

Whole-transcriptome analysis of chordoma of the skull base

Diana Bell^{1,2} · Shaan M. Raza^{2,3} · Achim H. Bell⁴ · Gregory N. Fuller¹ · Franco DeMonte³

Received: 7 May 2016 / Revised: 16 June 2016 / Accepted: 30 June 2016 / Published online: 11 July 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Fourteen skull base chordoma specimens and three normal specimens were microdissected from paraffin-embedded tissue. Pools of RNA from highly enriched preparations of these cell types were subjected to expression profiling using whole-transcriptome shotgun sequencing. Using strict criteria, 294 differentially expressed transcripts were found, with 28 % upregulated and 72 % downregulated. The transcripts were annotated using NCBI Entrez Gene and computationally analyzed with the Ingenuity Pathway Analysis program. From these significantly changed expressions, the analysis identified 222 cancer-related transcripts. These 294 differentially expressed genes and non-coding RNA transcripts provide here a set to specifically define skull base chordomas and to identify novel and potentially important targets for diagnosis, prognosis, and therapy of this cancer. *Significance* Genomic profiling to subtype skull base chordoma reveals potential candidates for specific biomarkers, with validation by IHC for selected candidates. The

highly expressed developmental genes T, LMX1A, ZIC4, LHX4, and HOXA1 may be potential drivers of this disease.

Keywords Skull base · Chordoma · Transcriptome

Introduction

Chordomas are rare malignant tumors arising from notochordal remnants along the axial skeleton [1]. Chordomas, which account for only 1–4 % of all primary skeletal tumors, have an incidence of 1 per 1 million individuals per year and are found predominantly in men (58.4 %). They are generally locally aggressive, invasive neoplasms that are resistant to conventional chemo- and radiotherapy [2]. Chordomas most commonly arise in the sacrococcygeal area (50 %), followed by the skull base (35 %) and vertebral bodies (15 %); however, they also have been reported to originate in an extra-axial distribution or within soft tissue [3–5]. Morphologically, chordomas are identified as classical (conventional), chondroid, or dedifferentiated.

Typical of chordoma are physaliferous cells, which appear in clusters of large cells separated by fibrous septa into lobules and surrounded by basophilic extracellular matrix rich in mucin and glycogen. By morphology, chordomas are divided into classical (conventional), chondroid, and dedifferentiated. Although chordomas have been studied histologically, very little is known about the molecular mechanisms that drive these tumors. The gold-standard treatment for these tumors is en bloc resection; however, within the skull base, surrounding critical neurovascular structures and challenges in surgical access render clinical management of these patients difficult [3–5]. Therapeutic options for chordoma are limited, due to lack of effective chemotherapy options and relative radio resistance of these tumors necessitating high radiation doses.

Electronic supplementary material The online version of this article (doi:10.1007/s00428-016-1985-y) contains supplementary material, which is available to authorized users.

✉ Diana Bell
diana.bell@mdanderson.org

✉ Franco DeMonte
fdemonte@mdanderson.org

¹ Department of Pathology, University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

² Department of Head and Neck Surgery, University of Texas MD Anderson Cancer Center, Houston, TX, USA

³ Department of Neurosurgery, University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

⁴ Lone Star College, Houston, TX, USA

Despite optimal surgical resection followed by charged-particle radiotherapy, recurrence is a common event, reported as 30–85 %, with a median survival of approximately 6 years [4, 6, 7]. Recent studies have suggested that there may be a potential role for molecular therapy in the treatment of chordomas [1]. Work from Barth et al. on chordoma cell lines and tissues argue for an extension of palbociclib therapy to chordoma patients with a potential responder phenotype defined in this study [8]. Palbociclib is a specific inhibitor of CDK4/6. Several clinical trials and studies with palbociclib for various kinds of cancer show favorable effects on the progression-free survival and only moderate side effects (reviewed in [8]). Flanagan's group just reported (p-)EGFR expression in clinical samples, entertaining the efficacy of EGFR inhibitors in the chordoma treatment [9].

Histologically and phenotypically, the chordomas are well characterized, but currently, there are no predictive or prognostic biomarkers for chordoma [1]. Identification of predictive biomarkers would help identify patients at risk for local and systemic progression; this information could potentially aid in the identification of targeted therapies for patients likely to fail standard surgical and radiation options.

Advances in DNA sequencing have allowed massive parallel throughput and data volumes that eclipse the nucleic acid information content possible with other technologies, making feasible extensive genome analysis of groups of individuals, including analysis of sequence differences, polymorphisms, mutations, copy number variations, epigenetic variations, and transcript abundance. Biomarker discovery is an attractive potential application of this new technology.

Aberrant transcript expression includes changes in expression levels, isoforms, and polymorphisms, which are commonly observed in cancer; these aberrations could alter biological pathways and disease phenotypes. Next-generation sequencing (NGS) of RNA (RNA-Seq) has become a superior tool for studying the comprehensive transcriptome [10]. Since its inception, array technology has improved in terms of its sensitivity and dynamic range for assessing gene expression. However, RNA-Seq technology provides sequence information for all expressed transcripts, including new and unknown transcripts, much more accurately than array technology ever could and thus provides much greater knowledge of the cancer transcriptome [10, 11].

Understanding the transcriptome is essential for interpreting the functional elements of the genome, revealing the molecular constituents of cells and tissues, and understanding development and disease [10]. The key aims of transcriptomics are to catalog all species of transcript, including messenger RNAs (mRNAs), non-coding RNAs, and small RNAs; to determine the

transcriptional structure of genes (start sites, 5' and 3' ends, splicing patterns, and other post-transcriptional modifications); and to quantify the changing expression levels of each transcript under different conditions, e.g., during development of diseases [12].

The NGS method described here enables transcriptome-wide cancer biomarker discovery with archival formalin-fixed paraffin-embedded (FFPE) tissue. The FFPE material is linked to mature clinical records in hospital pathology archives. This material can be used for tumor gene expression profiling and therefore may enable rapid clinical biomarker discovery in studies that are statistically well-powered.

The purpose of this retrospective study was to identify abnormally expressed genes in skull base chordomas. We used high-throughput mRNA sequencing (RNA-Seq) to characterize the differences and similarities in transcriptome expression between patient specimens of skull base chordoma tissue and healthy skull base tissue.

Materials and methods

Tissue specimens

We identified 37 patients for whom clinical outcome data and FFPE specimens of primary skull base tumors were available. Six non-matching FFPE specimens of normal nasal turbinate tissue were analyzed in parallel.

