Gene expression profiling of fast- and slow-growing non-functioning gonadotroph pituitary adenomas

Camilla Maria Falch[1,2,3,4], **Arvind Y M Sundaram**[5], **Kristin Astrid Øystese**[1,6], **Kjersti Ringvoll Normann**[1,2,6], **Tove Lekva**[2], **Ivars Silamikelis**[7], **Alexander Kirkeby Eieland**[1], **Marianne Andersen**[3], **Jens Bollerslev**[1,6] and **Nicoleta Cristina Olarescu**[1,2]

[1]Section of Specialized Endocrinology, Department of Endocrinology, [2]Research Institute for Internal Medicine, Oslo University Hospital, Oslo, Norway, [3]Department of Endocrinology and Metabolism, Odense University Hospital, Odense, Denmark, [4]University of Southern Denmark, Odense, Denmark, [5]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway, [6]Faculty of Medicine, University of Oslo, Oslo, Norway, and [7]Latvian Biomedical Research and Study Centre (LV BMC), Riga, Latvia

Correspondence should be addressed to C M Falch
Email
cafal14@student.sdu.dk

Abstract

Objective: Reliable biomarkers associated with aggressiveness of non-functioning gonadotroph adenomas (GAs) are lacking. As the growth of tumor remnants is highly variable, molecular markers for growth potential prediction are necessary. We hypothesized that fast- and slow-growing GAs present different gene expression profiles and reliable biomarkers for tumor growth potential could be identified, focusing on the specific role of epithelial-mesenchymal transition (EMT).

Design and methods: Eight GAs selected for RNA sequencing were equally divided into fast- and slow-growing group by the tumor volume doubling time (TVDT) median (27.75 months). Data were analyzed by tophat2, cufflinks and cummeRbund pipeline. 40 genes were selected for RT-qPCR validation in 20 GAs based on significance, fold-change and pathway analyses. The effect of silencing *MTDH* (metadherin) and *EMCN* (endomucin) on *in vitro* migration of human adenoma cells was evaluated.

Results: 350 genes were significantly differentially expressed (282 genes upregulated and 68 downregulated in the fast group, *P*-adjusted <0.05). Among 40 selected genes, 11 showed associations with TVDT ($-0.669 < R < -0.46$, $P < 0.05$). These were *PCDH18, UNC5D, EMCN, MYO1B, GPM6A* and six EMT-related genes (*SPAG9, SKIL, MTDH, HOOK1, CNOT6L* and *PRKACB*). *MTDH*, but not *EMCN*, demonstrated involvement in cell migration and association with EMT markers.

Conclusions: Fast- and slow-growing GAs present different gene expression profiles, and genes related to EMT have higher expression in fast-growing tumors. In addition to *MTDH*, identified as an important contributor to aggressiveness, the other genes might represent markers for tumor growth potential and possible targets for drug therapy.

Introduction

The pathogenesis of sporadic non-functioning pituitary adenomas (NFPAs) remains elusive. Regardless of intensive effort, specific genetic mutations are rarely found in NFPAs (1). There are no reliable molecular markers associated with aggressiveness or recurrence, making the prediction of tumor behavior a demanding challenge. Consequently, prolonged observation is necessary, constituting a significant burden in terms of medical resources and societal costs.

The silent gonadotroph NFPAs (GAs) are the most common (75%) of all the immunohistochemical subgroups of NFPAs (2). Gene expression profiling analyses have shown to be of importance in identifying molecular markers, as well as to determine essential signaling pathways for correct histological stratification and thereby adequate treatment of a wide range of cancers (3). Recent studies have investigated the differential gene expression according to tumor invasiveness or recurrence characteristics (i.e. comparing invasive vs non-invasive (4, 5) and early recurrent vs non-recurrent pituitary adenomas (6)). Tumor volume doubling time (TVDT) describes the growth velocity by measuring the time it takes for a tumor to double in volume. It is calculated presuming the tumor has an exponential growth curve, enabling the clinician to estimate a safe interval between follow-up investigations and is used to predict the clinical course of tumors in several cancer types. Information of adenoma growth kinetics and TVDT, collaborated with data about tumor biology, would substantially enhance the clinical knowledge of tumor behavior and help to adjust and individualize follow-up (7). Previous studies usually describe the growth of NFPAs by an exponential, logistic or linear growth model (7, 8, 9, 10, 11). The presence of tumor remnants following initial surgery indicates a high risk of regrowth (12), and we have recently presented that the calculation of early TVDT as a prediction of 'worst-case' scenario for future growth can be a tool to tailor safe follow-up intervals (13).

