitamin D receptor. The vitamin D receptor (VDR) is a transcription factor that regulates gene expression in a ligand-dependent manner. VDR belongs to the steroid hormone receptor superfamily. The VDR molecule contains several important functional and structural regions, such as a nuclear localization signal peptide, hormone ligand binding region, DNA binding sequence, dimerization region, and activation and transactivation regions [54]. After binding the ligand, VDR forms a heterodimer with the retinoid X receptor (RXR) and binds to target genes with a vitamin D receptor binding site, thereby exerting its multifaceted effects in many tissue and cell types of the body [54].

The VDR gene is located on the long arm of chromosome 12 (locus 12q13.1) and contains two promoters and two additional regulatory regions. The mapping of polymorphisms (variants) of the gene began in the late 1980s, using different restriction enzymes and Southern blot hybridization techniques. As a result of the tests, ApaI-[74], EcoRV-[75], BsmI-[75], TaqI-[76], FokI-[77], Tru9I-[78], and Cdx2-[79] polymorphisms were successively described [80]. In recent decades, the effect of polymorphisms has been investigated in connection with many different phenotypic characteristics and diseases, but VDR itself is a very polymorphic gene; more than 500 different gene variants have been described within the large gene. [55]

Hand grip strength (HGS) is a reliable marker of physical capability and motor function [19] and can be a predictive factor of disability [20], morbidity, and mortality in the elderly. [19, 56] HGS in youth is associated with general health [22], metabolic risk [23], and nutritional status. [24] HGS changes with age (an increase in childhood and a

decrease in the elderly can be noticed) and generally declines in cases of chronic disease, malnutrition, and disability. HGS can be measured using simple, digitalized, cheap, and reliable methods both in childhood and adulthood. [25] HGS is a typical multifactorial phenotype where genetic and environmental factors influence the phenotypic feature. The heritability of HGS is about 50-60%. [26, 27] This phenotype is also suitable for identifying the association between genetic variants and mid- and late-life physical functioning. [28] However, only a few genes and gene variants have been reported in association with HGS so far, and results have been explored mainly in adult and/or stratified populations. Different studies found an association between HGS and polymorphism of uncoupling protein 3 (UCP3) in elderly people. [29, 57] A significant effect of angiotensin-converting enzyme (ACE) gene variants on HGS was revealed by Chiu et al. in adolescent girls [58] and by Yoshihara et al. in elderly subjects. [59] The combination of genomic variants in ACE, ACTN3 (alpha-actinin-3), and PPARA (peroxisome proliferator-activated receptor alpha) genes showed an association with HGS in schoolchildren in the recent study by Ahmetov et al. [60]

The correlation between vitamin D and muscle function was already investigated in many studies. These results demonstrated that vitamin D, beyond its many other effects, can influence muscle performance and can have a significant role in maintaining balance. [61, 62] The vitamin D receptor (VDR) is a widely studied candidate gene for several musculoskeletal and extraskeletal phenotypes. Polymorphisms of VDR have been shown to be associated with bone mineral density and fracture risk. [63-65] The latter is related to the frequency of falls, which was also described in association with VDR genotypes. [66, 67] This can be explained by the influence of VDR genotypes on the function of different skeletal muscles. VDR candidate polymorphisms showed a significant association with quadriceps strength in different studies [66, 68-71]; however, others have not confirmed this relationship [72, 73]. Among these studies, some included HGS in the phenotypic measurements and reported its association with VDR genotypes, yielding contradictory results in different stratified populations. A positive association between HGS and VDR *BsmI* polymorphism was reported by Geusens et al. in non-obese elderly Belgian women [69], and similar associations with a trend were described in two other studies. [66, 71] A contrary result failed to find any association between *BsmI* and HGS in Swedish women. [72] The *ApaI* polymorphism

has been found to be associated with low HGS in elderly Chinese people in a recent study by Wu et al. [74], but no association was reported by Iki et al. in Japanese women. [73] The *TaqI* variant was found not to be related to HGS in the study by Iki et al. [73]; however, a positive association was demonstrated by Windelinckx et al. [71]

These incongruent reports about the effect of VDR polymorphisms on HGS can be explained by the relatively underpowered studies performed mainly on different subgroups of patients and the partial consideration of the confounding factors. To avoid these biasing factors, in this study, we measured HGS as the main study phenotype and determined the influence of VDR gene variants on that in a large sample of young schoolchildren.

3.2. Materials and Methods VDR

3.2.1. Subjects and Phenotypic Measurements

A total of 706 children were recruited from seven primary schools around the country (Table 1). The research protocol was designed and implemented with regard to the Helsinki Declaration on human subjects testing and the study was approved by the Scientific and Research Ethics Committee of the Medical Research Council (431/PI/2007). Pupils from the 2^{nd}-5^{th} grades were included in the study after their parents' informed consent. All the children were Caucasian. All subjects were healthy volunteers without any known disabling musculoskeletal or other chronic disease or functional limitation. Children with acute conditions influencing hand grip strength (e.g., hand trauma) were also excluded. Body height and weight were measured with calibrated tools. Maximum HGS in the dominant and non-dominant hands was determined using the recommended digital Jamar Plus+ hand dynamometer (Patterson Medical, Bolingbrook, IL, USA) [25, 75]. The handle was adjusted so that the child's proximal interphalangeal joints lay exactly on the top of the adjustable handle. The measurement was made in a sitting position with the arm close to the body. The arm was flexed to 90 degrees, and the forearm was in a neutral position. The measurements were taken three times for the dominant hand and non-dominant hand as well. One minute of resting time was applied between the measurements. The highest score of

each hand was used in further analyses. The reliability of the hand grip strength measurement was studied in a subsample of 105 pupils.

Table 1. Demographic data of the study cohort (N=706)

Variable	**Mean (SD)**
Age (y)	9.8 (1.2)
Gender (boy/girl %)	51% / 49%
Weight (kg)	36.0 (9.1)
Height (cm)	141.2 (9.1)
BMI (kg/m^2)	17.8 (3.0)
Dominant hand (right/left %)	91% / 9%
Hand grip strength dominant (N)	147.5 (42.1)
Hand grip strength non-dominant (N)	137.3 (38.3)

3.2.2. Selection of VDR Polymorphisms

For our research, we conducted a literature search using the PubMed database. In the database, using the keywords "vitamin D", "VDR gene", "polymorphism", "expression", "variants", "molecular", "pathway", "protein", "enzyme", "muscle", "exercise", we conducted the literature search using the keywords "muscle performance" and "strength". After that, because the VDR gene was previously studied a lot, so-called candidate polymorphisms were selected for multiplex genotyping.

3.2.3. Isolation of Genomic DNA

Genomic DNA was isolated from saliva samples collected from children. The Oragene kit (DNA Genotek Inc., Ontario, Canada) was used to collect the saliva samples, and the DNA was isolated according to the manufacturer's protocol. The first step was to collect and lyse the cells. The collected saliva samples were centrifuged and then treated with digestion buffer solution and protease enzyme to remove contaminating RNA and proteins. The samples were thoroughly mixed and incubated for 16 hours at 55°C. The next step was DNA precipitation. Isolation buffer was added to the samples, resulting in a whitish precipitate. Absolute ethanol was added to the samples, which cleared the solution. The collected material was removed from the sample in small parts by pipetting onto a filter, which was gradually centrifuged until the material was completely consumed. Washing followed, during which we washed the filters in several steps with two types of washing liquid. RNase mix was added for nuclease digestion, then the sample collected on the filters was incubated for 30 minutes. After this step, the filters were washed again with two types of washing liquid and continuously centrifuged. The last step is DNA elution. The filters were placed in clean collecting tubes, onto which the elution solution was measured. After centrifugation, the filtered liquid contained pure DNA.

3.2.4. Genotyping

Saliva samples were collected from 674 children using the Oragene OG-500 kit (DNA Genotek Inc., Ontario, Canada). Genomic DNA was purified from the samples following the manufacturer’s protocol. DNA quantity and quality were checked using PicoGreen reagent (Molecular Probes, Inc., Eugene, Oregon, USA). Six candidate SNPs (*A1012G, FokI, DdeI, BsmI, TaqI,* and rs10783215) across the VDR gene were selected for genotyping based on previous literature reports (Table 2.). Genotyping was performed by the Technology Centre, Institute for Molecular Medicine Finland (FIMM), University of Helsinki using Sequenom MassArray technology and the iPLEX Gold reagents (Sequenom Inc., San Diego, USA). Allele discrimination is based on primer extension with single mass–modified nucleotides followed by MALDI-TOF mass spectrometry. Genotyping quality was examined by a detailed QC procedure

consisting of success rate checks, duplicated samples, and positive and negative control samples. Genotyping was done blinded to the phenotypic data.

Table 2. Studied VDR SNPs and descriptive statistics of genotyping

Nº	rs number	Traditional name	Alleles	Region	Success rate (%)	MAF	HWp
1	rs4516035	*A1012G*	C/T	promoter	99.0	0.429	1.0
2	rs2228570	*FokI*	C/T	Exon 2	98.9	0.391	0.563
3	rs3782905	*Ddel*	C/G	Intron 2	98.3	0.295	0.402
4	rs1544410	*BsmI*	A/G	Intron 8	99.0	0.374	0.404
5	rs731236	*TaqI*	C/T	Exon 9	99.0	0.372	0.721
6	rs10783215	-	C/T	3' UTR	99.5	0.474	0.419

MAF: minimal allele frequency, HWp: *p*-value of Hardy-Weinberg equilibrium

The gene polymorphisms we selected were: rs4516035, rs1544410, rs731236, rs10783215, rs3782905, rs11568820, and rs2228570 (Table 2).

3.2.5. Measuring Hand Grip Strength

The grip strength of each child's hand was measured using a digital Jamar dynamometer, for both hands. During the examination, the children repeated the measurement three times in a sitting position, with their forearms flexed at 90 degrees. The average results were measured in N (Newtons) and recorded in a database.

3.2.6. Statistical Analysis

Associations between demographic variables and hand grip strength of dominant and non-dominant hands were analyzed using Pearson's correlation and backward linear regression models. In correlation analyses, $r>0.40$ was considered satisfactory ($r>0.80$

as excellent, 0.61-0.80 very good, 0.41-0.60 good, 0.21-0.40 fair, 0-0.20 poor). Allelic and genotype distribution, Hardy-Weinberg equilibrium, minimal allele frequency (MAF), as well as haplotype constructs were determined using Haploview 4.2 software (Haploblock presented in Figure 2.1) [76]. Associations between hand grip strength and genetic variants were analyzed using the "SNPassoc" R software package [77] and the THESIAS software [78]. Individual genotype-phenotype associations were studied using generalized linear models (GLM). Haplotype-phenotype associations were investigated applying log-likelihood ratio tests (LRT). The individual effect of the haplotypes was estimated and compared using LRT analysis. Rare haplotypes less frequent than 1% were excluded from the analyses. All genetic analyses were adjusted for age, sex, weight, and height. A p-value less than 0.05 was considered significant. Power of the study calculation was performed using Quanto 1.2.4 software. [79]

3.3. Potential Molecular Mechanism in Self-Renewal is Associated with miRNA Dysregulation in Sacral Chordoma: a Next-Generation RNA Sequencing Study

3.3.1. Characteristics of Chordoma

Chordoma is a primary malignant bone tumor the only one originating from the spine, with a frequency estimated between 1-4% of bone tumors [81]. Its incidence in the entire population is 0.08/100,000 inhabitants/year [82]. The most common localization is the sacrum (50%) and the skull base (clivus) (35%). In the remaining 15%, it can occur anywhere in the vertebral bodies. [82]. According to the gender distribution, chordoma located on the sacrum is twice as common in men, while there is no difference between the genders in the occurrence of chordoma at the skull base [82, 83]. The appearance of chordoma shows a wide variation by age of the patients; however, in terms of its development, its appearance can be attributed mainly to the fourth, sixth, and seventh decades [83].

Considering the place of its formation and the nature of the tumor, surgical resection is a challenge for clinicians. It is a typical low-grade malignancy characterized by slow, locally aggressive growth with variable distant metastatic potential but a very high local recurrence rate [84]. Wide margin en bloc resection gives the best oncological results [85, 86] but this is often extremely difficult or impossible to achieve due to large tumor

size, surrounding vital structures, and the high risk of severe neurologic damage. In order to improve the oncological and clinical outcome, to reduce the high rate of local recurrences after tumor resection, and to prolong survival, new, effective adjuvant therapeutic options are needed, since chordomas are extremely resistant to existing chemotherapy and radiotherapy options. [84]. The main obstacle to the development of new anti-chordoma adjuvant therapeutic agents is the limited knowledge of the molecular mechanism of chordoma tumorigenesis. One of the known starting points for tumor formation is the notochord. [87]. Chordoma usually manifests between the ages of 40 and 60 and is thought to arise from malignant transformation of notochordal cell remnants of the axial skeleton [88] During development, the embryonic notochord almost completely disappears, and the only notochord-derived differentiated cells in humans are the nucleus pulposus (NP) cells of the intervertebral discs. Neither the differentiation process of human NP cells nor the biogenesis of chordoma tumor cells have been fully characterized [88, 89]. During the differentiation of the embryonic notochord, after the transformation to nucleus pulposus, cell remnants may remain in some parts of the spine. Youssef et al. described that benign notochord cell tumors or chordomas can develop from notochord cell islands remaining in the bone, but in some cases no tumorous transformation occurs. [90] The exact mechanism of the tumor formation is not known. So far, no environmental risk factors have been described in the development of chordoma that would play a role in the development of the disease, or promote the growth of the tumor, or increase its tendency to relapse. Some research has shown that radiotherapy applied after surgery can contribute to the transformation of the tumor into a dedifferentiated form [91]. It is rare, but family accumulation also occurs [92, 93]; however, the molecular background of the causes of these differences could not be explained.

Molecular biological studies investigating the appearance of chordoma are placing increasing emphasis on the possible role of the brachyury gene in the development of the tumor. [94]. Brachyury is a gene that plays a role in the development of the notochord, and it is most likely responsible for the accumulation of chordoma within the family.

Examining 20 cases of chordoma, Pillay et al. linked a single nucleotide polymorphism (SNP) located on exon 4 of the brachyury gene to the development of chordoma [33,

95]. Based on their results, chordoma patients who carry the “A” allele have increased brachyury expression. Since 18 of the 20 patients they examined had the "A" allele, they concluded that there is a significant correlation between carrying the "A" allele of the polymorphism and the development of chordoma [96]. The given polymorphism was also investigated by our working group within the framework of an international consortium research[97]. Out of 128 examined cases, the "GG" genotype was detectable in 8 tumors; in the other cases, the presence of the "A" allele was detectable in the form of "AA" or "AG" genotypes. Comparing our results with the patients' clinical data, we noticed that the survival was significantly shorter in cases with the "GG" genotype [97].

In recent years, the T transcription factor (brachyury) and its genetic polymorphism have been described as critical chordoma-specific oncogenic drivers. [97, 98]. A cancer vaccine targeting brachyury has recently been developed [99], but its clinical efficacy in chordoma patients is not yet known. Activation of the PI3K/AKT/TSC1/TSC2/mTOR signaling pathway [100], overexpression of receptor tyrosine kinases (e.g., EGFR), as well as activation of IGF1R have been described in chordoma and considered as targets for costumised therapy [101, 102]. Despite these findings, sporadic case reports or small case series have shown that even targeted chemotherapy produces partial tumor responses against these oncogenic pathways at best. [103-106].