Sample preparation and RNA sequencing

RNA was extracted from the 37 tumor and 6 normal tissue specimens using the RNeasy Universal kit (Qiagen, Gaithersburg, MD, USA). Total RNA quality and quantity were verified spectrophotometrically with a NanoDrop 100 spectrometer (ThermoScientific, Wilmington, DE, USA) and electrophoretically with a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). To construct Illumina-compatible libraries, we used the TruSeq RNA Gold library preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. In brief, we used the Ribo-Zero method to remove ribosomal RNA from total RNA. The residual RNA was chemically fragmented and converted into single-stranded complementary DNA (cDNA) using random hexamer priming. Double-stranded cDNA synthesis with strand selection, adapter ligation, and PCR amplification was used to generate the TruSeq whole-transcriptome cDNA library.

The library samples of 21 tumor and 6 normal specimens could be subjected to RNA-Seq on a HiSeq 2000 platform (Illumina) with 50 bp paired-end reads from each end of the RNA insert. The range of the adjusted read depths was 6.4–67.8 million. Most samples had 40–60

million reads (80–120 million if counting each paired read separately); the median number of reads per sample was 52 million.

RNA-Seq analysis of differentially expressed genes and isoforms

RNA-Seq analysis was performed with the RNAv9 pipeline developed by EQ-quintiles using open source programs and a variety of internally developed programs; many of which are available at <https://code.google.com/p/ea-utils/>. These tools were coordinated to perform the analysis, which included both quality control and gene and isoform quantification.

In brief, RNA-Seq reads were mapped to the human genome using the RNA-STAR (Spliced Transcripts Alignment to a Reference) software program (version 2.4; developed by Dobin et al. at Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA; email: dobin@cshl.edu) [13] and quantified using the RSEM software program (version 1.2.14).

Gene definitions of the RNAv9 analysis pipeline consist of 34,495 gene transcripts and 88,933 isoform transcripts. Transcript counts for gene expression levels were calculated, and the relative transcript abundance was determined as fragments per kilobase of exon per million fragments mapped (FPKM) using the Cufflinks software program (version 1.2.1) with upper quartile normalization. Raw data were extracted as FPKM values across all samples, and samples with zero values across more than 50 % of the genes were excluded.

The statistical significance of the fold change (FC) in a gene's expression between tumor tissue and normal tissue was determined using paired *t* tests with the null hypothesis that no difference existed between the two values. The false discovery rate (FDR) was controlled by using the Benjamini-Hochberg algorithm to adjust *p* values.

These analyses yielded consistent and reliable data for 3 normal and 14 tumor samples and produced 28,265 gene transcripts and 73,485 isoform transcripts representing the whole human transcriptome of skull base chordoma, including non-coding RNA (ncRNA) and excluding ribosomal and control transcripts (Supporting Table 1). Transcripts were identified and annotated automatically during the alignment analysis, and the remaining unknown annotations were manually updated using the ENSEMBL, UCSC Genome Browser, and NCBI Entrez Gene databases.

Gene and isoform transcripts that are characteristic of only skull base chordoma and that fulfilled the strict criteria of having FDR values below 0.012 (corresponding to unadjusted *p* values of less than $1.3E-4$ for genes and less than $5.4E-5$ for isoforms) and log₂ FCs greater 1 are shown in (Supporting Table 2).

Cluster analysis and heatmaps

To show the dissimilarity among all samples, we performed a dendrogram cluster analysis of the expression data of all log-transformed transcripts (not just the differentially expressed ones) using the matrix-cluster-sample-analysis correlation procedure, in which dissimilarity between neoplastic and normal samples is calculated as (1—correlation) (Fig. 2a).

Heatmaps with cluster analyses were generated for the top 100 differentially expressed genes and the top 100 differentially expressed isoforms based on their log₂ FC expression values. Similarity calculations were performed as average linkage clustering, based on the Euclidean distance of the samples, to compare the top 100 gene and top 100 isoform expressions.

All cluster and heatmap analyses and the visualization of differentially expressed genes were conducted by EA-quintiles using the R2.15.1 software package (www.r-project.org).

Computational analysis of pathway networks and biomarkers

Skull base chordoma pathway networks and biomarkers were identified using the Ingenuity Pathway Analysis (IPA) program (content version, 24718999; release date, 2015-09-14) from Ingenuity Systems (QIAGEN Silicon Valley, Redwood, CA).

Immunohistochemical analysis

To validate the gene expression patterns detected by RNA-Seq analysis, we used immunohistochemical analysis to assess the protein expression of five genes highly upregulated in and considered important for skull base chordoma. Immunohistochemical analysis was performed with antibodies against human T brachyury (Abcam, Cambridge, MA, USA), LMX1A, ZIC4, LHX4, and HOXA1 (LifeSpan Biosciences, Seattle, WA, USA). All slides were analyzed by the same pathologist (DB). For each protein, the results were categorized as positive (staining of >10 % of the area of the entire tissue specimen) or negative (staining of <10 % of the area of the entire tissue specimen).

Results

Clinical and histologic data

The clinical characteristics of the 37 patients are given in Table 1. Patient's ages at diagnosis ranged from 6 to 76 years. Most patients (51 %) were women. The most frequent primary

tumor site was the clivus ($n = 26$). The morphologic subtypes of chordoma, conventional (54 %, $n = 20$), chondroid (41 %, $n = 15$), and spindle/ dedifferentiated (5 %, $n = 2$), are illustrated in Fig. 1.

Genome-wide RNA-Seq analysis

RNA-Seq yielded sequence data for 21 of the 35 skull base chordoma specimens and 6 of the 6 normal tissue samples. Statistical analysis of these 27 samples' data identified 3 normal tissue samples (N2, N3, N4) and 14 neoplastic tissue samples (T3, T8, T11, T17, T18, T19, T22, T26, T28, T30, T31, T32, T34, and T36) for comparative analysis.

Hierarchical cluster analysis for dissimilarities among these 17 samples, which was based on the expression data for the top 5000 isoform transcripts, revealed no major group separations between normal tissue samples and tumor samples; Kaplan-Meier curve of progression-free survival (Fig. 2a).

All samples' expression data were used to create a heatmap and a dendrogram of the 100 gene-isoforms with the highest

differential expression between normal tissue and tumor specimens (Fig. 2b).

We then created a more complete and characteristic gene-transcript panel for skull base chordoma. Subjecting our RNA-Seq analysis results to strict statistical and biological criteria identified 294 differentially expressed gene transcripts (28 % upregulated and 72 % downregulated in chordoma as compared to normal tissue) as being the most characteristic of the disease (Supporting Table 2). We further annotated these transcripts using NCBI Entrez Gene and analyzed them for gene disease assignments, pathway networks, and biomarkers using the IPA program. Of the 294 genes with significantly differential expression, 222 were identified by IPA as being related to cancer.