Epithelial-mesenchymal transition (EMT) is a well-studied regulatory programming process in various cancers, which also plays a role in the pathogenesis of pituitary adenomas (14, 15, 16). Epithelial cells undergo changes to show mesenchymal cell characteristics with increased motility, invasiveness and even distant metastasis (14, 15, 16). Loss of membranous E-cadherin is a hallmark of EMT (14). Studies of growth hormone and ACTH-producing adenomas have linked EMT to the progression of tumors. The reduction in E-cadherin content in these tumors has been associated with tumor size, invasiveness and poor response to medical treatment (15, 17, 18, 19, 20).

In this study, we hypothesized that GAs with different initial growth velocities (high vs low) have different gene expression profiles. We aimed to identify target genes in fast- and slow-growing tumors, estimated by early postoperative TVDT, and to find reliable biomarkers predicting growth potential of the residual tumor tissue, focusing on the specific role of EMT.

Subjects and methods

Patients and samples

The patients included in this study ($n=20$) were selected from a larger cohort of patients ($n=88$) based on the possibility to assess growth modeling (only tumors with regrowth and four or more consecutive magnetic resonance imaging (MRI) examinations or five years of follow-up, not interrupted by new interventions), the availability of stored tumor tissue and the positivity for gonadotroph hormones (follicle-stimulating hormone (FSH) and/or luteinizing hormone (LH)) and/or steroidogenic factor (SF)-1 on immunohistochemistry (IHC). Tumor volume, early TVDT and growth models (logistic, exponential or linear) were estimated for all samples (13). The patients underwent surgery at Oslo University Hospital between 2002 and 2009, and all samples were from the primary operation. The tumor tissue was immediately placed on ice and snap frozen in liquid nitrogen, and later stored at −80°C or embedded in Tissue-Tek.

Written informed consent was obtained from all patients. The study was approved by the regional ethics committee (REK no: 2014/635 and REK no: 2014/1680) and hospital authority. The genetic expression analyses were exclusively connected to tumor biology. Thus, no information about the patient's genotype was given, and the study did not influence treatment or patient follow-up.

RNA isolation and reverse transcription

After homogenization of tumor tissue, total RNA was extracted using TRIzol (Invitrogen) and later purified with the QIAGEN miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions, including the DNase digestion step for removal of genomic DNA. RNA integrity was determined by Agilent 2100 Bioanalyzer (Agilent Technologies) and the concentrations were measured by OD readings on a NanoDrop ND-1000 Spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). All samples had RNA integrity numbers (RIN) >7, indicating an adequate quality. A high-capacity cDNA Reverse Transcription Kit (Applied Biosystems) was used to perform reverse transcription in a Labnet MultiGene Gradient Thermal Cylinder (Labnet International, Edison, NJ, USA) using 1 µg RNA in reaction. After the reaction, the cDNA was diluted 1:10 and all real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) measurements were performed with the same batch of cDNA.

RNA sequencing

Library preparation and sequencing

The RNA sequencing (RNA-seq) was performed using 2 µg RNA for each sample (*n* = 8) at the Norwegian Sequencing Centre, Oslo, Norway. Eight RNA-seq libraries with unique indexes were prepared using TruSeq stranded mRNA library prep kit (Illumina, San Diego, CA, USA) by following manufacturer's instructions. Libraries were pooled together, and 125 nt paired end sequencing was performed on two lanes of HiSeq 2500 (Illumina). RTA v1.18.66.3 was used for base calling and was further processed using bc2fastq v2.17.1.14 to demultiplex and generate fastq data based on the indexes used during library preparation.

Data processing

Pre-processing and cleanup   Fastq data from two lanes were concatenated for each sample before further analysis. Low-quality reads and adaptors were removed using Trimmomatic v0.33 (21) with recommended parameters. BBMap v34.56 (http://sourceforge.net/projects/bbmap) was used to remove reads aligning to PhiX (RefSeq: NC_001422.1), which was added as a spike-in during sequencing.