3.3.2. Gene Expression

The gene expression differences described in chordoma research show similarities to tumors of cartilage origin, such as chondrosarcoma, chondroblastoma, and chondromyxoid fibroma [107]. Increased expression of the SOX9, fibronectin, MMP9, and MMP19 genes can be detected in chordoma; the function of these genes is responsible for the formation of hyaline cartilage and the extracellular matrix [107]. Gene expression changes that are expressed in chordomas, but not in other tumors of cartilage origin, have also been described. These include brachyury, CD24, cytokeratins 8, 15, 18, 19, DDR1 (discoidin domain receptor), and periplakin [108]. EGFR and EGF hyperexpression can also be found in the nucleus pulposus. Increased expression can also be observed in the case of genes such as collagen X, PDGFa, and reticulocalbin 3, which play an important role in the calcification of cartilage tissue and are responsible for the mitosis of the cells of connective tissue elements [108].

3.3.3. Protein Expression

Based on literature data, protein expression tests were mainly performed with immunohistochemical tests and reverse transcription polymerase chain reaction (RT-PCR). According to one of the most important observations, the expression of proteins responsible for tyrosine kinase activity (e.g., PDGFR-beta) was found in all chordomas [109, 110]. In a study analyzing 21 chordomas, EGFR expression was detectable in 67%, KIT in 33%, and C-MET in 66%. The active phosphorylated forms of the transduction molecules responsible for tyrosine kinase activity, such as ERK (86%), AKT (76%), and STAT3 (67%), were also expressed at a high rate [110]. Another study analysing 12 chordomas revealed that EGFR and C-met were expressed in 100% of the cases, while TrkA and NGF, which play a role in nerve cell development, are expressed at a significantly higher rate in chordomas than in cells of the notochord [110]. Cyclo-oxygenase-2 (COX2) showed increased expression in 90% of chordoma cases, ERβ (estrogen receptor β) in 100% of cases, and the androgen receptor also in 90% [110].

3.4. Materials and Methods chordoma

3.4.1. Patients and Tissue Samples

Eight chordoma (CH) and eight nucleus pulposus (NP) tissue samples were collected from surgically treated Caucasian patients. Tumor samples were obtained from three male and five female subjects (mean age: 61.1±12.5 years). All chordomas were localized to the sacrum and histologically classified as classic chordomas. Five cases were recurrent tumors, and none of the patients had received any adjuvant treatment before the surgery. The clinical characteristics of the patients are summarized in Table 3. The chordoma samples were frozen in liquid nitrogen immediately after removal and stored at -80 ºC until processing. NP tissues were surgically removed from eight patients (mean age: 60.3±10.5 years) diagnosed with degenerative disc disease (Pfirrmann grade III), undergoing lumbar intervertebral fusion surgery. All patients were informed and signed consent of approval for using their tissue samples for the study. The study was approved by the Scientific and Research Ethics Committee of the Medical Research Council (49777/2012/EKU (751/PI/12.)

Table 3. Chordoma cases selected in the study

Patient ID	Diagnosis	Age	Gender	primary/recurrent
CH4	classic chordoma	71	F	primary
CH6	classic chordoma	64	F	recurrent
CH7	classic chordoma	51	M	recurrent
CH8	classic chordoma	59	M	recurrent
CH10	classic chordoma	65	F	recurrent
CH11	classic chordoma	69	M	primary
CH12	classic chordoma	45	F	recurrent
CH13	classic chordoma	85	F	primary

3.4.2. Immunohistochemistry

Paraffin blocks were also prepared from the tumors, from which smears were prepared with hematoxylin eosinophilic staining. The positive chordoma diagnosis was confirmed using Vimentin (Linaris, Wertheim), S-100 protein (Dako, Glostrup), brachyury (Santa Cruz), KI67, and CKAE1 (Sigma-Aldrich) antibodies.

3.4.3. NP Tissue Cultures

Previously, we tested RNA isolation directly from surgically obtained NP tissues, but the quality and quantity of RNA were inadequate for next-generation sequencing, despite extensive optimization of the isolation protocol. Therefore, we established primary NP cell cultures from the NP tissue samples, applying the explantation method without any enzymatic digestions. NP cell cultures were maintained in Dulbecco's Modified Eagle's Medium/Ham's Nutrient Mixture F12 (Sigma-Aldrich) containing 20% fetal bovine serum (Sigma-Aldrich), and penicillin-streptomycin (100 U and 100 mg/mL, respectively). Cells were cultured in 25 cm^2 tissue culture flasks at 37ºC under a humidified atmosphere containing 5% (v/v) CO_2. The NP cells had a small polygonal shape with short processes in the monolayer cultures, which is the usual appearance of these cells in vitro [113].

3.4.4. RNA Isolation

RNA isolation from freshly frozen chordoma tissue was performed using a previously developed protocol due to its difficulties.

The tumor tissue removed during the operation was placed in a nitrogen container 1-2 minutes after the removal.

After crushing, the tissue piece frozen in liquid nitrogen is homogenized in an appropriate amount of ice-cooled acid-guanidine-phenol lysis buffer in a 50 ml centrifuge tube. 500 µl of Tri Reagent (Direct Zol RNA prep) is added to 50 mg of tissue.

After sonication, it was centrifuged for 15 minutes at 11,000 rpm. The supernatant is separated and placed in a 1 ml Eppendorf tube. Chloroform is added to the samples at a ratio of 200 µl/1 ml Tri reagent.

In the next step, we prepare 15 ml Phase Lock Gel tubes. These are first centrifuged at 1500 g for 1 minute. 1-6 ml of aqueous phase solution (PLG 15 ml) is added to the prepared column (the same amount as the amount of the sample), then the prepared sample is pipetted onto it. Centrifuge the columns for 5 minutes at 1500 g. PLG forms a barrier between the aqueous and organic phases. We carefully pipetted the upper aqueous phase containing nucleic acid into a new microcentrifuge tube.

RNA purification followed using the Direct-zol MiniPrep kit according to the manual protocol:

One volume of ethanol was added to one volume of sample homogenate taken in Tri Reagent. The mixture was pipetted onto a Zymo-Spin IIC column in a Collection Tube. This was centrifuged for one minute at 12,000 g. The step was repeated every 700 µl until the sample was used up.

400 µl RNA washing buffer (ethanol) was added to the Zymo Spin IIC column in the Collection Tube and centrifuged again for 1 minute.

We prepared the DNase mixture for DNase digestion: DNaseI (250 U/vial, lyophilised) 5 µl + 10 x DNase I Reaction buffer 8 µl + DNase-free water 3 µl + RNA wash buffer 64 µl.

The 80 µl DNase mix was pipetted onto the matrix of the prepared Zymo Spin IIC column, which was first centrifuged. Digestion followed at 37 °C for 15 minutes, then centrifuged for 30 seconds at 12,000 g.

400 µl of Direct-zol RNA pre-wash solution was pipetted onto the column, which was centrifuged for 1 minute at12000 g. This was followed by a washing phase using 700 µl RNA washing buffer, which was centrifuged at 12,000 g for 2 minutes.

25 µl DNase-free water was measured onto the column, which was centrifuged at maximum speed for one minute. The RNA eluted in this way could be used immediately or stored at < -70 °C.

3.4.5. Quantification and Quality Control of RNA

The quick-frozen chordoma tissue samples were frozen and pulverized, and the tissue powder was immediately lysed with TRIzolate Reagent (Sigma-Aldrich) (50 mg sample per 500 ul reagent), then the samples were homogenized with a rotor-stator homogenizer. An equal volume of chloroform was added to the homogenates and vortexed, and the aqueous phase was separated using Phase Lock Gel™ (QuantaBio) (PLG 15 ml Heavy) tubes. RNA was isolated from the aqueous phase using the Direct-Zol RNA Miniprep kit (Zymo Research) according to the manufacturer's instructions without DNase I treatment. RNA was prepared from NP cell cultures using TRIzolate Reagent and Direct-Zol RNA Miniprep kit, according to the manufacturer's instructions. RNA quality was checked with a Nanodrop 1000 spectrophotometer and an Agilent 2100 Bioanalyzer.

3.4.6. Library Preparation and Sequencing

For RNA-seq, sequencing libraries were prepared with the Illumina TruSeq RNA SamplePrep v2 kit, and for small RNA-seq, sequencing libraries were prepared with the Illumina TruSeq Small RNA SamplePrep kit, according to the manufacturer’s instructions. All libraries were indexed, allowing multiplexed sequencing. The RNA-seq libraries were sequenced as 2x100 bp paired-end reads, and the small RNA-seq libraries were sequenced as 1x50 bp single-end reads on Illumina HiSeq 2000. Raw data were preprocessed with the CASAVA 1.8.2 software (converts *.bcl files into *.fastq.gz files = compressed FASTQ files) and resulted in the demultiplexed, fastq.gz files used in all subsequent analyses.

3.4.7. Bioinformatic Analysis of RNA Next-Generation Sequencing Data

For most bioinformatic analyses, either the public Galaxy Main server or the Galaxy Mississippi server was used (https://usegalaxy.org/, https://mississippi.snv.jussieu.fr/). The sequencing quality of the raw data was first checked with FASTQC. Data were preprocessed using Trimmomatic (RNA-seq) and Adapter clipping + Trimmomatic with a size selection of 19-25 nt (small RNA-seq). RNA-seq for one sample (CH4) was unsuccessful, likely due to an index sequence mix-up or mislabeling during library preparation, resulting in a missing file after demultiplexing, but the small RNA-seq was successful for this sample as well. Based on the FASTQC-based quality control analysis, all small RNA-seq and RNA-seq reads were of high quality, and the libraries had the expected total read counts. Sequence length distribution plots of the small RNA-seq data revealed that the control (NP) sequencing libraries contained a higher ratio of longer RNA species (presumably snoRNA or tRNA fragments) and a population of shorter (14-18 bp) fragments. To ensure better normalization between chordoma and control sequencing libraries in the differential expression analysis, miRNA-seq data preprocessing included the selection of the 19-25 nt fragment population from all samples.

Preprocessed RNA-seq data were aligned to the hg19 reference genome using TopHat2, and the read count table was prepared with Featurecounts. Preprocessed small RNA-seq data were aligned to human miRNA hairpin sequences using Chimira[114], and the resulting plain counts table was used in the subsequent steps. Only miRNAs with a minimum average read count ≥ 5 in either the CH or the NP samples were considered for differential gene expression analysis. A value of 1 was added to every value in the Featurecounts tables to eliminate zero values. Differential gene expression analysis was done with EdgeR[115] using the read count tables, hierarchical clustering using the Euclidean distance measure and average linkage were performed with ClustVis[116], and a heatmap of relative expression levels was generated with GraphPad Prism. Principal component analysis was performed with DESeq2. Since clustering analysis of RNA-seq data suggested that the NP11 nucleus pulposus sample was an outlier, the sample was left out of the differential gene expression analysis (data not shown).

3.4.8. miRNA Target Prediction and Pathway Analysis

Differentially expressed miRNAs were selected from the EdgeR results table, using the following filters: FDR (adjusted p-value) ≤ 0.05, |log2FoldChange| ≥ 3. The top upregulated or downregulated miRNAs were those with |log2FoldChange| ≥ 5. Differentially expressed mRNAs were selected from the EdgeR results table, using the following filters: FDR (adjusted p-value) ≤ 0.01, |log2FoldChange| ≥ 2, and dispersion ≤ 1.54. Genes without functional annotation and partial cDNA genes were removed from the list. Targets of the differentially expressed miRNAs were predicted with the Targetscan 7.1 algorithm (http://www.targetscan.org/vert_71/)[117]. Gene ontology (GO) annotations for functional classification of target genes were based on the Panther Classification System (http://pantherdb.org/)[118] and literature searches. Pathway analysis was done with the Consensus PathDB human, release 32 (gene set overrepresentation analysis) (http://cpdb.molgen.mpg.de/CPDB)[119].

3.4.9. Real-Time Quantitative PCR Analysis and Data Processing

The expression of selected miRNAs and mRNAs was validated by real-time quantitative PCR (qPCR) method. TaqMan™ assay IDs and details of the RT-qPCR protocol are provided in the Supplementary Data. For miRNA data normalization, we tested RNU44, RNU6B, miR-106b, and miR-30d, and selected RNU44 as the best reference gene based on the analysis with the NormFinder software[120]. For mRNA data normalization, we tested GUSB, GAPDH, Cyclophilin A (PPIA), RPLP0, and HPRT1 as reference genes, and selected PPIA as the best reference gene. Samples were measured on StepOne™ Real-Time PCR System (Applied Biosystem®). Raw data analysis was performed using the StepOne Software v2.2.2 (Applied Biosystem®). Cutoff Ct values (limit of quantification) for each TaqMan® assay set at Ct=37. Ct values above this value, or "undetected" values were uniformly set to Ct=40. Normalized expression values were calculated using the 2^{-dCt} method[121]. Statistical analysis was done with GraphPad Prism, using Mann-Whitney U-tests on the normalized expression values.

3.4.10. Comparative Gene Expression Analysis

The initial gene expression analysis was performed using the Cuffdiff filtering algorithm. The DeSeq2 program was used for differentiation gene expression analysis,

which has been proven to have high specificity but lower sensitivity. In order to eliminate low sensitivity, we used an EdgeR program with good sensitivity as a supplement. In the continuation, only the gene expression changes detected by both programs were further investigated.

4. Results

4.1. Results VDR

4.1.1. Non-Genetic Determinants of Hand Grip Strength

Mean hand grip strength was 147.5±42.1 Newtons (N) and 137.3±38.3 N in dominant and non-dominant hands, respectively, in our cohort, and the correlation between the strength of the two hands was excellent ($r=0.83$, $p<0.01$). HGS in both dominant and non-dominant hands were distributed normally in the sample population ($p>0.05$ in the Kolmogorov-Smirnov test). Age, weight, height, and gender correlated with both hands' grip strength and all four demographic parameters proved to be significant predictors in multivariate regression models (data not shown). The measurement method proved to be reliable, characterized by a high ICC for both hands (0.98 and 0.95 for dominant and non-dominant hands, respectively) (Table 4.).

Table 4. Reproducibility of the hand grip strength measurements (N=105)

Measurement	Test mean (SD)	Retest mean (SD)	*p*-value (t-test)	ICC (95% CI)	SEM
Grip force dominant	230.0 (55.7)	228.0 (54.6)	0.141	0.98 (0.96-0.99)	7.8
Grip force non-dominant	216.1 (55.5)	215.4 (55.1)	0.725	0.95 (0.93-0.97)	12.4

ICC: Intraclass Correlation Coefficient, SEM: standard error of measurement

4.1.2. Genetic Associations

The genotyping success rate was more than 98.5% for all SNPs. DNA quality was not good enough for further genotyping in six cases. The genetic association tests were conducted only on subjects who were successfully genotyped for all polymorphisms (n=643). Minor allele frequencies of each SNP were more than 1% in the study population, and all genotyped polymorphisms were in Hardy-Weinberg equilibrium (Table 5.). A haploblock constructed by three SNPs in the 3' part of the gene (rs1544410-rs731236-rs10783215) was identified (Figure 1.).

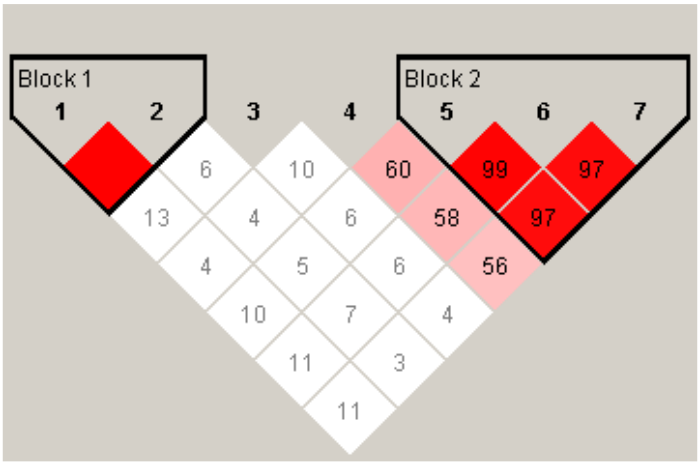

Figure 1. Linkage relationships of VDR SNPs (a larger number indicates a higher degree of linkage between the two polymorphisms)

Association analyses were adjusted for the non-genetic predictors such as age, gender, weight, and height. Significant associations were found between three individual SNPs (rs4516035, rs1544410, rs731236) and HGS (Table 5.A,B., Figure 2. A, B, C, D.). The 'T' allele of the rs4516035 (A1012G) polymorphism was associated with higher grip strength in both hands. Higher dominant hand grip strength was found in children carrying the 'A' allele of rs1544410 (*BsmI*) and in 'C' allele carriers of rs731236 (*TaqI*).