Based on the 294 gene transcripts we identified as being characteristic of skull base chordoma (Supporting Table 2), this disease seems to have an abnormally high number of upregulated genes involved in development and differentiation. The seven such genes with the highest expression were *T* (T brachyury transcription factor; log2FC: 14), *LMX1A* (LIM homeobox transcription factor 1; log2FC: 11), *ZIC1* (Zic family member 1; log2FC: 8.4), *ZIC4* (Zic family member 4; log2FC: 8), *LHX4*

Table 1 Clinical characteristics of 37 chordoma patients

	Sample ID	Age	Sex	Location	Histology chordoma	Adjuvant therapy	FU months	FU (status)
1	T1	6	M	Clival	Spindle/dedifferentiated	Proton	72	Deceased 2nd malignancy
2	T8	27	M	Clival	Chondroid	Proton, chemo	120	DOD recurrence, drop met
3	T3	69	M	Pituitary	Conventional	XRT	N/A	Deceased 2nd malignancy
4	T6	39	F	Clival	Chondroid	XRT	216	AWD
5	T7	53	F	Sphenoid	Chondroid	XRT	108	NED
6	T9	43	F	Clival	Conventional		108	NED
7	T10	48	F	Clival	Chondroid		96	NED
8	T11	10	M	Clival	Spindle/dedifferentiated	XRT chemo	48	DM DOD
9	T12	58	F	Clival	Chondroid	XRT	96	NED
10	T13	34	M	Clival	Chondroid	XRT	96	NED
11	T15	45	F	Clival	Conventional	XRT	84	DM DOD
12	T16	57	MM	Clival	Conventional	Proton	36	N/A
13	T17	72	M	Pharynx	Chondroid	Proton	60	AWD
14	T18	47	F	Nasopharynx	Conventional	XRT	60	AWD
15	T19	74	M	Sphenoid	Conventional	Proton	48	AWD
16	T20	62	F	Nasopharynx	Chondroid	Proton	84	NED
17	T22	36	F	Clival	Conventional	Proton chemo	48	DM AWD
18	T23	61	M	Sphenoid	Chondroid	Proton	24	DM DOD
19	T25	60	M	Nasopharynx	Conventional			
20	T26	69	M	Clival	Chondroid	Proton	72	AWD
21	T27	74	F	Clival	Chondroid	IMRT		N/A
22	T28	65	M	Clival	Conventional	XRT; gamma-knife	120	AWD
23	T29	52	F	Clival	Conventional	XRT; gamma-knife; chemo	156	DOD
24	T30	69	M	Clival	Conventional	XRT; cyber-knife	12	N/A
25	T31	65	M	Clival	Conventional	Proton	12	N/A
26	T32	39	F	Clival	Conventional	XRT	12	
27	T33	3	F	Nasopharynx/clival	Chondroid			AWD
28	T34	6	F	Clival	Conventional	XRT	12	AWD
29	T35	57	M	Retropharyngeal	Conventional	XRT	204	AWD
30	T36	72	F	Clival	Chondroid	XRT	72	AWD
31	T37	60	M	Clival	Conventional	XRT	12	NED
32	T38	40	F	Clival	Conventional	XRT	12	AWD
33	T39	13	M	Clival	Conventional	Proton	12	AWD
34	T40	40	F	Clival	Conventional	Proton	108	AWD
35	T42	51	M	Sellar	Conventional	XRT	12	N/A
36	T43	68	M	Clival	Chondroid	XRT	4	NED
37	T44	68	F	Clival	Chondroid		3	NED

(LIM homeobox 4; log₂FC: 7), *HOXA2* (homeobox A2; log₂FC: 6), and *HOXA1* (homeobox A1, log₂FC: 6).

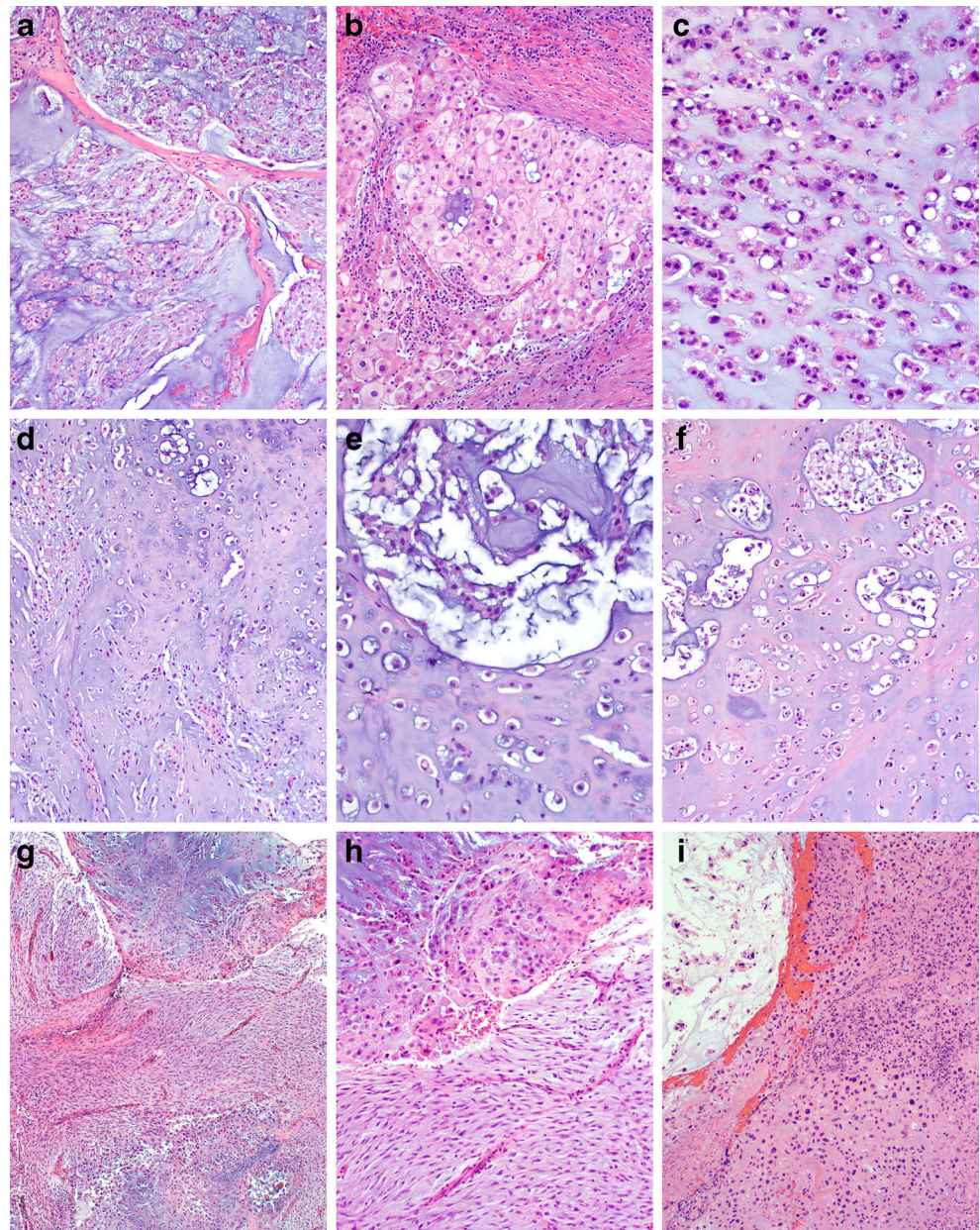
This disease also has many upregulated genes involved in cell matrix and cell membrane remodeling. The six such genes with the highest expression were *COL2A1* (collagen, type II, alpha 1; log₂FC: 11), *HRNR* (homerin; log₂FC: 11), *ACAN* (aggrecan; log₂FC: 8), *NIPAL4* (NIPA-like domain containing 4; log₂FC: 6), *FNI* (fibronectin 1; log₂FC: 6), and *CDSN* (corneodesmosin; log₂FC: 6).