Transcriptome alignment   Cleaned data were aligned against the Human ensemble GRCh38 genome and the transcriptome using Tophat2, v2.0.13 (22) using '--library-type fr-firststrand --no-mixed --no-novel-juncs --transcriptome-index' as parameters. Library size was estimated using picard, v1.112 (http://broadinstitute.github.io/picard/) CollectInsertSizeMetrics tool after aligning the first one million reads using bowtie, v2.2.3 (23) to Human ensemble GRCh38 cDNA sequences, and the output was also provided as parameters for tophat2 alignment. Cuffdiff, v2.2.1 (24) pipeline was used to calculate the differential expression of the known genes described in ensemble GTF and CummeRbund, v2.14 (25). R package was used to visualize expression data, and custom scripts were used to create tables and graphs. Gene expression plots were created for both gene and isoform expression. Hierarchical clustering was performed using Pearson correlation with complete linkage in TM4 MeV (http://mev.tm4.org/).

RT-qPCR validation

To verify the expression of the transcripts of the selected genes, RT-qPCR was performed in the ABI 7900 (Applied Biosystems) on the entire cohort (*n* = 20). Reaction mix (with Power SYBR Green PCR Master Mix, Applied Biosystems) and samples were dispensed in the corresponding wells by an automated pipetting system (ep*Motion*5070 CB, Hamburg, Germany). RT-qPCR amplification was performed as previously described (26).

We used the primer bank (https://pga.mgh.harvard.edu/primerbank/) or searched the literature to find the primer pairs (Supplementary file 1, see section on supplementary data given at the end of this article), and tested their specificity using BLAST analysis (NCBI). The genomic DNA sequence (downloaded from www.ensembl.org) was used to study exon–intron borders by matching the primers to their location to avoid false-positive amplification due to DNA contamination. Primers that could detect the same isoforms as detected by RNA-seq were carefully selected. Primers' amplification efficiency and correlation coefficients were acquired from the slope of the standard curves. All RT-qPCR experiments were in accordance to the MIQE guidelines. Gene expression was quantified using the delta-delta Ct (ΔΔCt) method. The expression was normalized to the index obtained from the geometric mean of *GAPDH* and *ALAS1* Ct levels, as previously shown to be some of the most stable reference genes in NFPAs, and expressed as relative mRNA levels (26).

In vitro experiments

Cell cultures

Primary cells were obtained from three patients following a previously described protocol (27). Briefly, tumor tissue from the primary operation was collected in the operation room in L-15 (Leibovitz) medium (Lonza, cat. no. 12-700F, Verviers, Belgium) and transported on ice. The blood was removed immediately and the tumor cut in small pieces. After washing, the pellet was enzymatically dispersed with papain 13.2 U/mL (Worthington Biochemical Corporation, cat. no. LS003118, Lakewood, NJ, USA) for 3 min at 37°C. The papain was stopped by adding human albumin and removed by washing, and the cells were filtered through 100 µm cell strainer (BD Falcon), counted and plated at ca. 10^5 cells/cm² in Nunc Flasks (Thermo Scientific cat. no. 156340). The attached cell fraction was cultured in DMEM/F12 (GIBCO, cat. no. 31331-028) supplemented with 10 mM Hepes Buffer, 1% fetal calf serum, penicillin/streptomycin (50 U/mL and 50 µg/mL), 2.5 µg/mL heparin and 1× B27 w/o vitamin A. The culture was fed two times a week with 10 ng/mL bFGF and 20 ng/mL TGFa (R&D Systems, cat. no. 233-FB

and cat. no. 239-A), and the media was replaced every 10 days until the cells were ready for passaging. The cells were between the 2nd and 9th passages. Our culture conditions select the growth of pituitary adenoma cells, and not fibroblasts. At the mRNA level, the cells showed strong positive expression of *SF-1* and low expression of the fibroblast gene collagen type I alpha 2 chain (*COL1A2*) (Supplementary File 6). The culture media should favor the growth of neural stem cells due to low FCS and presence of B27 w/o vitamin A. Moreover, in the beginning of the culturing, the cells grew in colonies as a result from the division of single cells, a typical characteristic of epithelial cells (Supplementary file 7). Lastly, we observed that after a few passages, even if the culture was heterogeneous just after the isolation, it became more homogenous and none of the cells had the classical fibroblast appearance. All *in vitro* studies were performed with 3–4 technical replicates/patient, normalized and the mean of each patient were pooled together to perform data analysis.