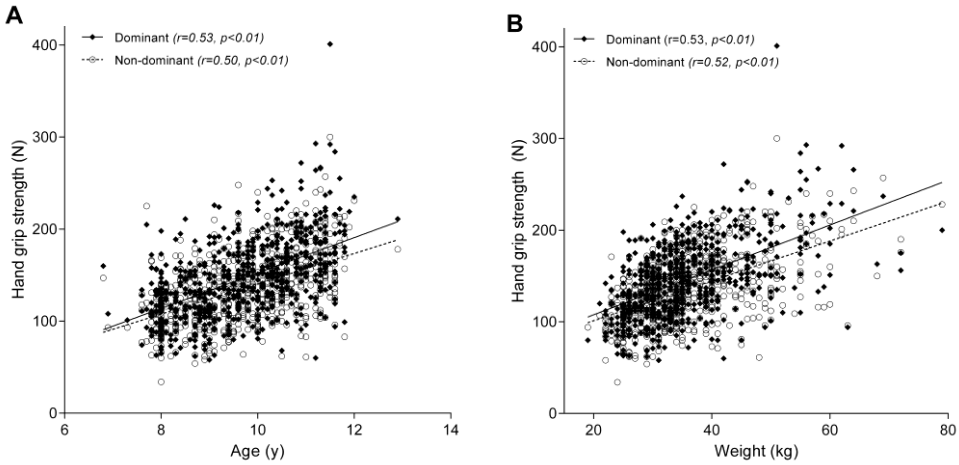

Figure 2.A, B. Association of demographic parameters with hand grip strength (A: age, B: weight,– mean, 95%CI, p < 0.05)

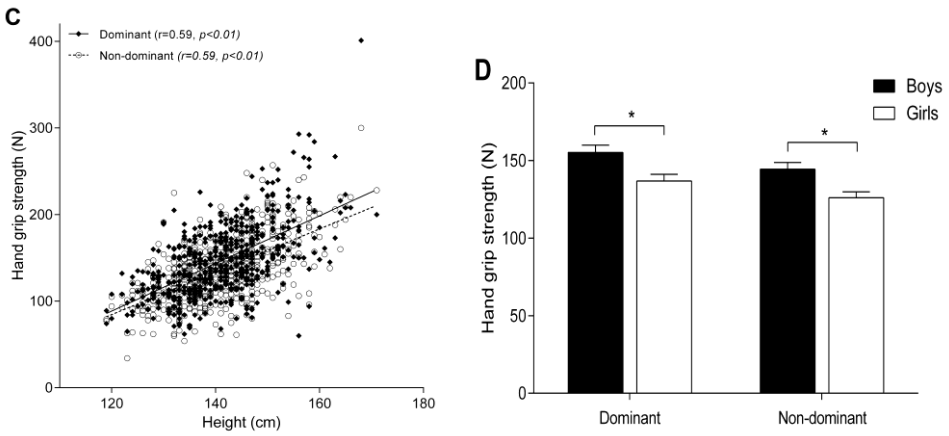

Figure 2.C, D. Association of demographic parameters with hand grip strength (C: height, D: gender – mean, 95%CI, p < 0.05)

Three haplotypes with more than 1% frequency were identified inside the VDR haploblock located in the 3' part of the gene. VDR haplotypes showed significant associations with hand grip strength in both hands ($p<0.005$ and $p<0.01$ for dominant and non-dominant hands, respectively) (Figure 3. Table 5.). The mean difference in hand grip strength between 'ACT' and 'GTT' haplotypes was 12.1 N and 10.0 N for the dominant and non-dominant hands, respectively.

Considering the mean of HGS in the dominant hand, our sample size provided 80% power to detect a 10N difference in a covariate-adjusted additive genetic model in case of 0.373 MAF.

Table 5..A, B. Hand grip strength relation to VDR genotypes (N=643)

A.

		Hand grip strength (N)							
		Dominant				*Non-dominant*			
SNP Genotypes	**Frequency (n, %)**	**Mean**	**SD**	**95% CI**	***p***	**Mean**	**SD**	**95% CI**	***p***
rs4516035									
CC	119 (18.5)	144.8	43.2	137.0-		135.6	40.7	128.2-143.0	
CT	324 (50.3)	145.1	42.6	152.7		134.0	36.9	130.0-138.0	
TT	200 (31.1)	148.9	40.8	140.5-	***0.009***	137.6	38.3	132.2-142.9	***0.034***
				149.8					
				143.2-					
				154.6					
rs2228570									
CC	241 (37.5)	146.2	43.8	140.7-		134.6	40.1	129.5-139.7	
CT	300 (46.6)	145.2	42.2	151.8		135.7	37.7	131.4-140.0	
TT	102 (15.8)	149.3	37.9	140.4-	0.534	136.6	34.1	129.9-143.3	0.520
				150.0					
				141.9-					
				156.8					
rs3782905									
CC	315 (48.9)	143.5	41.0	138.9-		132.0	36.7	127.9-136.1	
CG	279 (43.3)	148.0	40.3	148.0		138.5	37.5	134.1-142.9	
GG	49 (7.6)	154.0	56.4	143.3-	0.278	140.1	47.5	126.4-153.7	0.248
				152.8					
				137.8-					
				170.2					

B.

		Hand grip strength (N)							
		Dominant				*Non-dominant*			
SNP Genotypes	**Frequency (n, %)**	**Mean**	**SD**	**95% CI**	***p***	**Mean**	**SD**	**95% CI**	***p***
rs1544410									
AA	87 (13.5)	151.7	41.7	142.8-160.6	***0.010***	138.0	38.7	129.8-146.3	0.169
AG	310 (48.2)	148.2	44.7			136.5	38.7	132.2-140.9	
GG	246 (38.2)	141.9	38.6	143.2-153.1		133.1	36.9	128.5-137.7	
				137.1-146.8					
rs731236									
CC	87 (13.5)	151.9	41.9	143.0-160.9	***0.038***	138.5	39.3	130.1-146.8	0.116
CT	304 (47.2)	147.9	44.2			136.4	38.4	132.1-140.8	
TT	252 (39.1)	142.3	39.4	142.9-152.9		133.1	37.1	128.5-137.7	
				137.4-147.1					
rs10783215									
CC	175 (27.2)	143.5	39.8	137.6-149.5	0.372	135.2	37.7	129.6-140.9	0.654
CT	335 (52.1)	147.1	42.9			135.6	38.5	131.5-139.8	
TT	133 (20.6)	147.8	43.2	142.4-151.7		135.1	37.6	128.7-141.6	
				140.4-155.2					

Table 5. Association of VDR haploblock with hand grip strength (n=643)

		Hand grip strength (N)			
		*Dominant hand***		*Non-dominant**	
Haplotype	**Frequency (%)**	**Mean**	**95% CI**	**Mean**	**95% CI**
GTC	52.7	72.3	69.5-75.1	67.9	65.5-70.4
ACT	36.5	76.6	73.1-80.1	69.5	66.4-72.6
GTT	9.5	64.5	56.6-72.3	59.5	51.8-67.1

*p<0.01, **p<0.005

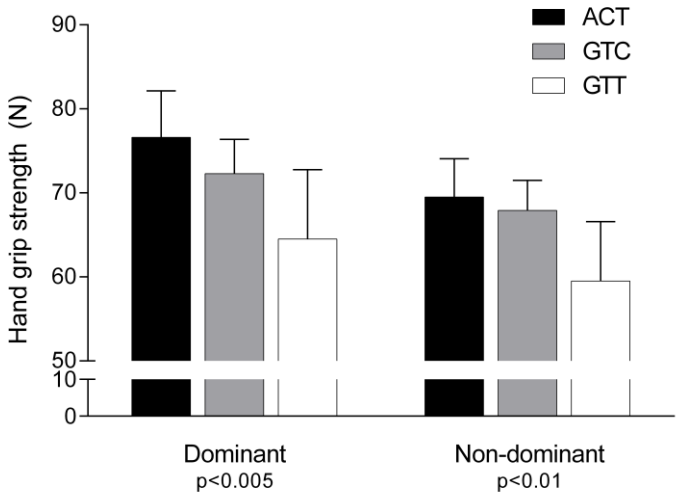

Figure 3. Estimated hand grip strength according to VDRhaplotypes (mean, 95%CI and p-values of the global haplotype effect on HGS are represented). (n=643)

4.2. Results chordoma

4.2.1. Immunohistochemistry Analysis of Chordoma

As a result of the immunohistochemical tests, all 8 tumor samples selected for sequencing confirmed classical chordoma. Intensive brachyury positivity was also confirmed in all examined samples (Figure 4).

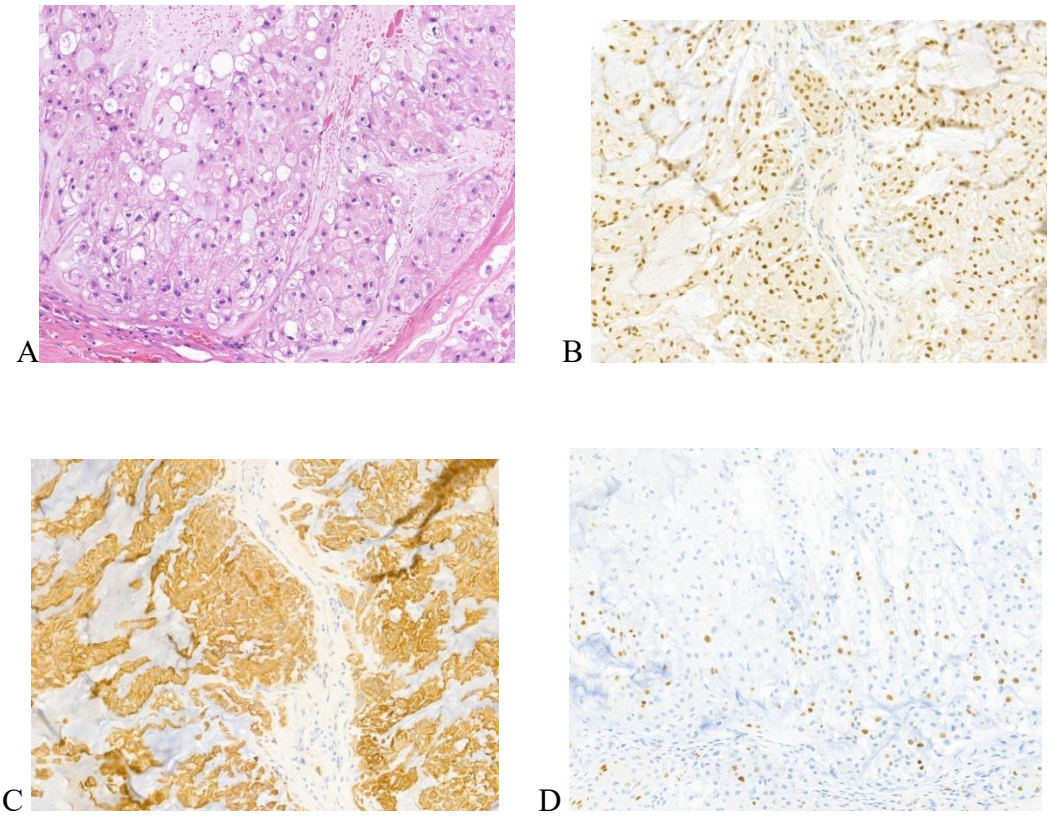

Figure 4. Immunohistochemistry analysis of chordoma (A) Hematoxylin and eosin, (B) brachyury, (C) CKAE, (D) Ki67

4.2.2. Results of RNA Isolation and Sequencing, Quality Control

After the isolation of the RNA, quality and quantity control was carried out, during which checks showed excellent results were measured in all samples (Table 6.).

Results of whole transcriptome sequencing of chordoma and nucleus pulposus using next generation equipment (Genome Analyzer II- Illumina) are shown in table 7.A, 7.B. Sequencing of a single chordoma sample was not successful. The determination of this sample sequence will be repeated in the future.

Table 6. Results of quantity control of RNA isolation of chordoma and nucleus pulposus (Nano Drop)

Sample ID	Sample origin	ng/uL	260/280	260/230	Agilent RIN
CH4	chordoma tumor. surgical specimen	223.19	2.08	2.32	8.2
CH6		468.01	1.95	1.95	8.2
CH7		462.8	1.94	1.27	8.3
CH8		3262.68	1.93	1.96	8.8
CH10		578.77	2.03	1.43	8.6
CH11		1357.92	2	1.82	7.3
CH12		938.61	2.03	1.78	8.0
CH13		1045.25	2.06	1.54	7.9
NP5	Nucleus pulposus cell culture	233.58	2.07	2	8.1
NP6		218.29	2.07	1.99	9.2
NP7		311.51	2.06	1.91	9.1
NP9		900.41	2.15	2.12	10
NP10		832.77	2.15	2.11	10
NP11		599.07	2.13	1.36	10
NP12		209.48	2.09	2.07	7.9
NP13		687.45	2.13	2.06	9.6

During the statistical analysis of paired ends in the TopHat2 series, it was confirmed that the sequence files were able to determine a mapping sequence with an overall accuracy of 77% on an average of 18,330,864 total reads. Furthermore, on average, 13,325,254 matches during end-to-end matching determined the accuracy of the sequence alignment with 72.7% consensus (Tables 7.A-B.). The number of accepted reads of the transcriptome in the tumor samples confirmed 24-26 fold coverage, while in the nucleus pulposus samples 20-27 fold coverage, the accuracy of which was considered the intended hit number even before the test. In the case of some control samples, the values confirming lower coverage (NP7, NP11, NP13) are acceptable, because the number of gene expression variations and chromosomal aberrations in the control samples is extremely small. The mapping rate of genetic information exceeded the required 60% in all samples. The sequence matching rate in all samples also exceeded 60%.