We also found upregulated genes involved in cell signaling. The three genes with the highest expression were *CPLX2* (complexin 2; log₂FC: 7), *PPEF1* (protein phosphatase, EF-

hand calcium binding domain 1; log₂FC: 6), and *IQGAP3* (IQ motif containing GTPase activating protein 3; log₂FC: 5).

Many of the genes we found to be downregulated in skull base chordoma are directly or indirectly involved in tumor suppression. The dominant nine genes were *STATH* (indirect tumor suppression by enabling a high concentration of intracellular calcium; log₂FC: -18); the *BPI* fold containing family members *BPIFB1* (log₂FC: -17), *BPIFB2* (log₂FC: -12), *BPIFB3* (log₂FC: -12), and *BPIFA1* (log₂FC: -17); *DMBT1* (deleted in malignant brain tumors 1; log₂FC: -14); *PIGR* (polymeric immunoglobulin receptor; log₂FC: -14); *MSMB*

Fig. 1 Morphologic chordoma subtypes: conventional (a–c), chondroid (d–f), and spindle type/dedifferentiated (g–i). Specimens were stained hematoxylin and eosin



(microseminoprotein, beta-; log₂FC: -13), and *MUC7* (mucin 7, secreted; log₂FC: -13).

Pathway analysis

IPA of the 294 gene transcripts revealed a pathway-network for embryonic, organismal, and tissue development that includes *T* acting on *SOX6* (Fig. 2c).

Biomarker analysis

To validate the gene expression patterns detected by RNA-Seq analysis, we used immunohistochemistry to assess the protein expression of five representative development-related genes highly upregulated in and considered important for skull base chordoma: *T*, *LMX1A*, *ZIC4*, *LHX4*, and *HOXA1*. These genes' corresponding gene log₂FCs were 14, 11, 7, and 6.

Of the 14 chordoma specimens with transcriptome data, 13 (81 % of 35 tumors tested) had *T* expression and 12 (68 %) had *HOXA1* expression. Respiratory mucosa and submucosal seromucinous glands within normal turbinates had weak positive *LMX1A*, *LHX4*, *ZIC4*, and *HOXA1* expression and no *T* expression; the cartilaginous component within turbinates lacked expression of *LMX1A*, *LHX4*, *ZIC4*, *HOXA1*, *T*. All 14 chordoma specimens with transcriptome analysis had *LMX1A*, *ZIC4*, and *LHX4* expression (78 and 81 % of 35 tumors tested (Fig. 3)). A detailed analysis of immunohistochemistry with various biomarkers is shown in Table 2.

Discussion

The present study demonstrates that RNA-Seq analysis of FFPE chordoma tissue is feasible and can provide insights into skull base chordoma progression.

Of 294 genes whose expression is characteristic of skull base chordoma, at least 222 have been associated with some other type of cancer (Supporting Table 3). This supports the view that skull base chordoma is in fact a cancer.

There is tremendous need for robust biomarkers to identify and predict which tumors are more likely to result in specific clinical outcomes to optimize treatment decisions.

Biomarker studies in chordoma in general (and in particular in skull base chordomas) have faced a challenge in which the number of specimens available, associated with very long-term follow-up (>8 years) is scarce.

Our RNA-Seq analysis identified a prominent group of genes that are involved in embryonic, organismal, and tissue development and whose expression is characteristic of chordoma (Supporting Table 2). From this group, we identified the five genes whose expression was most upregulated in chordoma compared with normal tissue and that may serve as biomarkers for the disease *T*, *LMX1A*, *ZIC4*, *LHX4*, and *HOXA1*.

T (*T* brachyury transcription factor, also known as brachyury, *TFT*, and *SAVA*) is an embryonic nuclear transcription factor that binds to a palindromic T-site on DNA. *T* is normally expressed in notochord-derived cells, where it affects the transcription of genes required for mesoderm