Silencing and RT-qPCR

After the cells reached 70–80% confluence, they were transfected with the Lipofectamine RNAiMAX (ThermoFisher Scientific, cat. no. 13778075) and incubated for two or three days before performing the experiments. siRNA (Ambion Silencer Select Pre-designed siRNA, ThermoFisher Scientific) specific for human *MTDH*, *EMCN*, *GAPDH* (as a positive control, Supplementary file 8) and scramble control was used. The sense strand sequence was 5′-GAAUCUCCCAAACAAAUAAtt-3′ for *MTDH* and 5′-GCUGUUUCAACAUUACAAAtt-3′ for *EMCN*. The silencing was validated by RT-qPCR and by Western blot (WB) and IHC (Supplementary file 3). Cytotoxicity was measured using Cytotoxicity Detection Kit (Sigma-Aldrich cat. no. 11644793001).

WB

Protein extraction reagents were from ThermoFisher Scientific. The cells, kept on ice, were washed with cold PBS, lysed using 100 μL/well M-PER (cat. no. 78501) mixed with Halt Protease Inhibitor Cocktail (100×) (cat. no. 78430), scraped, and the protein was collected. The amount was measured with Pierce BCA Protein Assay Kit (cat. no. 23225). Pierce Lane Marker Reducing Sample Buffer (cat. no. 39000), was added, and the samples were boiled at 95°C for 5 min. WB was performed with

20 μg protein per lane. The following antibodies were used: Lyric/Metadherin (D5Y8R) XP Rabbit mAb #14065 (1:2000) (*MTDH*) and *GAPDH* (14C10) Rabbit mAb (1:5000) both from Cell Signaling, and Anti-Endomucin antibody ab96315, Abcam (1:500) and Endomucin Polyclonal Antibody (PA5-21395), ThermoFisher Scientific (1:1000) (*EMCN*). Anti-rabbit IgG, HRP-linked Antibody (Cell Signaling) was used as secondary antibody (1:10.000). Protein levels were adjusted to *GAPDH* and analyzed with the Multi Gauge software (Fuji film, Tokyo, Japan).

Wound healing assay

Images were taken with Motic AE200, objective motic 4×, ocular motic N-WF 10×, mitocam 2300 after the scratch was made, as well as every 3rd to 12th hour depending on the observed migration rate, until the wound was completely healed or up to 33 hours when silencing was expected to lose its function. Cell migration (μm between wound borders) was measured using Fiji ImageJ (https://imagej.net/Fiji/Downloads).

Statistics

All measurements were checked for a normal distribution by visual methods (histograms and Q-Q plots). The data are presented as means ± S.E.M. or as medians and IQR for continuous variables and as counts for categorical data. Differences in patient demographics were analyzed by Mann–Whitney *U* test. A non-parametric univariate association analysis was used to quantify bivariate independent associations. One-way ANOVA and unpaired Student's sample *t* test (two-tailed) were used to evaluate the differences between silenced cells and scramble controls. To assess the differences in the wound-healing assay, repetitive measures model with Bonferroni correction was used. The analyses were performed using SPSS, version 21. A *P* value <0.05 was considered significant .

Results

Characterization of patient samples for RNA-seq and RT-qPCR analysis

The samples included in this study (*n* = 20) were GAs macroadenomas from the primary operation, and all except two samples showed tumor remnant at three months postoperatively. The samples were divided

Table 1 Clinical characteristics of patients.

	RT-qPCR (*n* = 20)	**RNA seq** (*n* = 8)*	**Fast** (*n* = 10)	**Slow** (*n* = 10)
Gender (female/male)	7/13	4/4	2/8	5/5
Age (years)	57 (20)	68 (21)	60 (37)	56 (17)
Tumor volume doubling time (TVDT) (months)	27.8 (21.6)	29.6 (73.9)	19.7 (10.8)	38.6 (55.0)
Invasive** (yes/no), *n* = 14	6/8	5/2	3/5	3/3
Growth model (logistic/linear/exponential)	10/5/5	5/1/2	7/2/1	3/3/4
Preoperative volume (cm³), *n* = 14	6.02 (4.97)	6.69 (2.41)	5.23 (4.94)	5.44 (3.72)
Growth velocity*** (fast/slow)	10/10	4/4	10/0	0/10