Table 7.A. Results of whole transcriptome sequencing of chordoma and nucleus pulposus

		Left reads		Right reads	
Sample ID	Tissue type	All accepted reads	Mapped reads (% of input)	All accepted reads	Mapped reads (% of input)
CH4	tumor	sequencing was unsuccessful			
CH6	tumor	18369649	14849045 (80.8%)	18369649	14021174 (76.3%)
CH7	tumor	18284510	14765809 (80.8%)	18284510	14008803 (76.6%
CH8	tumor	18238291	14136624 (77.5%)	18238291	13421589 (73.6%
CH10	tumor	18797429	15037662 (80.0%)	18797429	14274028 (75.9%)
CH11	tumor	19349383	14411471 (74.5%)	19349383	13664557 (70.6%
CH12	tumor	19855625	15353196 (77.3%)	19855625	14563183 (73.3%
CH13	tumor	18048073	13578229 (75.2%)	18048073	12884589 (71.4%
NP5	normal	18567461	14739372 (79.4%)	18567461	13925793 (75.0%
NP6	normal	19416597	15762699 (81.2%)	19416597	14891436 (76.7%
NP7	normal	16709134	13206963 (79.0%)	16709134	12500064 (74.8%
NP9	normal	19065870	15363582 (80.6%)	19065870	14395381 (75.5%
NP10	normal	20370354	16418370 (80.6%)	20370354	15411488 (75.7%
NP11	normal	16482810	12674606 (76.9%)	16482810	11880093 (72.1%
NP12	normal	18152445	14539912 (80.1%)	18152445	13628977 (75.1%
NP13	normal	15255331	12866219 (84.3%)	15255331	12034558 (78.9%

Table 7.B. Results of whole transcriptome sequencing of chordoma and nucleus pulposus

Sample ID	Tissue type	Overall read mapping rate	Aligned pairs	Concordant pair alignment rate
CH4	tumor	-	-	-
CH6	tumor	78.6%	13646400	74.2%
CH7	tumor	78.7%	13627713	74.4%
CH8	tumor	75.6%	13004960	71.2%
CH10	tumor	78%	13872992	73.7%
CH11	tumor	72.6%	13253182	68.4%
CH12	tumor	75.3%	14160732	71.2%
CH13	tumor	73.3%	12529207	69.4%
NP5	normal	77.2%	13560750	73%
NP6	normal	78.9%	14496195	74.6%
NP7	normal	76.9%	12170942	72.8%
NP9	normal	78%	13995591	73.4%
NP10	normal	78.1%	14975557	73.5%
NP11	normal	74.5%	11551344	70.1%
NP12	normal	77.6%	13280481	73.1%
NP13	normal	81.6%	11752767	77%

4.2.3. Expression Profile of miRNAs

Our analysis identified 126 differentially expressed miRNAs, with 74 miRNAs being upregulated and 52 downregulated in chordoma (Figure 4.A. and Supplementary Table 1). The PCA analysis based on the miRNA differential expression profile demonstrated a clear separation of CH samples from NP samples and also identified the CH4 chordoma sample as an outlier (Figure 4.B.). It should be noted that clinical follow-up showed that the CH4 patient is still in complete remission, as opposed to the other chordoma patients who all succumbed to the disease or its complications. 52% (N=66) of the dysregulated miRNAs were intragenic, and 39% of these (26 miRNAs) were co-regulated with their host genes. Interestingly, several downregulated ncRNAs clustered to chromosome 14 at the border of chromosomal bands 14q32.2 and 14q32.3, including miRNAs located between hsa-miR-493 and hsa-miR-409, the small nucleolar RNAs SNORD112-114, and the long intergenic noncoding RNAs MEG3/8/9.

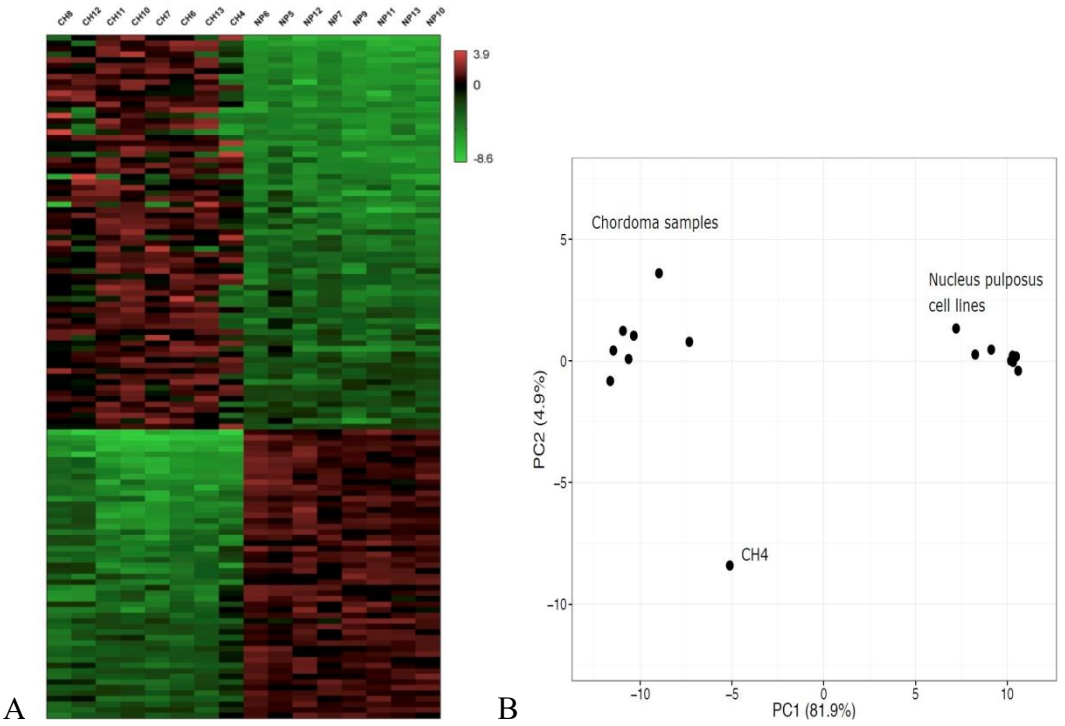

Figure 4.A. Differentially expressed miRNAs in chordoma tumors (CH) and nucleus pulposus (NP) cell lines. a) Heatmap of the 126 differentially expressed (DE) miRNAs in chordoma. Normalized expression values from Chimira/EdgeR analysis were scaled and log2 transformed. The color scale of the heatmap

illustrates the relative expression levels of the DE miRNAs 126 differentially expressed miRNAs, 4.B. CH4 chordoma sample as an outlier

4.2.4. miRNA Target Prediction and Pathway Analysis

Next, we selected the top upregulated and downregulated miRNAs for mRNA target prediction (Figure 5).

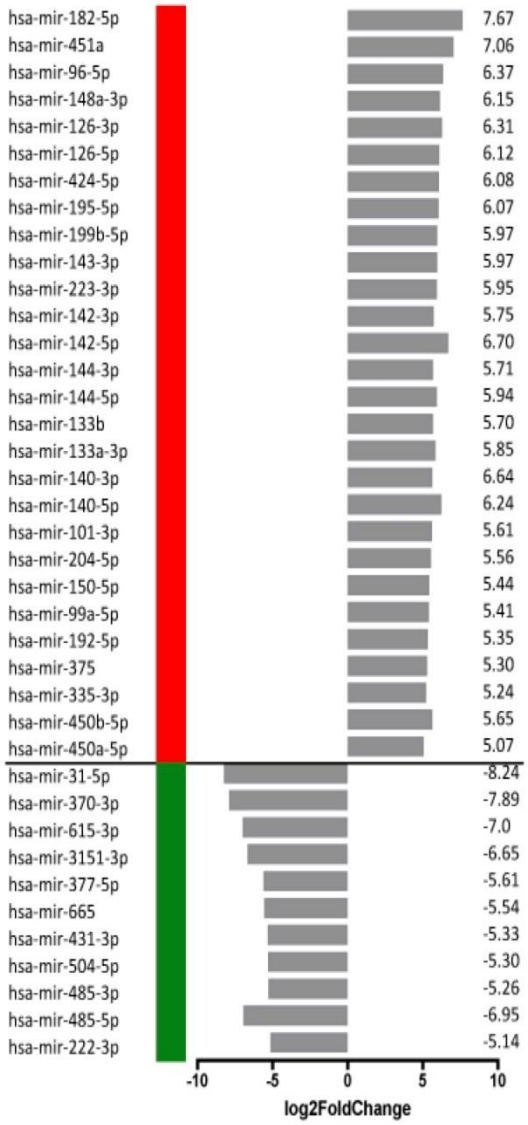

Figure 5. Top dysregulated miRNAs in chordoma. Relative expression (log2-FoldChange) and direction of the dysregulation are represented (red: upregulated, green: downregulated miRNAs).

Gray cells indicate Targetscan predictions of the miRNA-mRNA interaction. Anti-tumor or oncogenic classification was based on literature searches. Functional classification was based on PantherDB analysis and literature searches. Only the most prominent categories are shown – annotated genes outside these categories are lumped together in the "Mixed" category. Genes in the "mixed" category but without a connection with cancer are not shown. miRNAs with the highest number of predicted targets are highlighted in orange/pale orange. Targets of miRNAs in the same miRNA family are marked in the same column. Regulated miRNAs with only a few predicted targets are not shown.

Predicted target genes (based on Targetscan analysis) were filtered by their Cumulative weighted context++ scores (score≤-0.3) and overlapped with the list of significantly up- or downregulated genes from the RNA-seq analysis. Based on the literature, the majority of downregulated predicted target genes of upregulated miRNAs have anti-tumor activities – on the other hand, the upregulated genes targeted by downregulated miRNAs are mostly oncogenic. Functional classification of the target genes based on Panther DB analysis revealed that both up- and downregulated miRNAs target transcription factors and components of signal transduction pathways. In addition, upregulated miRNAs may target regulators of chromatin structure and 16 protein kinases as well, as opposed to downregulated miRNAs, which do not target genes with chromatin-associated functions, and may target just a single protein kinase, ITK (Figure 6).

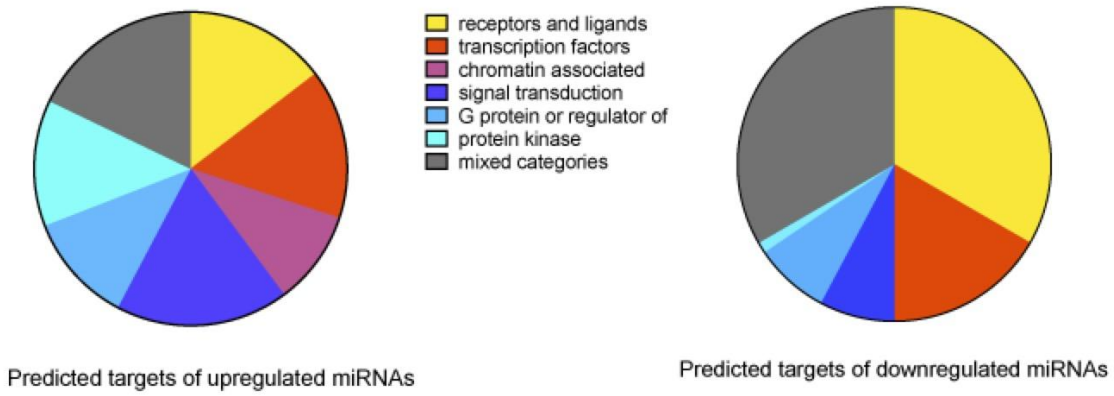

Figure 6. MiRNA Target Regulators

Overrepresentation analysis of predicted target genes identified different enriched pathways that may be regulated by the up- or downregulated miRNA sets (Figure 7).

Similarly to target gene functional classification, upregulated miRNAs are likely to target various signal transduction pathways, whereas downregulated miRNAs appear to target pathways related to immune response regulation. A notable target is the TNF signaling pathway, which was shown to activate metastatic and tumor-promoting inflammatory pathways in chordoma. We observed significant overexpression of TNF family members and several TNF receptors (TNFRSF1B/4/8/18) in the chordoma samples (data not shown). Interestingly, NFATC2 (nuclear factor of activated T-cells), a transcriptional regulator of TNF and TNF receptor expression in melanoma cell lines[122] is also overexpressed in the chordoma samples and may be targeted by the downregulated miRNA miR-665.

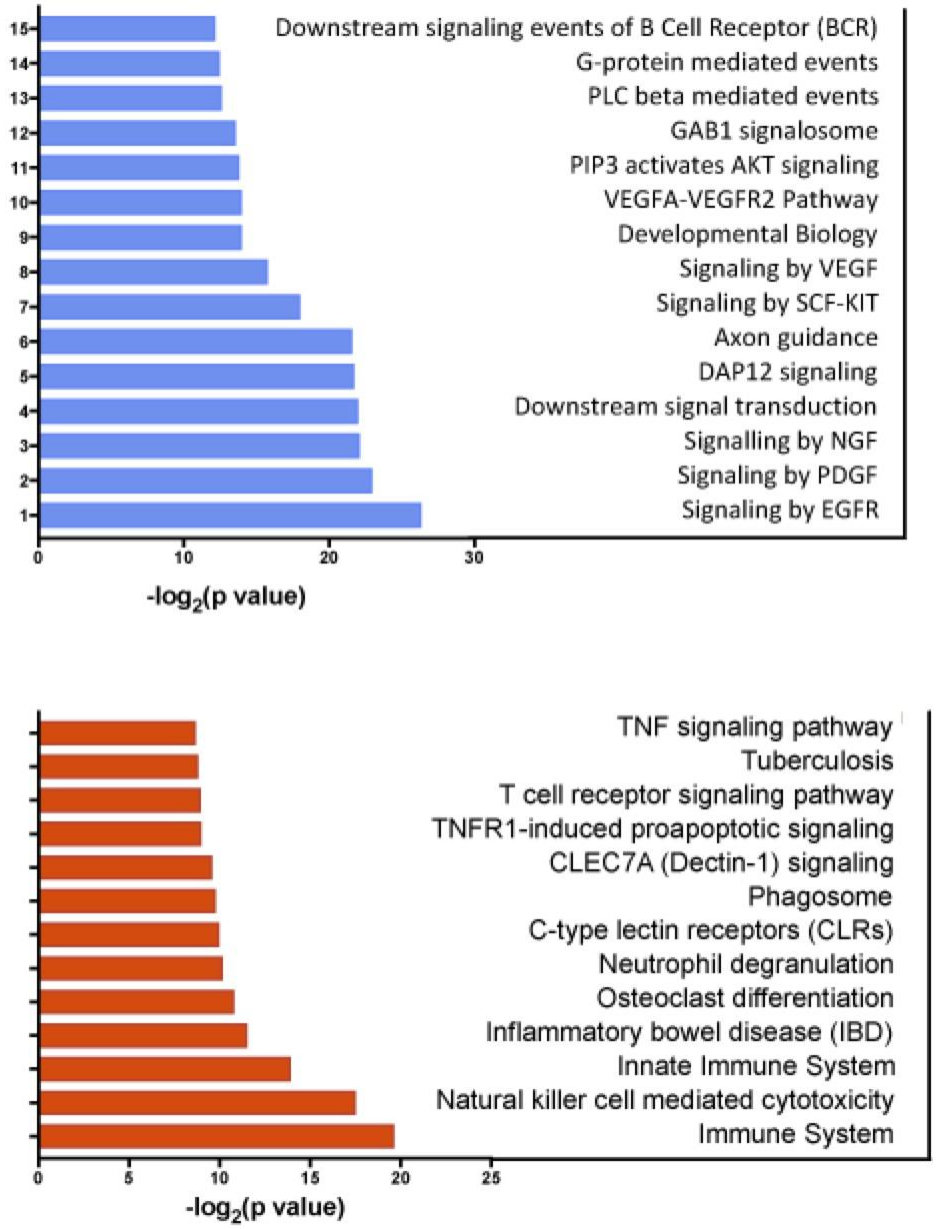

Figure 7. Results of functional classification analysis a) and pathway analysis(b) of the predicted targets of upregulated and downregulated miRNAs inchordoma. (a) The protein function GO categories identified by Panther DB functional analysis (GO-Slim Molecular function) are shown. (b) Overrepresented pathways were identified with the ConsensusPath DB algorithm.The x-axis shows the -log2(p value) of the enrichment. Top panel: Enriched pathways for downregulated genes predicted to be targeted by upregulated miRNAs. Bottom panel: Enriched pathways for upregulated genes predicted tobe targeted by downregulated miRNAs.

4.2.5. Identification of a Possible Self-Renewal Molecular Mechanisms in Chordoma

To identify the genes that are the most likely candidates for regulation by differentially expressed miRNAs in chordoma, we selected several dysregulated genes that may be targeted by multiple dysregulated miRNAs, and target genes with the best cumulative weighted context++ scores (score≤-0.6). This approach suggested several previously validated regulatory mRNA/miRNA interactions between LCOR/miR-199a-5p, MITF/miR-182-5p/miR-101/miR-148a-3p, HIF3a/miR-485-5p, HMGA1/miR-142-3p, RASA1/miR-182-5p/miR-223-3p, RGS17/miR-182-5p and RICTOR/miR-142-3p. Functional annotation of the mRNA targets predicted a miRNA-mRNA network regulating the self-renewal potential of chordoma tumors, through the regulation of the transcription factor LCOR, and the potential activation of the Hippo pathway through targeting its regulators, TAOK and MOB1 (Figure 8).