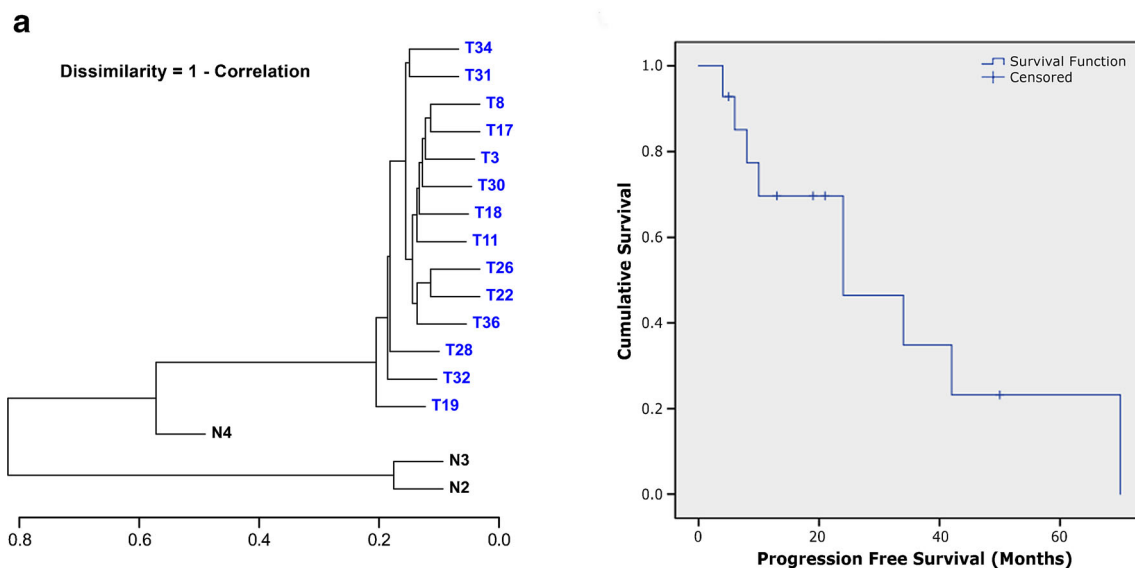


Fig. 2 **a** The cluster dendrogram shows the dissimilarities between neoplastic and normal tissue samples. **b** Kaplan-Meier curve of progression-free survival. **c** Heatmap of 100 genes with the highest differentially transcribed RNA transcript isoforms with dendrograms. *Red* indicates upregulation; *green* indicates downregulation. **c** IPA

revealed a pathway network that includes genes of differentially expressed RNA related to embryonic, organismal, and tissue development, including *T* brachyury acting on *SOX6*. *Red* indicates upregulation; *green* indicates downregulation

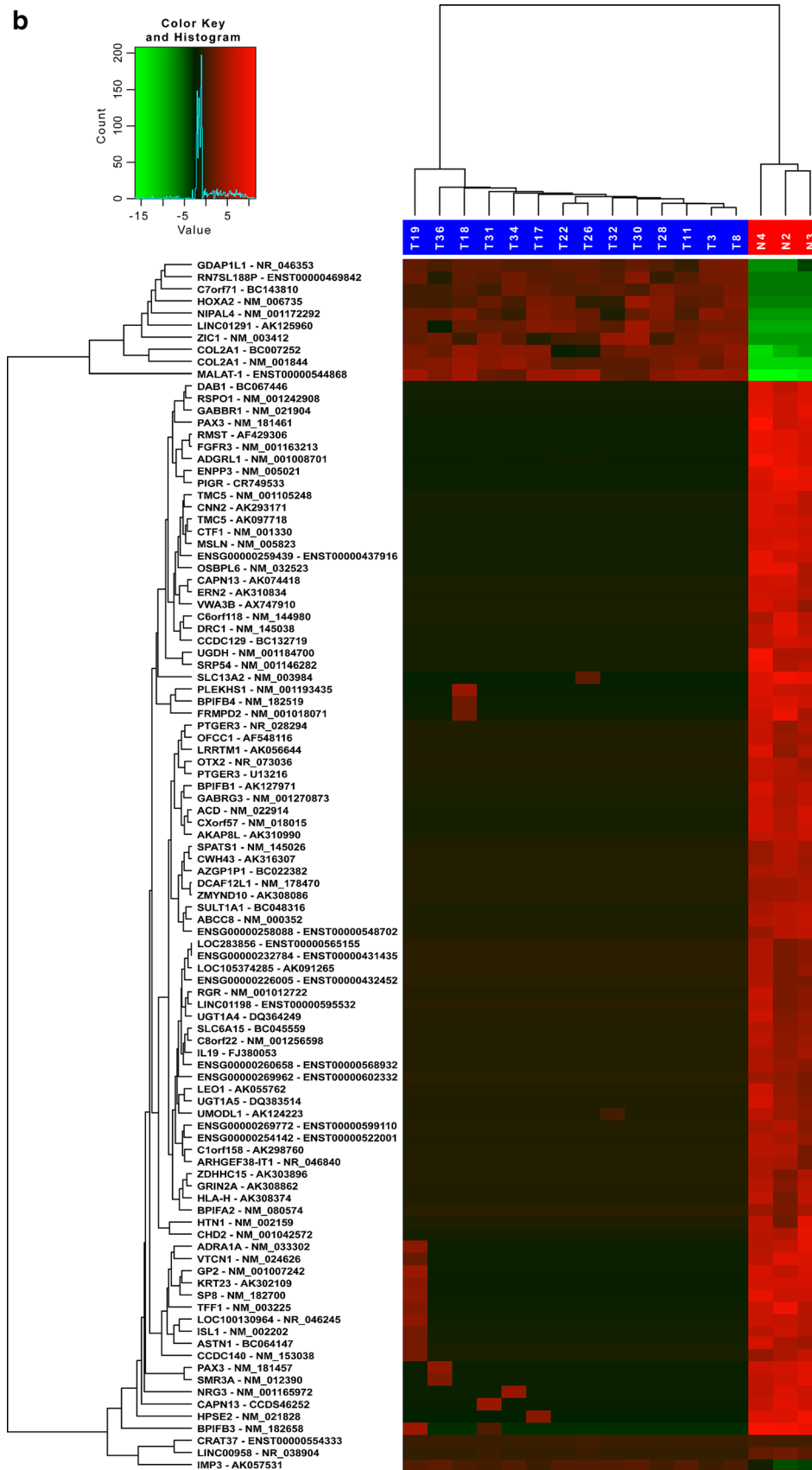


Fig. 2 (continued)