*Samples for RNA sequencing are included in samples for validation with RT-qPCR; **Invasiveness measured by Knosp–Steiner criteria; ***Growth velocity divided into fast- or slow-growing tumors by postoperative TVDT median.

into a fast- and slow-growing group, respectively, by the TVDT median (27.75 months). Eight samples (four fast and four slow) were selected for RNA-seq. Selected genes from the sequencing analysis were validated with RT-qPCR on the entire cohort. Table 1 provides a descriptive baseline data of the samples used in both RT-qPCR and RNA-seq studies and of the fast- and slow-growing tumors. The patients with sequenced tumors were older (68 (21) vs 54 (12), (P = 0.02)) and had more invasive tumors (5 vs 1 patients showing a Knosp–Steiner score ≥3 on the left/right side). There was no difference in gender, preoperative volume, initial TVDT or growth model between sequenced and the remaining samples.

RNA-seq

RNA-seq analysis identified 350 significantly differentially expressed genes (P-adjusted <0.05) between the fast and the slow group (Supplementary file 2): 282 genes were upregulated in the fast group (fold-change between 1.9 and 32.0), and 68 genes were upregulated in the slow group (fold-change between 1.9 and 18.4) (Fig. 1A). 40 genes were selected for validation and investigation with RT-qPCR based on their significance (P-adjusted <0.05), fold-change, expression plots in each sample (Supplementary file 4) and pathway analyses (g:Profiler, http://biit.cs.ut.ee/gprofiler/). There were no significantly differentially regulated pathways between the fast and the slow group; therefore, we also based the selection on the literature search with the main focus on cancer- and EMT-related pathways (Fig. 1B and C). Of these, many genes were involved in cell adhesion (*UACA*, *LGALS3BP*, *PXN*, *ADGRB2*, *FBLN1*, *PCDH18*, *EMCN*, *CD44*, *MTDH*, *ROCK2*, *UNC5D* and *ATP1B2*), cytoskeleton (*HOOK1*, *MTDH*, *ROCK2* and *FNBP1L*) and migration (*GPM6A*, *MYO1B*, *CD44*, *UNC5D*, *MDK* and *FBLN1*). 17 genes were related to EMT (Fig. 1C).

RT-qPCR

Figure 2A shows the associations between selected genes and initial TVDT, pre- and post- operative tumor volume and different imagistic invasiveness parameters in the cohort of GAs.

11 of the 40 genes showed significant negative associations with TVDT ($-0.669 ≤ R ≥ -0.466$, $P < 0.05$) (Fig. 2B). These were *PCDH18*, *UNC5D*, *EMCN*, *MYO1B*, *GPM6A* and six genes known as related to EMT (*SPAG9*, *SKIL*, *MTDH*, *HOOK1*, *CNOT6L* and *PRKACB*).