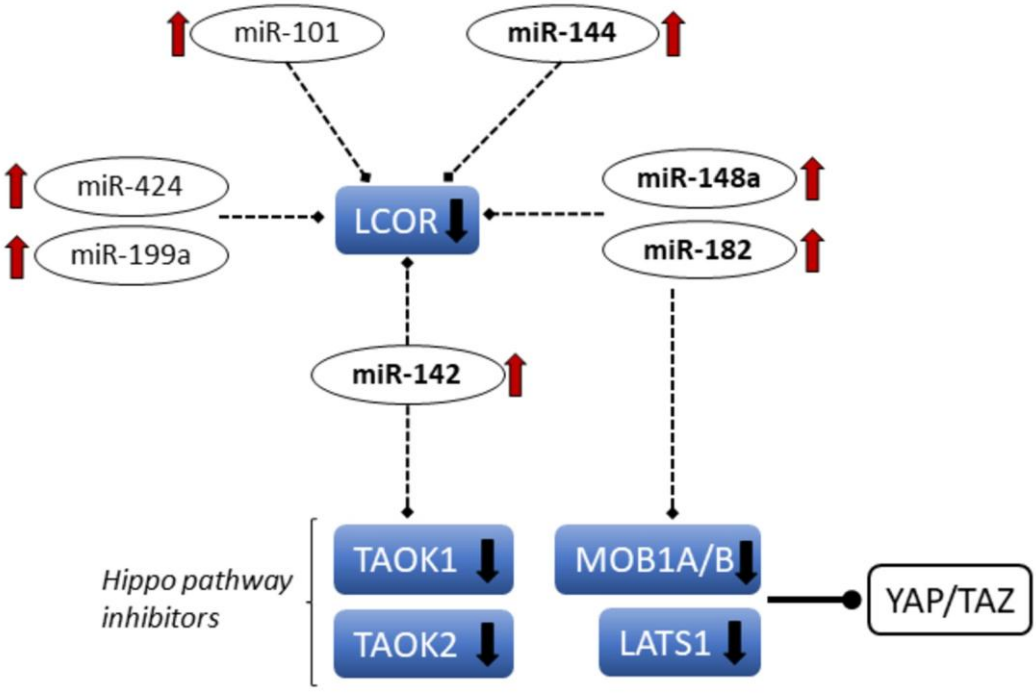

Figure 8. Predicted miRNA-mRNA regulatory network of self-renewal in chordoma. Bold arrows indicate the direction of RNA expression dysregulation: red ¼ upregulated, black ¼ downregulated. Upregulated miRNAs validated by RT-qPCR are in bold

3.4.6. qPCR Validation of the Dysregulated Self-Renewal Pathway

We selected 4 miRNAs and 6 mRNAs from the predicted self-renewal regulatory network for RT-qPCR validation. The qPCR analyses confirmed the significantly different expression and the reciprocal regulation of the selected miRNAs (Figure 9) and their target mRNAs (Figure 10) in chordoma vs. nucleus pulposus samples, similar to the results of the NGS analysis. We also validated the reciprocal regulation of NFATC2 and miR-665 (data not shown).

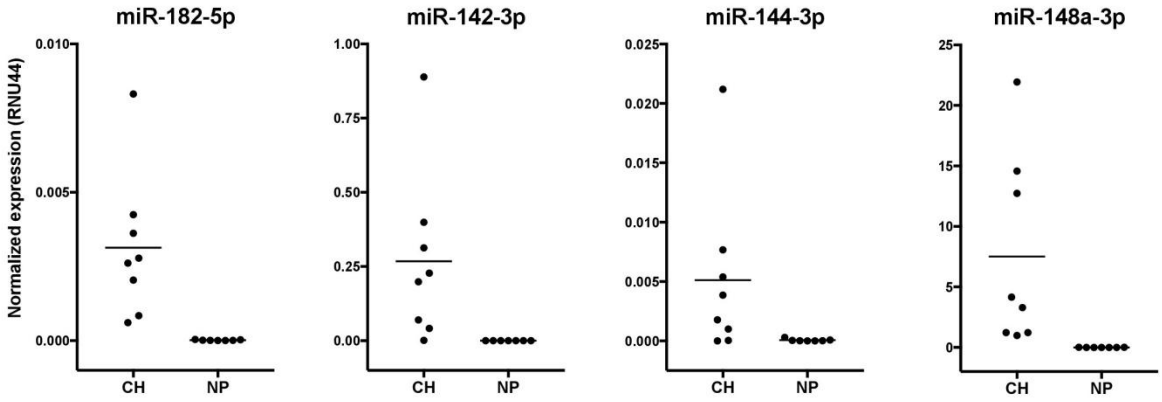

Figure 9. RT-qPCR validation of miRNA expression in chordoma (CH) and nucleus pulposus (NP) samples. Normalized expression was calculated with the 2 dCt method, with RNU44 as the reference gene. All differences were statistically significant (Mann-Whitney test, p 0.001).

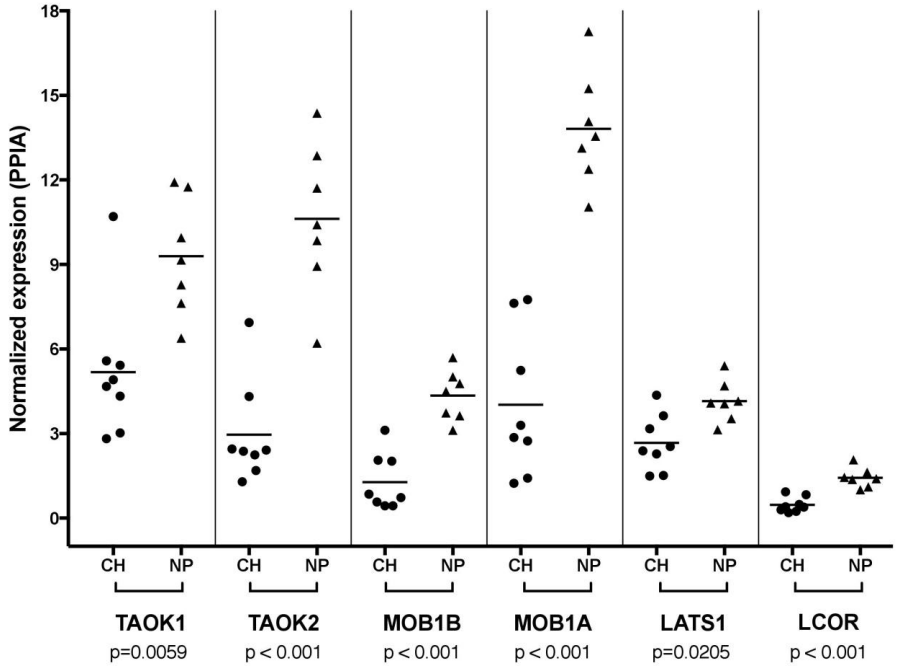

Figure 10 RT-qPCR validation of mRNA expression in chordoma (CH) and nucleus pulposus (NP) samples. Normalized expression was calculated with the 2dCt method, with PPIA as the reference gene (p-values are from Mann-Whitney tests)

5. Discussion

5.1. Discussion VDR

The relationship between vitamin D and muscle performance has already been investigated in numerous studies. The results of these studies proved that vitamin D, in addition to its many other effects, affects muscle function and performance and plays a key role in maintaining balance. However, the correlations between polymorphisms of the vitamin D receptor gene and muscle performance have not been investigated so far, and there is little information about how polymorphisms of the VDR gene affect muscle function in a healthy, young, average population.

In single-marker association tests, the "AA" genotype of rs1544410 (BsmI), and the "CC" genotype of rs731236 (TaqI) were associated with stronger hand grip in the dominant hand, while the "TT" genotype of rs4516035 (A1012G) was associated with stronger hand grip in both hands. Consistently, the VDR haplotype constructed by three SNPs (Bsml, TaqI, and rs10783215) located in the 3′ part of the gene showed a strong relationship with hand grip strength for both hands. We found that the rarest haplotype “GTT” (commonly known as “bTT”) is associated with the weakest HGS, while the more common haplotype “ACT” (BtT) is associated with the strongest HGS. Polymorphisms of the VDR gene related to muscle function are less studied, and considerably less data are available on the relationship between VDR polymorphisms and muscle strength in a healthy young population. In the literature, only a few studies have analyzed the associations between VDR polymorphisms and HGS, an important hereditary muscle phenotype, and to our knowledge, the association between VDR genotypes and HGS in both (dominant and non-dominant) hands has not been published yet. The BsmI polymorphism in the eighth intron is in strong linkage disequilibrium with two other genotypic SNPs (TaqI and rs10783215) from the 3' part of the gene. Carrier women, traditionally called “BB” (“AA genotype”) have greater fat-free mass and hamstring strength than “bb” (“GG”) genotype carriers [26]. Contradictory results were published by Geusens et al. [25], who first described the association between VDR polymorphisms and different muscle phenotypes and found that the BsmI allele “B” is associated with lower quadriceps strength in non-obese women. This study reported the only significant association of BsmI with HGS to date, where the “B” allele was associated with lower HGS in non-obese women, but this effect was reversed in obese

subjects. The “C” allele of the TaqI polymorphism (rs731236) is conventionally called “t” by many studies. TaqI is a non-synonymous SNP located in exon 9. Iki et al.[73] studied this in relation to different muscle phenotypes, and no correlation was found between the studied muscle parameters and the SNP. In their study, a borderline non-significant ($p = 0.06$) trend was observed for HGS, where 'Bt' homozygotes showed higher HGS than 'bT' carriers. This haplotype association was significant for isometric and concentric knee extensor strength.

Our results can explain the differences in previous results. First of all, in our study, we minimized the impact of disturbing environmental factors that accumulate with aging by analyzing a sample of young children. Second, all genotype–phenotype associations were adjusted for demographic variables shown to influence HGS in both arms. Furthermore, in addition to the single-marker association tests, we also performed haplotype analyzes for a deeper analysis of possible genetic associations.

Previously published molecular studies can provide a reasonable explanation for the significant association between genetic markers located in the two main regulatory regions of the HGS and VDR genes. The rs45160335 (A1012G) polymorphism is located in the 1e promoter region of the 5′ part of the gene. The transition from "A" to "G" in this SNP negatively modifies the ability of the VDR promoter to bind the transcription factor GATA-3. [74] The “A” allele results in increased promoter activity, which was confirmed by Fang et al. [80] using luciferase activity measurements. The 3′ UTR haplotype appears to be associated with VDR mRNA stability. In cell line studies, the "baT" haplotype of the BsmI-ApaI-TaqI construct was associated with a 15% higher mRNA level and a 30% slower mRNA decay compared to the "BAt" haplotype. [75] Based on these previous molecular results, the higher HGS associated with the “T” allele (A1012G “A”) of the rs4516035 promoter SNP and the “ACT” 3′ UTR haplotype (BtT) may be a direct genomic effect of the variants, resulting in higher VDR expression and more stable VDR.

However, the direct biological effect of the examined genomic variants has not been reported in muscle cells, and further studies are necessary to confirm the above genetic associations. Animal models have confirmed that vitamin D deficiency and congenital disorders of the vitamin D endocrine system can cause muscle weakness. VDR is a

ligand-dependent transcription factor that mediates the genomic effect of vitamin D on muscle cells. On the other hand, VDR can act as a membrane receptor for vitamin D, mediating its non-genomic effects.[76, 77, 80] Several RCTs have investigated the effects of vitamin D deficiency and vitamin D supplementation on muscle function and balance, but results are inconsistent and available evidence is limited by methodologically heterogeneous studies.[76, 80]

The conflicting results can be related to the possible modulating effect of VDR gene variants. None of the published clinical studies took into account the genetic background, but it is able to modify the molecular effect of the vitamin D ligand at the cellular level.

From the literature comparison above, it can be seen that genetic research related to muscle performance, even with regard to a more frequently studied gene such as VDR, does not provide clear results. Different methodologies, distorted study populations, and population genetic differences may be behind the different, contradictory results, but in accordance with the above literature results, our research definitely supports the role of VDR in muscle function and the genetic determination of the function.

5.2. Declarations

The research leading to these results received funding from the European Community's Seventh Framework Programme (FP7, 2007–2013) under grant agreement no. HEALTH-F2-2008-201626 (GENODISC project).

5.3. Discussion chordoma

In the present study, we used next-generation sequencing for the combined analysis of global miRNA and mRNA gene expression changes in sacral chordoma. This multi-omics approach and the subsequent detailed bioinformatic analysis predicted a novel molecular mechanism for the regulation of self-renewal in chordoma, never identified before. Altered expression of the miRNA regulators and their target genes in the self-renewal pathway was validated using quantitative PCR measurements.

5.3.1. Dysregulated miRNA Pattern in Chordoma

A large set (N=126) of significantly dysregulated miRNAs in chordoma was identified in our study. A positive correlation between the expression changes for many miRNAs

and their host genes suggests that several miRNAs are regulated predominantly at the transcriptional level, whereas for other miRNAs the transcriptional changes are counteracted or buffered by regulating the processing steps. Reciprocally regulated sets of differentially expressed genes and miRNAs, combined with miRNA target prediction, highlighted a set of genes in chordoma that may be regulated by the altered miRNA network. Upregulated miRNAs appear to target elements of growth factor-mediated signaling networks, including the axon guidance pathway, which is often activated in other cancer types. Downregulated miRNAs may target a very different network: most of the enriched pathways are associated with the regulation of immune function, notably with immune suppression. For example, suppressed miR-665 expression may contribute to the activation of TNF signaling pathways, which may help the immune evasion of the chordoma cells through stabilizing the undifferentiated phenotype, and/or through immunoediting in the tumor microenvironment. Other putative miRNA-regulated genes in these pathways may reflect in part the immunosuppressive phenotype of chordoma cells, but also the presence of inactivated tumor-infiltrating immune cells. We hypothesize that the dominant miRNA regulators of chordoma are those with the most target genes in the differentially expressed gene set: the upregulated miR-182-5p, the miR-15/16/195/424/497-5p family, miR-148a, miR-140, miR-142, and miR-199a/b-5p, and the downregulated miR-3151-3p, miR-377-5p, and miR-665. miR-182-5p is a well-characterized oncomir promotingaggressive phenotypes [123-125], and miR-377-5p acts as a tumor suppressor in different tumor types [126, 127]. The other dysregulated miRNAs, however, may have controversial, context-dependent effects in tumor cells. For instance, miR-424 may act either as a tumor suppressor or as an oncogene [128, 129], overexpression of miR-497 may contribute to chemoresistance or metastasis inhibition in different model systems [130, 131], and miR-223 may regulate chemoresistance differently in different tumor types [132, 133]. Similar to melanoma cells, we observed the reciprocal regulation of miR-126 (overexpressed) and miR-221/222 (downregulated) in chordomas [134].

5.3.2. Dysregulated miRNA/mRNA Network can be Associated with Chordoma Recurrence

A gene set targeted by several dysregulated miRNAs, having a possible role in the self-renewal process of chordoma stem cells, was identified through a detailed bioinformatic analysis. According to the cancer stem cell theory, standard chemotherapy can kill most cells in a tumor; however, cancer stem cells are resistant and remain viable. Despite the small number of such cells in the tumor mass, they might be the cause of tumor recurrence, sometimes many years after the "successful" treatment of the primary tumor, which is very frequent in chordoma [135, 136].