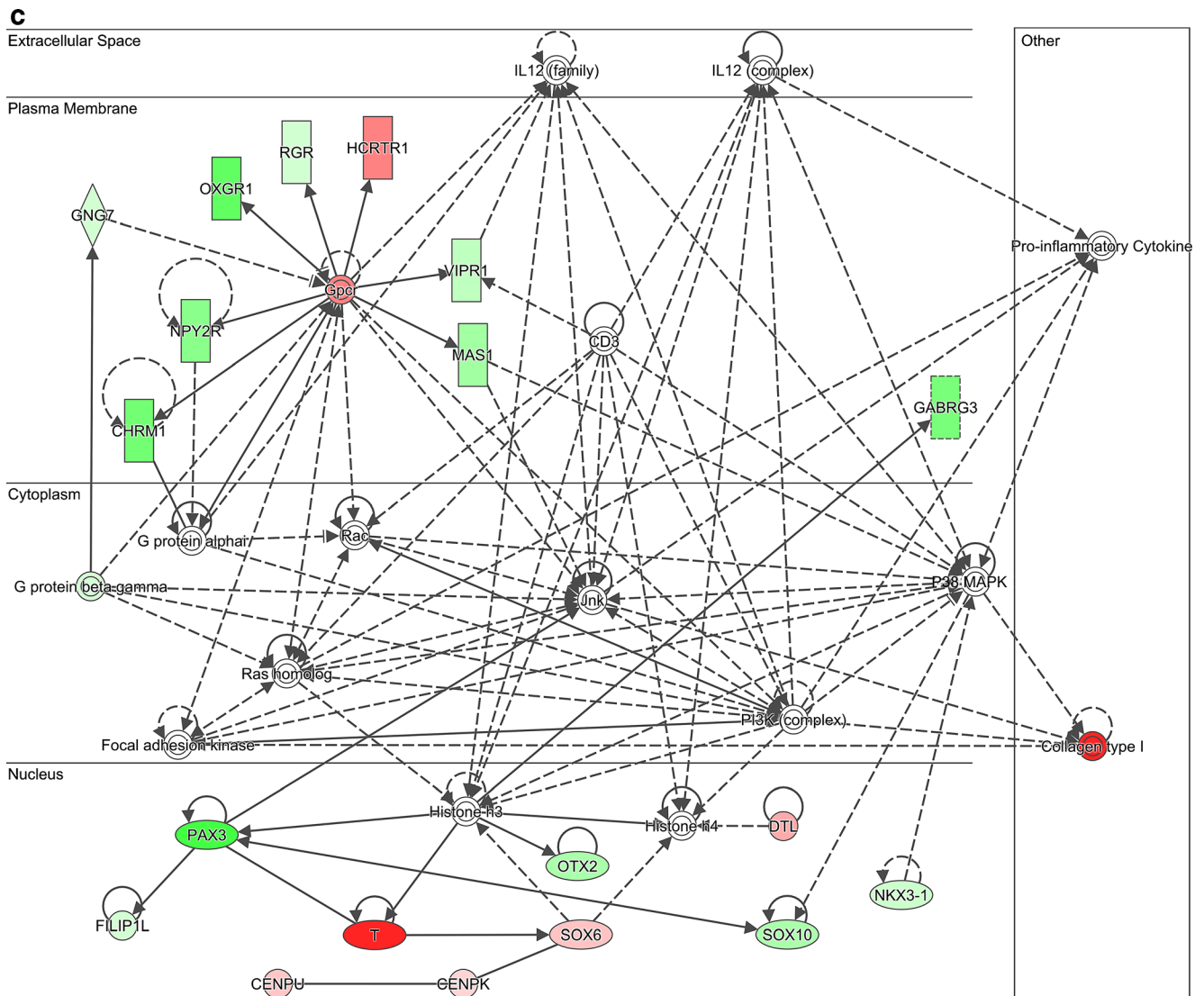


Fig. 2 (continued)

formation and differentiation. T is already used as diagnostic marker for the differential diagnosis of chordoma and hemangioblastoma versus neoplastic histological mimickers [14–17]. T's deregulated expression, its gene sequence variants, and its germline duplications are known to be involved in different types of cancer, including prostate cancer progression and aggressiveness [18], familial and sporadic chordoma [19], and poor prognosis-associated breast carcinomas [20].

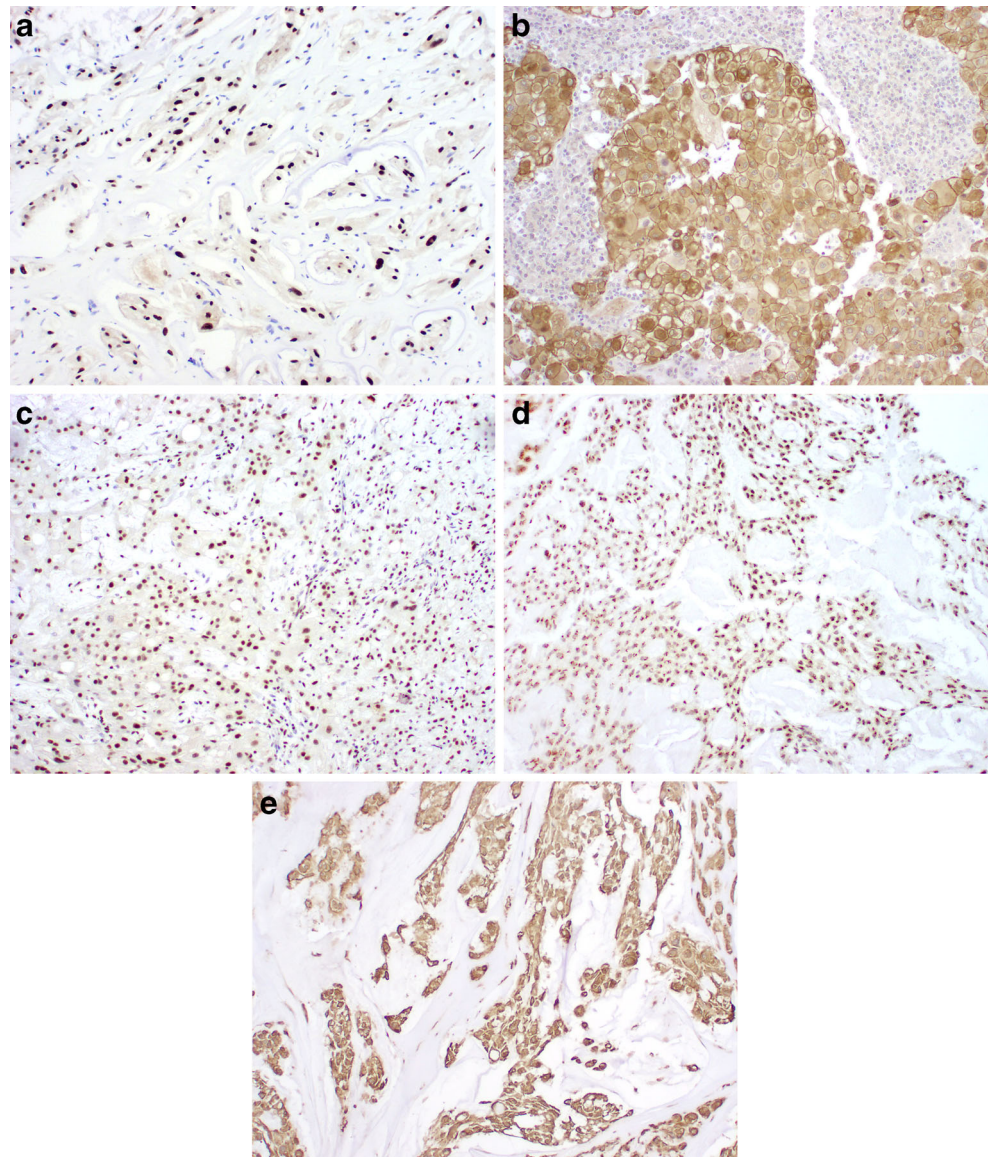
LMX1A (LIM homeobox transcription factor 1, alpha, also known as *LMX1.1*) is a homeodomain- and LIM domain-containing transcription factor that functions as an activator of insulin gene transcription and plays a role in the development of dopamine-producing neurons during embryogenesis. [21] *LMX1A* has also been reported to have a role in the differentiation of human embryonic stem cells into midbrain dopamine neurons in a model of Parkinson disease [22].