Preoperative volume was associated with postoperative volume ($R = 0.747$, $P = 0.003$). The only gene associated with preoperative volume was *FOS* ($R = 0.609$, $P = 0.021$). There was no significant association between selected genes and gender or age, with *ANXA11* as the only exception correlating with age ($R = -0.473$, $P = 0.035$). Postoperative tumor volume was negatively associated with *HOOK1* ($R = -0.525$, $P = 0.025$) and *CNOT6L* ($R = -0.633$, $P = 0.005$), in addition to a positive association with *FOS* ($R = 0.556$, $P = 0.020$). Of these, only *CNOT6L* was also associated negatively with postoperative invasiveness. Some of the selected genes (Fig. 2A) showed a positive association with growth models, meaning that they exhibited a low expression in the exponential, an intermediate expression in the linear and a high expression in the logistic growth. We found positive associations with the preoperative invasiveness for *ANXA11*, *LRCH2*, *KIF5B* and *FNBP1L* and negative associations with *MDK*, *LGALSSBP*, *WFDC2*, *FBLN1*, *P3H3* and *BSG*. Among these, only *LRCH2* and *KIF5B* (positively), and *P3H3* (negatively), remained associated with postoperative invasiveness. There was no difference in gene expression between invasive and non-invasive tumors, with the exception of LRCH2 that was higher expressed in the invasive tumors. Between invasiveness and pre- and post-operative volume no significant association was found. There was no association between the initial TVDT, *MTDH*, *EMCN* and *MKI67* (Ki-67) and *TP53* (tumor protein p53) gene expression (data not shown).
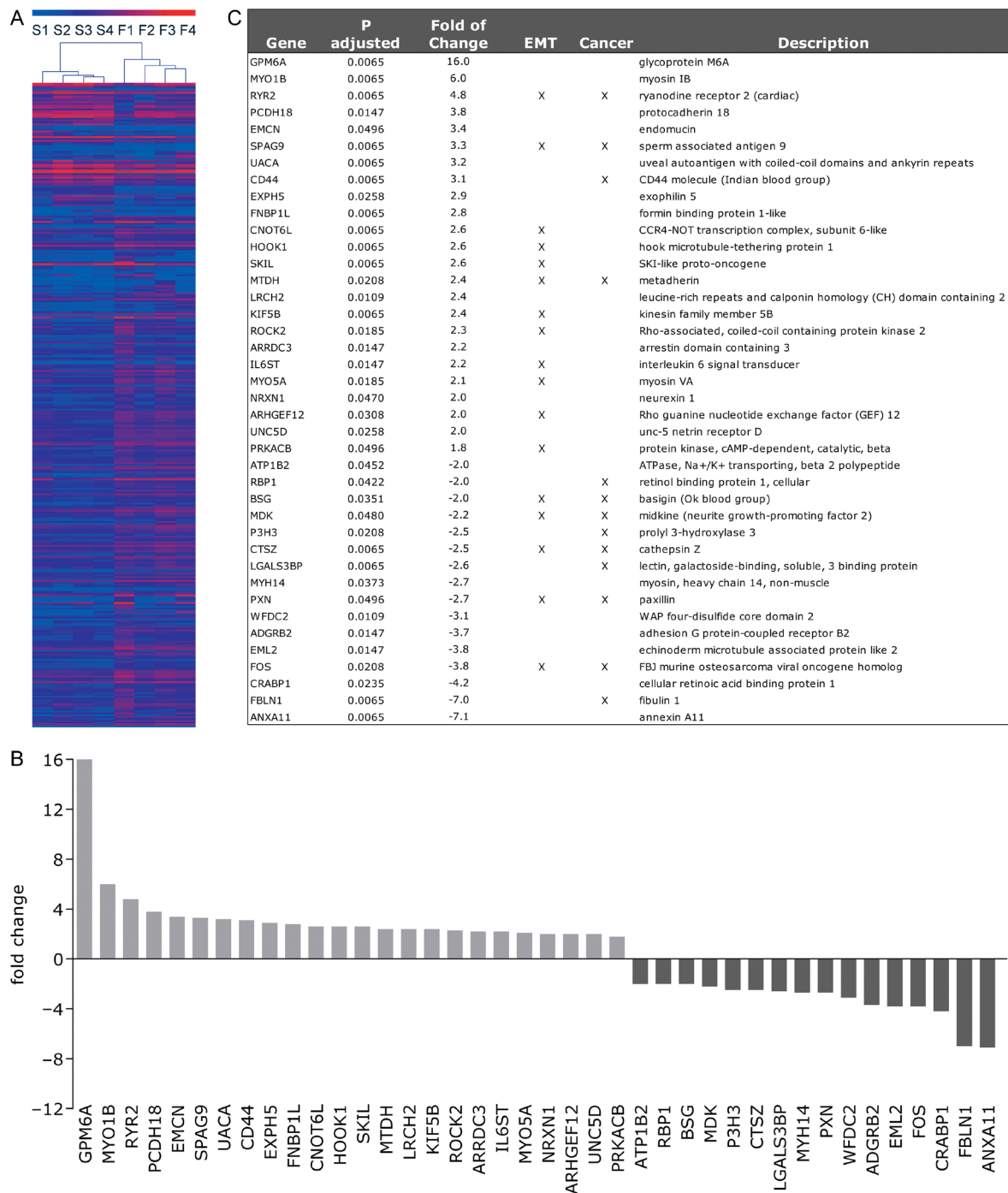
C