A dysregulated network identified in our study includes 7 miRNAs that may regulate 5 target genes, all of which have potential roles in stem cell renewal. Significantly lower expression of the LCOR (ligand-dependent nuclear receptor corepressor) gene in chordoma can be caused in part by the upregulation of seven miRNAs (miR-101, miR-142, miR-144, miR-148, miR-182, miR-199a, and miR-424), which are predicted or validated regulators of the LCOR mRNA. LCOR had been identified as a tumor suppressor in prostate cancer [137], and a recent study showed that the miR-199a-LCOR-IFN axis is a critical regulator maintaining the stem cell phenotype in breast cancer [138]. Some of the same upregulated miRNAs may target the elements of the Hippo signaling pathway - importantly, most of these miRNA-mRNA interactions have been validated in different model systems, as summarized in Supplementary Table 6. Five regulators of this pathway (TAOK1, TAOK2, MOB1A/B, and LATS1) were significantly suppressed in chordoma samples, and two of them are predicted targets for upregulated miRNAs (TAOK1 for miR-142, and MOB1B for miR-148a and miR-182). All five genes are responsible for the inhibition of the YAP/TAZ complex, a central effector regulating tissue proliferation and self-renewal in normal and cancer tissues [139-141] (Figure 4). Although we did not observe altered expression of previously identified YAP/TAZ target genes in chordoma, it is well known that the target gene set of YAP/TAZ is highly context-dependent [141]. Other regulators of stemness, such as MYBL1 [142] and RECK [143] are also predicted targets for multiple upregulated miRNAs in our study, and, together with LCOR and YAP/TAZ, may form a regulatory network maintaining the stem-cell phenotype in chordoma. Importantly, a recent study established that the YAP/TAZ pathway is indeed activated in chordoma tumors and

contributes to cancer growth and stemness [144]. This study identified the YAP gene as one of the Brachyury targets, activated at the transcriptional level. Based on our results, the overexpressed oncomirs can also contribute to the activation of the pathway.

This feature of stemness regulation in chordoma tumors, suggested by the bioinformatics analysis of the NGS results, can be related to the unique origin of this tumor (i.e., originated from notochordal remnants) and could explain the high rate of local recurrence as well as its resistance to approved chemotherapeutic agents.

5.3.3. Strength and Limitations

The strength of our study is the integrated analysis of the noncoding and coding RNA expression profile, leading to the prediction of dominant miRNA regulators in chordoma, and the prediction of a novel miRNA-mRNA regulatory network. Another strength of our study is the validation of the miRNA-mRNA gene expression changes by RT-qPCR. Differences in miRNA expression profiles may be variable depending on the location of chordoma and even among different chordoma cell lines [41].

The CH and NP samples in our study came from patients forming homogenous groups based on their clinical characteristics, to minimize the inter-donor heterogeneity. To avoid additional differences in the tumor transcriptomes, all chordoma samples were harvested from sacral chordoma tumors, and the patients did not receive any previous neo-adjuvant treatment. The small number of chordoma samples is a limitation of the study, but given that chordoma is a rare tumor (less than 1/million person/year new cases in Hungary), larger sample sets can be collected only in international collaborations, or in large countries. That said, even our small sample set corroborated previous findings, such as downregulation of miR-31 or overexpression of cytokeratin genes, and provides a reliable molecular portrait of chordoma.

Another important factor in the comparative chordoma studies is the type of the control sample. As chordoma originates from the notochord, the use of muscle [39, 43], fibroblast [44] or nasal turbinate tissue [145] as controls can lead to biased results, revealing tissue-specific differences of miRNA expression profiles, rather than tumor-specific differences. Similar to the majority of other authors [38, 41, 42, 45] we compared the transcriptome profile of chordoma to nucleus pulposus cells, which were derived from the single mature, non-dysplastic tissue type originating from the

notochord in humans. We obtained NP tissue samples from moderately degenerated intervertebral discs during fusion procedures to avoid the contamination of the NP tissue with annulus fibrosus or cartilage parts, which is usually present in herniated disc materials. Several extraction methods were tested, but none yielded high-quality RNA suitable for NGS from surgical NP tissue samples; therefore, we used primary NP cell cultures as controls. The same setup was applied in the studies of Bayrak et al. [38] and Gulluoglu et al. [41], and our findings strongly supported their results: namely, the overexpression of miR-140 and miR-148a, as well as the decreased expression of miR-31 and miR-222 in chordoma. In fact, miR-31 was the most downregulated miRNA in chordoma in our study.

Although our results are in good agreement with previous studies, which used similar sample sets but different RNA profiling methods [38, 41], further functional validation is needed to confirm our findings. This might be especially important in understanding the complex miRNA-dependent regulation of LCOR if novel therapeutic approaches are considered to target the LCOR-mediated IFN-responsiveness in chordoma. Also, different therapeutic approaches may be used to target the YAP/TAZ complex, affecting specific protein-protein interactions or downstream targets [146]. However, given the complexity and variability of the YAP/TAZ network, the best and critical targets may well be different in different tumor types. Importantly, our study revealed several regulatory miRNAs, the antimiR-based therapeutic suppression of which may lead to the reactivation of the Hippo pathway and the inhibition of YAP/TAZ and may target other stemness genes as well. Lastly, the clinical consequences of the ncRNA-regulated gene expression changes and the possible prognostic value of ncRNAs need further investigation on large, multicenter sample sets.

5.4. Ethics Approval and Consent to Participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments. The study was approved by the Scientific and Research Ethics Committee of the Medical Research Council (49777/2012/EKU (751/PI/12.)). Informed consent was obtained from all individual participants included in the study.

6. Conclusion

6.1. Conclusion VDR

Based on our results, a direct effect of VDR polymorphisms on muscle function can be assumed; however, this hypothesis needs to be further investigated. Determining the individual genetic background may not only be important in estimating the risk of pathology related to muscle function, but may also explain the discrepancy in published clinical results regarding the relationship between vitamin D and locomotor phenotypes. Our study requires further replication on other cohorts, but on the basis of the above correlations, we can recommend the genotyping of VDR candidate variants in future clinical trials focusing on vitamin D metabolism and muscle function.

Genetic associations with muscle strength-related phenotypes have already been described for more than 40 candidate genes. The present study demonstrates a significant association between VDR gene variants and handgrip strength – a well-defined muscle strength phenotype – in a primary school population.

Our study has some important limitations. First of all, we used a cluster sampling technique involving different primary schools from different regions of the country, and from these schools we included the entire healthy school population belonging to the age group under study. However, the lack of parental consent and the exclusion of subjects with missing genotypes reduced the studied population by 8.9%. Second, we did not apply α-level correction during the genetic association testing process. We followed this method because we used a hypothesis-driven approach where we calculated the effect of the candidate SNPs on the phenotype.

Moreover, an important “internal” validation was used to confirm the correlations of the study, even if the type I error rate was not reduced conservatively. A significant association between the VDR promoter SNP, the 3′ UTR haplotype, and HGS was found for both hands, which supports our working hypothesis that the genetic background of HGS related to muscle biology should be uniform and independent of hemispheric dominance. The above limitations may affect the reliability of our results, therefore it is strongly recommended to repeat the study independently on different populations.

As a conclusion of our research, we can state that mapping the genetic background of hand grip strength may be important in sports and general health research in the future.

6.2. Conclusion chordoma

In chordoma development, crucial cancer driver genes have not been identified so far. Despite the identification of numerous genes with dysregulated expression in vitro or in vivo in chordoma (e.g., brachyury, EGFR, Ezrin, MMP-9, c-MET, FHIT, etc. [147]), conventional or targeted chemotherapeutic agents showed a partial response at best-case scenarios in clinical case series. Significant genomic rearrangements were found in several chordoma samples in previous studies [32, 148], suggesting that the development and recurrence of the tumor can be related to the dysregulation of hundreds of genes and their complex regulatory networks such as non-coding RNAs. The recent development of high-throughput molecular profiling techniques (i.e., next-generation RNA sequencing) has further increased the possibility of performing comparative and integrated RNA profiling studies facilitating the detection of the alterations in the transcriptome and related regulatory networks [52, 149, 150].

Applying these techniques on surgical sacral chordoma tumor samples, several miRNAs with significant regulatory potential and a previously unidentified miRNA/mRNA network involved in stemness regulation have been identified in our study. Molecular elements of these pathways are possible targets for future therapy development in chordoma.

7. Summary

A better understanding of the functioning and healing of the musculoskeletal system allows for the development of specific treatment strategies. Many of the interesting techniques discussed in this thesis are in stages of development. Although these emerging technologies may develop into significant clinical treatment options, their full impact must be scientifically critically evaluated.

7.1. Vitamin D receptor gene

- A positive correlation between muscle strength and some biologically significant candidate VDR gene variants was revealed.
- "AA" genotype of BsmI, and "CC" genotype of TaqI were associated with stronger hand grip in the dominant hand. The "TT" genotype of A1012G was associated with stronger hand grip in both hands.
- The VDR haplotype constructed by Bsml, TaqI, and rs10783215 showed a strong relationship with hand grip strength for both hands.

7.2. Chordoma

- Until now, RNA isolation from freshly frozen chordoma has encountered obstacles due to the mucinous material of the tumor. The advantage of the present research is that, after overcoming the difficulties of isolating RNA from a freshly frozen tumor sample, the authors succeeded in setting up a new RNA isolation protocol, which for the first time enabled the sequencing of total RNA in chordoma.
- Succesful next-generation sequencing for the combined analysis of global miRNA and mRNA gene expression changes was done in sacral chordoma
- We successfully identified a large number of differentially expressed miRNAs between chondroma and nucleus pulopsus, where 74 miRNAs were upregulated and 52 downregulated in chordoma.
- We successfully validated that activation of the Hippo pathway is having role in the self-renewal mecahnism of chordoma.

8. Bibliography

1. International Human Genome Sequencing, C., *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.
2. Puzas, J.E., R.J. O'Keefe, and J.R. Lieberman, *The orthopaedic genome: what does the future hold and are we ready?* J Bone Joint Surg Am, 2002. **84**(1): p. 133-41.
3. Pasquinelli, A.E., *MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship.* Nat Rev Genet, 2012. **13**(4): p. 271-82.
4. Anderson, P. and N. Kedersha, *RNA granules: post-transcriptional and epigenetic modulators of gene expression.* Nat Rev Mol Cell Biol, 2009. **10**(6): p. 430-6.
5. Finnegan, E.F. and A.E. Pasquinelli, *MicroRNA biogenesis: regulating the regulators.* Crit Rev Biochem Mol Biol, 2013. **48**(1): p. 51-68.
6. Aicale, R., et al., *Genetics in orthopaedic practice.* J Biol Regul Homeost Agents, 2019. **33**(2 Suppl. 1): p. 103-117 XIX Congresso Nazionale S I C O O P Societa' Italiana Chirurghi Ortopedici Dell'ospedalita' Privata Accreditata.
7. Cech, T.R. and J.A. Steitz, *The noncoding RNA revolution-trashing old rules to forge new ones.* Cell, 2014. **157**(1): p. 77-94.
8. Evans, C.H. and R.N. Rosier, *Molecular biology in orthopaedics: the advent of molecular orthopaedics.* J Bone Joint Surg Am, 2005. **87**(11): p. 2550-64.
9. September, A.V., et al., *Variants within the COL5A1 gene are associated with Achilles tendinopathy in two populations.* Br J Sports Med, 2009. **43**(5): p. 357-65.
10. Panoutsopoulou, K., et al., *Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study.* Ann Rheum Dis, 2011. **70**(5): p. 864-7.
11. Valdes, A.M., et al., *Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee.* Arthritis Rheum, 2007. **56**(1): p. 137-46.

12. Raine, E.V., et al., *Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues.* BMC Musculoskelet Disord, 2013. **14**: p. 85.

13. Thieme, W.T. and R. Wynne-Davies, *Clinical examination and urinary oestrogen assays in newborn children with congenital dislocation of the hip.* J Bone Joint Surg Br, 1968. **50**(3): p. 546-50.

14. Matzko, M.E., T.R. Bowen, and W.R. Smith, *Orthogenomics: an update.* J Am Acad Orthop Surg, 2012. **20**(8): p. 536-46.

15. Albagha, O.M., et al., *Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone.* Nat Genet, 2010. **42**(6): p. 520-4.

16. Mokone, G.G., et al., *The COL5A1 gene and Achilles tendon pathology.* Scand J Med Sci Sports, 2006. **16**(1): p. 19-26.

17. Bond, M., et al., *A phase II study of imatinib mesylate in children with refractory or relapsed solid tumors: a Children's Oncology Group study.* Pediatr Blood Cancer, 2008. **50**(2): p. 254-8.

18. Daroszewska, A., et al., *Susceptibility to Paget's disease of bone is influenced by a common polymorphic variant of osteoprotegerin.* J Bone Miner Res, 2004. **19**(9): p. 1506-11.

19. Cooper, R., et al., *Objectively measured physical capability levels and mortality: systematic review and meta-analysis.* BMJ, 2010. **341**: p. c4467.

20. den Ouden, M.E., et al., *Physical performance characteristics related to disability in older persons: a systematic review.* Maturitas, 2011. **69**(3): p. 208-19.

21. Norman, A.W., *From vitamin D to hormone D: fundamentals of the vitamin D endocrine system essential for good health.* Am J Clin Nutr, 2008. **88**(2): p. 491S-499S.

22. Hager-Ross, C. and B. Rosblad, *Norms for grip strength in children aged 4-16 years.* Acta Paediatr, 2002. **91**(6): p. 617-25.

23. Cohen, D.D., et al., *Low muscle strength is associated with metabolic risk factors in Colombian children: the ACFIES study.* PLoS One, 2014. **9**(4): p. e93150.

24. Silva, C., et al., *Handgrip strength and nutrition status in hospitalized pediatric patients.* Nutr Clin Pract, 2014. **29**(3): p. 380-5.

25. Dekkers, K.J., et al., *Upper extremity strength measurement for children with cerebral palsy: a systematic review of available instruments.* Phys Ther, 2014. **94**(5): p. 609-22.

26. Katzmarzyk, P.T., et al., *Familial aggregation of 7-year changes in musculoskeletal fitness.* J Gerontol A Biol Sci Med Sci, 2001. **56**(12): p. B497-502.

27. Silventoinen, K., et al., *Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men.* Genet Epidemiol, 2008. **32**(4): p. 341-9.

28. Frederiksen, H., et al., *Hand grip strength: a phenotype suitable for identifying genetic variants affecting mid- and late-life physical functioning.* Genet Epidemiol, 2002. **23**(2): p. 110-22.

29. Crocco, P., et al., *A common polymorphism in the UCP3 promoter influences hand grip strength in elderly people.* Biogerontology, 2011. **12**(3): p. 265-71.

30. Stratton, M.R., P.J. Campbell, and P.A. Futreal, *The cancer genome.* Nature, 2009. **458**(7239): p. 719-24.

31. Stratton, M.R., *Journeys into the genome of cancer cells.* EMBO Mol Med, 2013. **5**(2): p. 169-72.

32. Stephens, P.J., et al., *Massive genomic rearrangement acquired in a single catastrophic event during cancer development.* Cell, 2011. **144**(1): p. 27-40.

33. Pillay, N., et al., *A common single-nucleotide variant in T is strongly associated with chordoma.* Nat Genet, 2012. **44**(11): p. 1185-7.

34. Xue, M., Y. Zhuo, and B. Shan, *MicroRNAs, Long Noncoding RNAs, and Their Functions in Human Disease.* Methods Mol Biol, 2017. **1617**: p. 1-25.

35. Ling, H., M. Fabbri, and G.A. Calin, *MicroRNAs and other non-coding RNAs as targets for anticancer drug development.* Nat Rev Drug Discov, 2013. **12**(11): p. 847-65.

36. Piraino, S.W. and S.J. Furney, *Beyond the exome: the role of non-coding somatic mutations in cancer.* Ann Oncol, 2016. **27**(2): p. 240-8.