Allelic variation in the *LMX1A* gene influences the improvement of training-related working memory, and the hypermethylation-mediated reduction of *LMX1A* expression has been found in gastric cancer [23].

ZIC4 (Zic family member 4, also known as zinc finger protein of the cerebellum 4) is a member of the ZIC family of C2H2-type zinc finger proteins. Members of this family are important during development and have been associated with X-linked visceral heterotaxy and holoprosencephaly type 5. *ZIC4* is closely linked to *ZIC1*, a related family member also located on chromosome 3. Deregulated *ZIC4* expression has been found in patients with paraneoplastic neurologic disorders and small-cell lung cancer [24].

LHX4 (LIM homeobox 4, also known as *CPHD4*) encodes a member of a large protein family that contains the LIM domain, a unique cysteine-rich zinc-binding domain. The *LHX4* protein is a transcription factor involved in the control

Fig. 3 Expression of T brachyury (a), HOAX (b), LHX4 (c), LMX (d), and ZIC4 (e) in tumor tissue from chordoma patients



of the differentiation and development of the pituitary gland. Mutations in this gene can cause panhypopituitarism [25], combined pituitary hormone deficiency, and congenital hypopituitarism. In colorectal cancer, LHX4 has oncogenicity through the Wnt/ β -catenin/TCF4 cascade [26].

The class of homeobox genes is located in clusters named A, B, C, and D on four separate chromosomes and encodes transcription factors. Expression of the homeobox proteins is spatially and temporally regulated during embryonic development. The *HOXA1* gene homeobox A1 (aliases: BSAS, HOX1F) is part of the A cluster on chromosome 7 and encodes a transcription factor which regulates gene expression, morphogenesis, and differentiation. HOXA1 overexpression has been identified in oral squamous cell carcinoma and hepatocellular carcinoma, and its expression is correlated with poor prognosis [27, 28]. Furthermore, HOXA1 drives the growth and metastasis of melanoma and elicits an invasion

gene expression signature that prognosticates the clinical outcome [29].

Although this study is done on a small cohort of patients, the cluster dendrogram shows correlation from neoplastic samples versus normal samples, based on the log values of all expressed RNA, and patients with similar clinical outcome grouped together.

RNA-Seq may be too complex and expensive for clinical translation; thus, the development of TaqMan, digital PCR, or NanoString assays for the identification of chordoma may be more realistic. In addition, a percentage of FFPE tissue samples may yield RNA that fails to meet quality control criteria for RNA-Seq; however, whether alternative assays can be used to analyze specimens with more degraded RNA remains unknown. Generally applicable to other tumor types, additional validation of a biomarker panel using another independent,

expanded set of patient samples will increase the robustness of the specific panel and speed the translation of the biomarkers into a laboratory test for widespread clinical application. Such a test could be used to discriminate between aggressive and indolent disease, thereby

enabling the avoidance of unnecessary treatments and improving patients' treatment course.

In conclusion, the 294 differentially expressed gene transcripts we identified provide here, for the first time, a sufficiently large set to define skull base chordoma. They also provide a basis for the identification of novel and potentially important targets for the diagnosis, prognosis, and treatment of this cancer.

Table 2 Detailed immunohistochemical analysis of tissue samples with various biomarkers

Sample ID	Brachyury	LX4	LMX1A	HOXA1	ZIC4
1 T1	n/a	n/a	n/a	n/a	n/a
2 T8	Positive	Positive	Positive	Negative	Positive
3 T3	Positive	Positive	Positive	Positive	Positive
T5	Positive	Positive	Positive	Positive	Positive
4 T6	Positive	Positive	Positive	Positive	Positive
5 T7	Positive	Positive	Positive	Positive	Positive
6 T9	n/a	n/a	n/a	n/a	n/a
7 T10	n/a	n/a	n/a	n/a	n/a
8 T11	Positive	Positive	Positive	Positive	n/a
9 T12	n/a	n/a	n/a	n/a	n/a
10 T13	Positive	Positive	Positive	Positive	Positive
11 T15	Positive	n/a	Positive	Negative	Positive
12 T16	Positive	Positive	Positive	Positive	Positive
13 T17	Focally positive	Positive	Positive	Positive	Positive
14 T18	Positive	Positive	Positive	Positive	Positive
15 T19	Positive	Positive	Positive	Positive	Positive
16 T20	Positive	Positive	Positive	Positive	Positive
17 T22	Positive	Positive	Positive	Positive	Positive
18 T23	Positive	Negative	Positive	Negative	Positive
19 T25	Positive	Positive	Positive	Positive	Positive
20 T26	Positive	Positive	Positive	Positive	Positive
21 T27	n/a	n/a	n/a	n/a	n/a
22 T28	Positive	Positive	Positive	Positive	Positive
23 T29	Positive	Positive	Positive	Positive	n/a
24 T30	Positive	Positive	Positive	Positive	Positive
25 T31	Positive	Positive	Positive	Positive	Positive
26 T32	Negative	Positive	Positive	Negative	Positive
27 T33	Positive	Positive	Positive	Positive	Positive
28 T34	Positive	Positive	Positive	Positive	Positive
29 T35	n/a	n/a	n/a	n/a	n/a
30 T36	Positive	Positive	Positive	Positive	Positive
31 T37	n/a	n/a	n/a	n/a	n/a
32 T38	Positive	Positive	Positive	Positive	Positive
33 T39	Positive	Positive	Positive	Negative	Positive
34 T40	Positive	Positive	Positive	Positive	Positive
35 T42	Positive	Positive	Positive	Positive	Positive
36 T43	Positive	Positive	Positive	Positive	Positive
37 T44	Positive	Positive	Positive	Positive	n/a

n/a not available

Acknowledgments This study was supported by funds from the Mary Beth Pawelek Chair in Neurosurgery (FD) and by start-up funds from MDACC (DB).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