37. Yoon, J.H., K. Abdelmohsen, and M. Gorospe, *Functional interactions among microRNAs and long noncoding RNAs.* Semin Cell Dev Biol, 2014. **34**: p. 9-14.

38. Bayrak, O.F., et al., *MicroRNA expression profiling reveals the potential function of microRNA-31 in chordomas.* J Neurooncol, 2013. **115**(2): p. 143-51.

39. Duan, Z., et al., *Differential expression of microRNA (miRNA) in chordoma reveals a role for miRNA-1 in Met expression.* J Orthop Res, 2010. **28**(6): p. 746-52.

40. Duan, Z., et al., *Prognostic significance of miRNA-1 (miR-1) expression in patients with chordoma.* J Orthop Res, 2014. **32**(5): p. 695-701.

41. Gulluoglu, S., et al., *The potential function of microRNA in chordomas.* Gene, 2016. **585**(1): p. 76-83.

42. Kuang, L., et al., *Overexpression of adenosine deaminase acting on RNA 1 in chordoma tissues is associated with chordoma pathogenesis by reducing miR-125a and miR-10a expression.* Mol Med Rep, 2015. **12**(1): p. 93-8.

43. Wei, W., et al., *miR-219-5p inhibits proliferation and clonogenicity in chordoma cells and is associated with tumor recurrence.* Oncol Lett, 2016. **12**(6): p. 4568-4576.

44. Zhang, Y., et al., *MicroRNA-608 and microRNA-34a regulate chordoma malignancy by targeting EGFR, Bcl-xL and MET.* PLoS One, 2014. **9**(3): p. e91546.

45. Zou, M.X., et al., *Reduced expression of miRNA-1237-3p associated with poor survival of spinal chordoma patients.* Eur Spine J, 2015. **24**(8): p. 1738-46.

46. Huang, W., et al., *Development and Validation of a 6-miRNA Prognostic Signature in Spinal Chordoma.* Front Oncol, 2020. **10**: p. 556902.

47. Li, L., et al., *Long Noncoding RNA LINC00525 Promotes the Aggressive Phenotype of Chordoma Through Acting as a microRNA-505-3p Sponge and Consequently Raising HMGB1 Expression.* Onco Targets Ther, 2020. **13**: p. 9015-9027.

48. Meng, T., et al., *Molecular Targeted Therapy in the Treatment of Chordoma: A Systematic Review.* Front Oncol, 2019. **9**: p. 30.

49. Zhang, H., et al., *miR-16-5p inhibits chordoma cell proliferation, invasion and metastasis by targeting Smad3.* Cell Death Dis, 2018. **9**(6): p. 680.

50. Choi, P.J., R.J. Oskouian, and R.S. Tubbs, *The Current Understanding of MicroRNA's Therapeutic, Diagnostic, and Prognostic Role in Chordomas: A Review of the Literature.* Cureus, 2018. **10**(12): p. e3772.

51. Yao, K., et al., *A competing endogenous RNA network identifies novel mRNA, miRNA and lncRNA markers for the prognosis of diabetic pancreatic cancer.* Tumour Biol, 2017. **39**(6): p. 1010428317707882.

52. Serrati, S., et al., *Next-generation sequencing: advances and applications in cancer diagnosis.* Onco Targets Ther, 2016. **9**: p. 7355-7365.

53. Close, G.L., et al., *Assessment of vitamin D concentration in non-supplemented professional athletes and healthy adults during the winter months in the UK: implications for skeletal muscle function.* J Sports Sci, 2013. **31**(4): p. 344-53.

54. Mohri, T., et al., *MicroRNA regulates human vitamin D receptor.* Int J Cancer, 2009. **125**(6): p. 1328-33.

55. Uitterlinden, A.G., et al., *Genetics and biology of vitamin D receptor polymorphisms.* Gene, 2004. **338**(2): p. 143-56.

56. Norman, K., et al., *Hand grip strength: outcome predictor and marker of nutritional status.* Clin Nutr, 2011. **30**(2): p. 135-42.

57. Dato, S., et al., *UCP3 polymorphisms, hand grip performance and survival at old age: association analysis in two Danish middle aged and elderly cohorts.* Mech Ageing Dev, 2012. **133**(8): p. 530-7.

58. Chiu, L.L., et al., *ACE I/D, ACTN3 R577X, PPARD T294C and PPARGC1A Gly482Ser polymorphisms and physical fitness in Taiwanese late adolescent girls.* J Physiol Sci, 2012. **62**(2): p. 115-21.

59. Yoshihara, A., et al., *Physical function is weakly associated with angiotensin-converting enzyme gene I/D polymorphism in elderly Japanese subjects.* Gerontology, 2009. **55**(4): p. 387-92.

60. Ahmetov, II, et al., *The association of ACE, ACTN3 and PPARA gene variants with strength phenotypes in middle school-age children.* J Physiol Sci, 2013. **63**(1): p. 79-85.

61. Girgis, C.M., et al., *The roles of vitamin D in skeletal muscle: form, function, and metabolism.* Endocr Rev, 2013. **34**(1): p. 33-83.

62. Halfon, M., O. Phan, and D. Teta, *Vitamin D: a review on its effects on muscle strength, the risk of fall, and frailty.* Biomed Res Int, 2015. **2015**: p. 953241.

63. Jia, F., et al., *Vitamin D receptor BsmI polymorphism and osteoporosis risk: a meta-analysis from 26 studies.* Genet Test Mol Biomarkers, 2013. **17**(1): p. 30-4.

64. Mohammadi, Z., et al., *Association between vitamin D receptor gene polymorphisms (Fok1 and Bsm1) and osteoporosis: a systematic review.* J Diabetes Metab Disord, 2014. **13**(1): p. 98.

65. Ji, G.R., et al., *BsmI, TaqI, ApaI and FokI polymorphisms in the vitamin D receptor (VDR) gene and risk of fracture in Caucasians: a meta-analysis.* Bone, 2010. **47**(3): p. 681-6.

66. Barr, R., et al., *Association between vitamin D receptor gene polymorphisms, falls, balance and muscle power: results from two independent studies (APOSS and OPUS).* Osteoporos Int, 2010. **21**(3): p. 457-66.

67. Onder, G., et al., *Vitamin D receptor polymorphisms and falls among older adults living in the community: results from the ilSIRENTE study.* J Bone Miner Res, 2008. **23**(7): p. 1031-6.

68. Bahat, G., et al., *BsmI polymorphism in the vitamin D receptor gene is associated with leg extensor muscle strength in elderly men.* Aging Clin Exp Res, 2010. **22**(3): p. 198-205.

69. Geusens, P., et al., *Quadriceps and grip strength are related to vitamin D receptor genotype in elderly nonobese women.* J Bone Miner Res, 1997. **12**(12): p. 2082-8.

70. Hopkinson, N.S., et al., *Vitamin D receptor genotypes influence quadriceps strength in chronic obstructive pulmonary disease.* Am J Clin Nutr, 2008. **87**(2): p. 385-90.

71. Windelinckx, A., et al., *Polymorphisms in the vitamin D receptor gene are associated with muscle strength in men and women.* Osteoporos Int, 2007. **18**(9): p. 1235-42.

72. Grundberg, E., et al., *Genetic variation in the human vitamin D receptor is associated with muscle strength, fat mass and body weight in Swedish women.* Eur J Endocrinol, 2004. **150**(3): p. 323-8.

73. Iki, M., et al., *Greater trunk muscle torque reduces postmenopausal bone loss at the spine independently of age, body size, and vitamin D receptor genotype in Japanese women.* Calcif Tissue Int, 2002. **71**(4): p. 300-7.

74. Wu, F.Y., et al., *Vitamin D receptor variability and physical activity are jointly associated with low handgrip strength and osteoporosis in community-dwelling elderly people in Taiwan: the Taichung Community Health Study for Elders (TCHS-E).* Osteoporos Int, 2014. **25**(7): p. 1917-29.

75. Roberts, H.C., et al., *A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach.* Age Ageing, 2011. **40**(4): p. 423-9.

76. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps.* Bioinformatics, 2005. **21**(2): p. 263-5.

77. Gonzalez, J.R., et al., *SNPassoc: an R package to perform whole genome association studies.* Bioinformatics, 2007. **23**(5): p. 644-5.

78. Tregouet, D.A. and V. Garelle, *A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies.* Bioinformatics, 2007. **23**(8): p. 1038-9.

79. Gauderman, W.J., *Sample size requirements for matched case-control studies of gene-environment interaction.* Stat Med, 2002. **21**(1): p. 35-50.

80. Gabriel, S.B., et al., *The structure of haplotype blocks in the human genome.* Science, 2002. **296**(5576): p. 2225-9.

81. Gottschalk, D., et al., *Matrix gene expression analysis and cellular phenotyping in chordoma reveals focal differentiation pattern of neoplastic cells mimicking nucleus pulposus development.* Am J Pathol, 2001. **158**(5): p. 1571-8.

82. McMaster, M.L., et al., *Chordoma: incidence and survival patterns in the United States, 1973-1995.* Cancer Causes Control, 2001. **12**(1): p. 1-11.

83. Weber, A.L., et al., *Cartilaginous tumors and chordomas of the cranial base.* Otolaryngol Clin North Am, 1995. **28**(3): p. 453-71.

84. Walcott, B.P., et al., *Chordoma: current concepts, management, and future directions.* Lancet Oncol, 2012. **13**(2): p. e69-76.

85. Gokaslan, Z.L., et al., *Mobile spine chordoma: results of 166 patients from the AOSpine Knowledge Forum Tumor database.* J Neurosurg Spine, 2016. **24**(4): p. 644-51.

86. Varga, P.P., et al., *Surgical treatment of sacral chordoma: prognostic variables for local recurrence and overall survival.* Eur Spine J, 2015. **24**(5): p. 1092-101.

87. Yamaguchi, T., M. Yamato, and K. Saotome, *First histologically confirmed case of a classic chordoma arising in a precursor benign notochordal lesion: differential diagnosis of benign and malignant notochordal lesions.* Skeletal Radiol, 2002. **31**(7): p. 413-8.

88. Yakkioui, Y., et al., *A comparison of cell-cycle markers in skull base and sacral chordomas.* World Neurosurg, 2014. **82**(1-2): p. e311-8.

89. Hunter, C.J., J.R. Matyas, and N.A. Duncan, *Cytomorphology of notochordal and chondrocytic cells from the nucleus pulposus: a species comparison.* J Anat, 2004. **205**(5): p. 357-62.

90. Youssef, C., et al., *Recent advances in understanding and managing chordomas.* F1000Res, 2016. **5**: p. 2902.

91. Hanna, S.A., et al., *Dedifferentiated chordoma: a report of four cases arising 'de novo'.* J Bone Joint Surg Br, 2008. **90**(5): p. 652-6.

92. Mizoe, J.E., et al., *Carbon ion radiotherapy for skull base chordoma.* Skull Base, 2009. **19**(3): p. 219-24.

93. Kelley, M.J., et al., *Familial chordoma, a tumor of notochordal remnants, is linked to chromosome 7q33.* Am J Hum Genet, 2001. **69**(2): p. 454-60.

94. Ghebranious, N., et al., *A missense T (Brachyury) mutation contributes to vertebral malformations.* J Bone Miner Res, 2008. **23**(10): p. 1576-83.

95. Pillay, N., et al., *P63 does not regulate brachyury expression in human chordomas and osteosarcomas.* Histopathology, 2011. **59**(5): p. 1025-7.

96. Nelson, A.C., et al., *An integrated functional genomics approach identifies the regulatory network directed by brachyury (T) in chordoma.* J Pathol, 2012. **228**(3): p. 274-85.

97. Bettegowda, C., et al., *Spinal column chordoma: prognostic significance of clinical variables and T (brachyury) gene SNP rs2305089 for local recurrence and overall survival.* Neuro Oncol, 2017. **19**(3): p. 405-413.

98. Presneau, N., et al., *Role of the transcription factor T (brachyury) in the pathogenesis of sporadic chordoma: a genetic and functional-based study.* J Pathol, 2011. **223**(3): p. 327-35.

99. Heery, C.R., et al., *Phase I Trial of a Yeast-Based Therapeutic Cancer Vaccine (GI-6301) Targeting the Transcription Factor Brachyury.* Cancer Immunol Res, 2015. **3**(11): p. 1248-56.

100. Presneau, N., et al., *Potential therapeutic targets for chordoma: PI3K/AKT/TSC1/TSC2/mTOR pathway.* Br J Cancer, 2009. **100**(9): p. 1406-14.

101. Scheipl, S., et al., *EGFR inhibitors identified as a potential treatment for chordoma in a focused compound screen.* J Pathol, 2016. **239**(3): p. 320-34.

102. Scheipl, S., et al., *Does insulin-like growth factor 1 receptor (IGF-1R) targeting provide new treatment options for chordomas? A retrospective clinical and immunohistochemical study.* Histopathology, 2012. **60**(6): p. 999-1003.

103. Lebellec, L., et al., *Advanced chordoma treated by first-line molecular targeted therapies: Outcomes and prognostic factors. A retrospective study of the French Sarcoma Group (GSF/GETO) and the Association des Neuro-Oncologues d'Expression Francaise (ANOCEF).* Eur J Cancer, 2017. **79**: p. 119-128.

104. Lipplaa, A., S. Dijkstra, and H. Gelderblom, *Efficacy of pazopanib and sunitinib in advanced axial chordoma: a single reference centre case series.* Clin Sarcoma Res, 2016. **6**: p. 19.

105. Hindi, N., et al., *Imatinib in advanced chordoma: A retrospective case series analysis.* Eur J Cancer, 2015. **51**(17): p. 2609-14.

106. Schuetze, S.M., et al., *Phase 2 study of dasatinib in patients with alveolar soft part sarcoma, chondrosarcoma, chordoma, epithelioid sarcoma, or solitary fibrous tumor.* Cancer, 2017. **123**(1): p. 90-97.

107. Schwab, J.H., et al., *Chordoma and chondrosarcoma gene profile: implications for immunotherapy.* Cancer Immunol Immunother, 2009. **58**(3): p. 339-49.

108. Vujovic, S., et al., *Brachyury, a crucial regulator of notochordal development, is a novel biomarker for chordomas.* J Pathol, 2006. **209**(2): p. 157-65.

109. Tamborini, E., et al., *Molecular and biochemical analyses of platelet-derived growth factor receptor (PDGFR) B, PDGFRA, and KIT receptors in chordomas.* Clin Cancer Res, 2006. **12**(23): p. 6920-8.

110. Fasig, J.H., et al., *Immunohistochemical analysis of receptor tyrosine kinase signal transduction activity in chordoma.* Neuropathol Appl Neurobiol, 2008. **34**(1): p. 95-104.

111. Szoverfi, Z., et al., *Primary Spinal Tumor Mortality Score (PSTMS): a novel scoring system for predicting poor survival.* Spine J, 2014. **14**(11): p. 2691-700.

112. Chen, K., et al., *A novel three-round multiplex PCR for SNP genotyping with next generation sequencing.* Anal Bioanal Chem, 2016. **408**(16): p. 4371-7.

113. Lazary, A., et al., *Effect of vertebroplasty filler materials on viability and gene expression of human nucleus pulposus cells.* J Orthop Res, 2008. **26**(5): p. 601-7.

114. Vitsios, D.M. and A.J. Enright, *Chimira: analysis of small RNA sequencing data and microRNA modifications.* Bioinformatics, 2015. **31**(20): p. 3365-7.

115. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.

116. Metsalu, T. and J. Vilo, *ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap.* Nucleic Acids Res, 2015. **43**(W1): p. W566-70.

117. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* Elife, 2015. **4**.

118. Mi, H., et al., *PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements.* Nucleic Acids Res, 2017. **45**(D1): p. D183-D189.

119. Kamburov, A., et al., *The ConsensusPathDB interaction database: 2013 update.* Nucleic Acids Res, 2013. **41**(Database issue): p. D793-800.