- Sun X, Hornicek F, Schwab JH (2015) Chordoma: an update on the pathophysiology and molecular mechanisms. *Curr Rev Musculoskelet Med* 8:344–352
- Lee J, Bhatia NN, Hoang BH, Ziogas A, Zell JA (2012) Analysis of prognostic factors for patients with chordoma with use of the California Cancer Registry. *J Bone Joint Surg Am* 94:356–363
- Mitchell A, Scheithauer BW, Unni KK, Forsyth PJ, Wold LE, McGivney DJ (1993) Chordoma and chondroid neoplasms of the sphenoid-occiput. An immunohistochemical study of 41 cases with prognostic and nosologic implications. *Cancer* 72:2943–2949
- Noel G, Habrand JL, Jauffret E, de Crevoisier R, Dederke S, Mammari H, et al. (2003) Radiation therapy for chordoma and chondrosarcoma of the skull base and the cervical spine. Prognostic factors and patterns of failure. *Strahlenther Onkol* 179:241–248
- Ruosi C, Colella G, Di Donato SL, Granata F, Di Salvatore MG, Fazioli F (2015) Surgical treatment of sacral chordoma: survival and prognostic factors. *Eur Spine J* 24(Suppl 7):912–917
- Thieblemont C, Biron P, Rocher F, Bouhour D, Bobin JY, Gerard JP, et al. (1995) Prognostic factors in chordoma: role of postoperative radiotherapy. *Eur J Cancer* 31A:2255–2259
- Varga PP, Szoverfi Z, Fisher CG, Boriani S, Gokaslan ZL, Dekutoski MB, et al. (2015) Surgical treatment of sacral chordoma: prognostic variables for local recurrence and overall survival. *Eur Spine J* 24:1092–1101
- von Witzleben A, Goertler LT, Marienfeld R, Barth H, Lechel A, Mellert K, et al. (2015) Preclinical characterization of novel chordoma cell systems and their targeting by pharmacological inhibitors of the CDK4/6 cell-cycle pathway. *Cancer Res* 75:3823–3831
- Scheipl S, Barnard M, Cottone L, Jorgensen M, Drewry DH, Zuercher WJ, et al. (2016) EGFR inhibitors identified as a potential treatment for chordoma in a focused compound screen. *J Pathol* 239(3):320–334
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63
- Mardis ER (2014) Sequencing the AML genome, transcriptome, and epigenome. *Semin Hematol* 51:250–258

12. Wolf JB (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-Seq tutorial. *Mol Ecol Resour* 13:559–572
13. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2013) STAR: ultrafast universal RNA-Seq aligner. *Bioinformatics* 29:15–21
14. Barresi V, Ieni A, Branca G, Tuccari G (2014) Brachyury: a diagnostic marker for the differential diagnosis of chordoma and hemangioblastoma versus neoplastic histological mimickers. *Dis Markers* 2014:514753
15. Faial T, Bernardo AS, Mendjan S, Diamanti E, Ortmann D, Gentsch GE, et al. (2015) Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. *Development* 142:2121–2135
16. Tirabosco R, Mangham DC, Rosenberg AE, Vujovic S, Bousdras K, Pizzolitto S, et al. (2008) Brachyury expression in extra-axial skeletal and soft tissue chordomas: a marker that distinguishes chordoma from mixed tumor/myoepithelioma/parachordoma in soft tissue. *Am J Surg Pathol* 32:572–580
17. Vujovic S, Henderson S, Presneau N, Odell E, Jacques TS, Tirabosco R, et al. (2006) Brachyury, a crucial regulator of notochordal development, is a novel biomarker for chordomas. *J Pathol* 209:157–165
18. Pinto F, Pertega-Gomes N, Pereira MS, Vizcaino JR, Monteiro P, Henrique RM, et al. (2014) T-box transcription factor brachyury is associated with prostate cancer progression and aggressiveness. *Clin Cancer Res* 20:4949–4961
19. Yang XR, Ng D, Alcorta DA, Liebsch NJ, Sheridan E, Li S, et al. (2009) T (brachyury) gene duplication confers major susceptibility to familial chordoma. *Nat Genet* 41:1176–1178
20. Palena C, Roselli M, Litzinger MT, Ferroni P, Costarelli L, Spila A, et al. (2014) Overexpression of the EMT driver brachyury in breast carcinomas: association with poor prognosis. *J Natl Cancer Inst* 106
21. Fathi A, Rasouli H, Yeganeh M, Salekdeh GH, Baharvand H (2015) Efficient differentiation of human embryonic stem cells toward dopaminergic neurons using recombinant LMX1A factor. *Mol Biotechnol* 57:184–194
22. Cai J, Donaldson A, Yang M, German MS, Enikolopov G, Iacovitti L (2009) The role of Lmx1a in the differentiation of human embryonic stem cells into midbrain dopamine neurons in culture and after transplantation into a Parkinson's disease model. *Stem Cells* 27:220–229
23. Dong W, Feng L, Xie Y, Zhang H, Wu Y (2011) Hypermethylation-mediated reduction of LMX1A expression in gastric cancer. *Cancer Sci* 102:361–366
24. Batailler L, Wade DF, Graus F, Stacey HD, Rosenfeld MR, Dalmau J (2004) Antibodies to Zic4 in paraneoplastic neurologic disorders and small-cell lung cancer. *Neurology* 62:778–782
25. Filges I, Bischof-Renner A, Rothlisberger B, Potthoff C, Glanzmann R, Gunthard J, et al. (2012) Panhypopituitarism presenting as life-threatening heart failure caused by an inherited microdeletion in 1q25 including LHX4. *Pediatrics* 129:e529–e534
26. Cha N, Liu W, Yang N, Xie S, Gao Y, Chen X, et al. (2014) Oncogenicity of LHX4 in colorectal cancer through Wnt/beta-catenin/TCF4 cascade. *Tumour Biol* 35:10319–10324
27. Bitu CC, Destro MF, Carrera M, da Silva SD, Graner E, Kowalski LP, et al. (2012) HOXA1 is overexpressed in oral squamous cell carcinomas and its expression is correlated with poor prognosis. *BMC Cancer* 12:146
28. Zha TZ, BS H, HF Y, Tan YF, Zhang Y, Zhang K (2012) Overexpression of HOXA1 correlates with poor prognosis in patients with hepatocellular carcinoma. *Tumour Biol* 33:2125–2134
29. Wardwell-Ozgo J, Dogruluk T, Gifford A, Zhang Y, Heffernan TP, van Doorn R, et al. (2014) HOXA1 drives melanoma tumor growth and metastasis and elicits an invasion gene expression signature that prognosticates clinical outcome. *Oncogene* 33:1017–1026