120. Andersen, C.L., J.L. Jensen, and T.F. Orntoft, *Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.* Cancer Res, 2004. **64**(15): p. 5245-50.

121. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

122. Perotti, V., et al., *NFATc2 is an intrinsic regulator of melanoma dedifferentiation.* Oncogene, 2016. **35**(22): p. 2862-72.

123. Hirata, H., et al., *MicroRNA-182-5p promotes cell invasion and proliferation by down regulating FOXF2, RECK and MTSS1 genes in human prostate cancer.* PLoS One, 2013. **8**(1): p. e55502.

124. Hirata, H., et al., *Oncogenic miRNA-182-5p targets Smad4 and RECK in human bladder cancer.* PLoS One, 2012. **7**(11): p. e51056.

125. Xue, J., et al., *miR-182-5p Induced by STAT3 Activation Promotes Glioma Tumorigenesis.* Cancer Res, 2016. **76**(14): p. 4293-304.

126. Li, B., et al., *MicroRNA-377 suppresses initiation and progression of esophageal cancer by inhibiting CD133 and VEGF.* Oncogene, 2017.

127. Sun, C., et al., *Long non-coding RNA NEAT1 promotes non-small cell lung cancer progression through regulation of miR-377-3p-E2F3 pathway.* Oncotarget, 2016. **7**(32): p. 51784-51814.

128. Dallavalle, C., et al., *MicroRNA-424 impairs ubiquitination to activate STAT3 and promote prostate tumor progression.* J Clin Invest, 2016. **126**(12): p. 4585-4602.

129. Rodriguez-Barrueco, R., et al., *miR-424(322)/503 is a breast cancer tumor suppressor whose loss promotes resistance to chemotherapy.* Genes Dev, 2017. **31**(6): p. 553-566.

130. Li, J., et al., *microRNA-497 overexpression decreases proliferation, migration and invasion of human retinoblastoma cells via targeting vascular endothelial growth factor A.* Oncol Lett, 2017. **13**(6): p. 5021-5027.

131. Zhu, D., et al., *Up-regulation of miR-497 confers resistance to temozolomide in human glioma cells by targeting mTOR/Bcl-2.* Cancer Med, 2017. **6**(2): p. 452-462.

132. Masciarelli, S., et al., *Gain-of-function mutant p53 downregulates miR-223 contributing to chemoresistance of cultured tumor cells.* Oncogene, 2014. **33**(12): p. 1601-8.

133. Huang, B.S., et al., *MiR-223/PAX6 Axis Regulates Glioblastoma Stem Cell Proliferation and the Chemo Resistance to TMZ via Regulating PI3K/Akt Pathway.* J Cell Biochem, 2017. **118**(10): p. 3452-3461.

134. Felli, N., et al., *AP2alpha controls the dynamic balance between miR-126&126* and miR-221&222 during melanoma progression.* Oncogene, 2016. **35**(23): p. 3016-26.

135. Ravasio, R., E. Ceccacci, and S. Minucci, *Self-renewal of tumor cells: epigenetic determinants of the cancer stem cell phenotype.* Curr Opin Genet Dev, 2016. **36**: p. 92-9.

136. Nassar, D. and C. Blanpain, *Cancer Stem Cells: Basic Concepts and Therapeutic Implications.* Annu Rev Pathol, 2016. **11**: p. 47-76.

137. Asim, M., et al., *Ligand-dependent corepressor acts as a novel androgen receptor corepressor, inhibits prostate cancer growth, and is functionally inactivated by the Src protein kinase.* J Biol Chem, 2011. **286**(43): p. 37108-17.

138. Celia-Terrassa, T., et al., *Normal and cancerous mammary stem cells evade interferon-induced constraint through the miR-199a-LCOR axis.* Nat Cell Biol, 2017. **19**(6): p. 711-723.

139. Noto, A., et al., *Stearoyl-CoA-desaturase 1 regulates lung cancer stemness via stabilization and nuclear localization of YAP/TAZ.* Oncogene, 2017.

140. Zanconato, F., M. Cordenonsi, and S. Piccolo, *YAP/TAZ at the Roots of Cancer.* Cancer Cell, 2016. **29**(6): p. 783-803.

141. Piccolo, S., M. Cordenonsi, and S. Dupont, *Molecular pathways: YAP and TAZ take center stage in organ growth and tumorigenesis.* Clin Cancer Res, 2013. **19**(18): p. 4925-30.

142. Guo, H., et al., *O-Linked N-Acetylglucosamine (O-GlcNAc) Expression Levels Epigenetically Regulate Colon Cancer Tumorigenesis by Affecting the Cancer Stem Cell Compartment via Modulating Expression of Transcriptional Factor MYBL1.* J Biol Chem, 2017. **292**(10): p. 4123-4137.

143. Hong, K.J., et al., *RECK inhibits stemness gene expression and tumorigenicity of gastric cancer cells by suppressing ADAM-mediated Notch1 activation.* J Cell Physiol, 2014. **229**(2): p. 191-201.

144. Shah, S.R., et al., *Brachyury-YAP Regulatory Axis Drives Stemness and Growth in Cancer.* Cell Rep, 2017. **21**(2): p. 495-507.

145. Bell, D., et al., *Whole-transcriptome analysis of chordoma of the skull base.* Virchows Arch, 2016. **469**(4): p. 439-49.

146. Juan, W.C. and W. Hong, *Targeting the Hippo Signaling Pathway for Tissue Regeneration and Cancer Therapy.* Genes (Basel), 2016. **7**(9).

147. Diaz, R.J. and M.D. Cusimano, *The biological basis for modern treatment of chordoma.* J Neurooncol, 2011. **104**(2): p. 411-22.

148. Diaz, R.J., et al., *High-resolution Whole-Genome Analysis of Skull Base Chordomas Implicates FHIT Loss in Chordoma Pathogenesis.* Neoplasia, 2012. **14**(9): p. 788-IN4.

149. Veneziano, D., et al., *Noncoding RNA: Current Deep Sequencing Data Analysis Approaches and Challenges.* Hum Mutat, 2016. **37**(12): p. 1283-1298.

150. Guo, L., et al., *An integrated analysis of miRNA, lncRNA, and mRNA expression profiles.* Biomed Res Int, 2014. **2014**: p. 345605.

9. References

9.1. Original Publications in the Topic of the Dissertation

Muscle strength is associated with vitamin D receptor gene variants.

Bozsodi A, Boja S, Szilagyi A, Somhegyi A, Varga PP, Lazary A. J Orthop Res. 2016 Nov;34(11):2031-2037. doi: 10.1002/jor.23220. Epub 2016 Mar 14.

Potential molecular mechanism in self-renewal is associated with miRNA dysregulation in sacral chordoma - A next-generation RNA sequencing study.

Bozsodi A, Scholtz B, Papp G, Sapi Z, Biczo A, Varga PP, Lazary A. Heliyon. 2022 Aug 13;8(8):e10227. doi: 10.1016/j.heliyon.2022.e10227. eCollection 2022 Aug.

9.2. Original Publications not in the Topic of the Dissertation

Primary Spinal Tumor Mortality Score (PSTMS): a novel scoring system for predicting poor survival. Szövérfi Z, Lazáry A, **Bozsódi A**, Klemencsics I, Eltes P, Varga PP

Spine J. 2014 Mar 17. pii: S1529-9430(14)00250-2. doi: 10.1016/j.spinee.2014.03.009.

Cross-cultural adaptation and validation of the Hungarian version of the Core Outcome Measures Index for the back (COMI Back).

Klemencsics I, Lazary A, Valasek T, Szoverfi Z, **Bozsodi A**, Eltes P, Fekete TF, Varga PP.

Eur Spine J. 2015 Jan 9

Validation of the Hungarian version of the Roland-Morris disability questionnaire.

Valasek T, Varga PP, Szövérfi Z**, Bozsodi A**, Klemencsics I, Fekete L, Lazary A.

Disabil Rehabil. 2015;37(1):86-90. doi: 10.3109/09638288.2014.909536. Epub 2014 Apr 21

[MEASURING THE FUNCTIONAL IMPAIRMENT OF THE LUMBAR SPINE]. Valasek

T, Varga PP, Klemencsics I, Szövérfi Z**, Bozsódi Á**, Éltes PE, Lazáry Á. Ideggyogy Sz. 2015 Mar 30;68(3-4):135-41.

Clinical Decision Making: Integrating Advances in the Molecular Understanding of Spine Tumors.

Goodwin CR, Abu-Bonsrah N, Bilsky MH, Reynolds JJ, Rhines LD, Laufer I, Disch AC, **Bozsodi A**, Patel SR, Gokaslan ZL, Sciubba DM, Bettegowda C. Spine (Phila Pa 1976). 2016 Oct 15;41 Suppl 20:S171-S177. doi: 10.1097/BRS.0000000000001836.

Risk factors for surgical site infection in elective routine degenerative lumbar surgeries.

Klemencsics I, Lazary A, Szoverfi Z, **Bozsodi A**, Eltes P, Varga PP. Spine J. 2016 Nov;16(11):1377-1383. doi: 10.1016/j.spinee.2016.08.018. Epub 2016 Aug 9

10. Acknowledgements

At the end of this doctoral thesis, I would like to gratefully thank all those who helped me in starting my scientific career.

First of all, I would like to express my gratitude to my supervisor, Dr. Áron Lazáry, the Scientific Director of the National Center for Spinal Disorders, whose continuous help enabled me to get involved in scientific life, allowed me to progress with my scientific work under his daily encouragement and supervision, and who set an example with his own life for the scientifically demanding cultivation of spine medicine.

I am indebted to Dr.Péter Pál Varga, Chief Physician, through whose help I got the opportunity to learn about orthopedics and spine medicine, and through whose kind support I got involved in the scientific life.

I would like to thank dr. Beáta Scholtz, Associate Professor of the Institute of Biochemistry and Molecular Biology of the University of Debrecen, who helped us in the implementation of our work. Her ideas and extensive knowledge boosted the work process.

Special thanks to Dr Zolta Sápi, chief physician of the 1st Institute of Pathological and Experimental Cancer Research, and Dr. Gergő Papp's persistent help and support in the field of laboratory work. Thank you for your openness, patience, and motivation, which allowed me to gain insight into molecular biological knowledge.

Thank you to my fellow doctoral students, Dr. To Ádám Biczó, Dr. To Péter Éltes, Dr. László Kiss, István Klemencsics, Júlia Szita, Hajnal Benjamin and Dr. Zsolt Szövérfi, the doctoral student years rich in memories.

I would also like to thank the Research and Developement staff of the National Spine Center, especially Zita Bognár and Edit Blaskó, for their daily and selfless help, which was also indispensable during my scientific work.

I am grateful to my wife Andrea Bozsódi-Almási, my children Máté Bozsódi and Milán Bozsódi, my parents and my brother, whose encouragement, love and patience were an extremely great resource during my PhD student years and still are for me today.

Supplementary Table 1

Most important candidate target genes of differentially expressed miRNAs in chordoma

Predicted target gene	log2FoldChange of target gene	miRNA	Cumulative weighted context++ score
LCOR	-3.77	hsa-miR-424-5p	-0.32
		hsa-miR-101-3p	-0.78
		hsa-miR-142-3p	-0.66
		hsa-miR-144-3p	-0.73
		hsa-miR-199a-5p	-0.67
MOB1B	-3.73	hsa-miR-148a-3p	-0.39
		hsa-miR-182-5p	-0.34
TAOK1	-4.50	hsa-miR-142-3p	-0.67
MITF	-3.94	hsa-miR-101-3p	-0.35
		hsa-miR-144-3p	-0.3
		hsa-miR-148a-3p	-0.74
		hsa-miR-182-5p	-0.73
MYBL1	-3.61	hsa-miR-424-5p	-0.5
		hsa-miR-144-3p	-0.49
RECK	-2.69	hsa-miR-497-5p	-0.45
		hsa-miR-182-5p	-0.46
ABL2	-3.44	hsa-miR-15a-5p	-0.37
		hsa-miR-143-3p	-0.39
		hsa-miR-148a-3p	-0.42
CAMK2N1	-3.48	hsa-miR-140-5p	-0.36

		hsa-miR-182-5p	-0.57
		hsa-miR-450b-5p	-0.78
DST	-5.26	hsa-miR-335-3p	-0.9
		hsa-miR-126-5p	-0.33
ECE1	-3.36	hsa-miR-140-3p	-0.52
		hsa-miR-199a-5p	-0.63
FAM73A	-3.36	hsa-miR-424-5p	-0.55
		hsa-miR-101-3p	-0.46
		hsa-miR-142-3p	-0.4
		hsa-miR-182-5p	-0.44
HMGA2	-9.19	hsa-miR-16-5p	-0.55
		hsa-miR-142-3p	-0.72
		hsa-miR-204-5p	-0.44
JMJD1C	-4.87	hsa-miR-96-5p	-0.58
		hsa-miR-148a-5p	-0.32
		hsa-miR-182-5p	-0.77
MIPOL1	-2.74	hsa-miR-497-5p	-0.43
		hsa-miR-140-5p	-0.34
		hsa-miR-192-5p	-0.46
NBEAL1	-4.58	hsa-miR-126-5p	-0.62
		hsa-miR-335-3p	-0.59
NDP	-6.90	hsa-miR-424-5p	-0.47
		hsa-miR-148a-3p	-0.4
		hsa-miR-223-3p	-0.43
PDPN	-4.14	hsa-miR-182-5p	-0.37
		hsa-miR-199a-5p	-0.7

RASA1	-2.07	hsa-miR-182-5p	-0.63
		hsa-miR-223-3p	-0.39
RGS17	-3.17	hsa-miR-144-3p	-0.54
		hsa-miR-182-5p	-1
SPIN1	-2.29	hsa-miR-96-5p	-0.86
		hsa-miR-182-5p	-0.44
SPTSSB	-4.40	hsa-miR-96-5p	-0.43
		hsa-miR-182-5p	-0.45
		hsa-miR-450b-5p	-0.57
TFRC	-3.04	hsa-miR-15a-5p	-0.46
		hsa-miR-148a-5p	-0.41
		hsa-miR-335-3p	-0.37
		hsa-miR-450b-5p	-0.75
RICTOR	-3.62	hsa-miR-142-3p	-0.72
ARHGAP21	-4.99	hsa-miR-199b-5p	-0.66
ARHGEF12	-4.61	hsa-miR-182-5p	-0.74
HMCN1	-4.96	hsa-miR-199a-5p	-0.69
HOXB6	-3.83	hsa-miR-126-5p	-0.64
NEK10	-4.62	hsa-miR-497-5p	-0.75
PRKACB	-3.42	hsa-miR-182-5p	-0.76
PTPN3	-3.97	hsa-miR-424-5p	-0.64
NFATC2	4.41	hsa-miR-431-3p	-0.35
		hsa-miR-665	-0.24
HIF3A	6.73	hsa-miR-377-5p	-0.63
		hsa-miR-485-5p	-0.64
PRELP	2.33	hsa-miR-31-5p	-0.4

		hsa-miR-665	-0.67
AKR1B10	7.73	hsa-miR-3151-3p	-0.89
C21orf119	2.19	hsa-miR-539-3p	-0.63
HEYL	7.35	hsa-miR-665	-0.62
IL15	2.55	hsa-miR-3151-3p	-0.82
LYPD6	4.84	hsa-miR-431-5p	-0.76
MYEOV2	2.53	hsa-miR-431-3p	-1.15
MYO5C	3.64	hsa-miR-377-5p	-0.9
P2RY13	7.60	hsa-miR-431-3p	-0.7
RADIL	4.65	hsa-miR-665	-0.72
SH2D1A	5.83	hsa-miR-31-5p	-0.87
SYNGR2	2.92	hsa-miR-665	-0.61
YPEL2	2.41	hsa-miR-3151-3p	-0.68