Gene	P adjusted	Fold of Change	EMT	Cancer	Description
GPM6A	0.0065	16.0			glycoprotein M6A
MYO1B	0.0065	6.0			myosin IB
RYR2	0.0065	4.8	X	X	ryanodine receptor 2 (cardiac)
PCDH18	0.0147	3.8			protocadherin 18
EMCN	0.0496	3.4			endomucin
SPAG9	0.0065	3.3	X	X	sperm associated antigen 9
UACA	0.0065	3.2			uveal autoantigen with coiled-coil domains and ankyrin repeats
CD44	0.0065	3.1		X	CD44 molecule (Indian blood group)
EXPH5	0.0258	2.9			exophilin 5
FNBP1L	0.0065	2.8			formin binding protein 1-like
CNOT6L	0.0065	2.6	X		CCR4-NOT transcription complex, subunit 6-like
HOOK1	0.0065	2.6	X		hook microtubule-tethering protein 1
SKIL	0.0065	2.6	X		SKI-like proto-oncogene
MTDH	0.0208	2.4	X	X	metadherin
LRCH2	0.0109	2.4			leucine-rich repeats and calponin homology (CH) domain containing 2
KIF5B	0.0065	2.4	X		kinesin family member 5B
ROCK2	0.0185	2.3	X		Rho-associated, coiled-coil containing protein kinase 2
ARRDC3	0.0147	2.2			arrestin domain containing 3
IL6ST	0.0147	2.2	X		interleukin 6 signal transducer
MYO5A	0.0185	2.1	X		myosin VA
NRXN1	0.0470	2.0			neurexin 1
ARHGEF12	0.0308	2.0	X		Rho guanine nucleotide exchange factor (GEF) 12
UNC5D	0.0258	2.0			unc-5 netrin receptor D
PRKACB	0.0496	1.8	X		protein kinase, cAMP-dependent, catalytic, beta
ATP1B2	0.0452	-2.0			ATPase, Na+/K+ transporting, beta 2 polypeptide
RBP1	0.0422	-2.0		X	retinol binding protein 1, cellular
BSG	0.0351	-2.0	X	X	basigin (Ok blood group)
MDK	0.0480	-2.2	X	X	midkine (neurite growth-promoting factor 2)
P3H3	0.0208	-2.5		X	prolyl 3-hydroxylase 3
CTSZ	0.0065	-2.5	X	X	cathepsin Z
LGALS3BP	0.0065	-2.6		X	lectin, galactoside-binding, soluble, 3 binding protein
MYH14	0.0373	-2.7			myosin, heavy chain 14, non-muscle
PXN	0.0496	-2.7	X	X	paxillin
WFDC2	0.0109	-3.1			WAP four-disulfide core domain 2
ADGRB2	0.0147	-3.7			adhesion G protein-coupled receptor B2
EML2	0.0147	-3.8			echinoderm microtubule associated protein like 2
FOS	0.0208	-3.8	X	X	FBJ murine osteosarcoma viral oncogene homolog
CRABP1	0.0235	-4.2			cellular retinoic acid binding protein 1
FBLN1	0.0065	-7.0		X	fibulin 1
ANXA11	0.0065	-7.1			annexin A11

Figure 1

Significantly differentially expressed genes between fast and slow non-functioning gonadotroph adenomas by RNA-seq. (A) Heatmap showing gene expression profiles clustered across samples and genes. Gene expression levels for each adenoma ($n=8$, S=slow, F=fast) are presented in horizontal rows with colors indicating upregulated (red) or downregulated (blue) genes. A total of 350 differentially expressed genes, 282 genes upregulated and 68 genes downregulated in the fast group were found (*P*-adjusted <0.05). Each column represents a single gene. (B) Distribution of fold-change for selected genes ($n=40$) in fast vs slow non-functioning gonadotroph adenomas. Positive fold-change indicates higher gene expression in the fast group. Negative fold-change indicates higher gene expression in the slow group. (C) Epithelial-mesenchymal transition (EMT) and cancer involvement of selected genes ($n=40$). The symbol 'x' marks the EMT- and cancer-related genes respectively.

A

Gene	TVDT	Growth model (exponential / linear / logistic)	Preoperative volume	Preoperative invasiveness			Postoperative volume	Postoperative invasiveness		
	n = 20	n = 20	n = 14	n = 14			n = 18	n = 18		
				Superior	KNOSP right	KNOSP left		Superior	KNOSP right	KNOSP left
PCDH18	-.669**	.368	.068	.054	.116	.250	-.255	-.658**	-.224	-.010
SPAG9	-.627**	.576**	-.226	.113	-.451	.193	-.288	-.316	-.383	.253
SKIL	-.594**	.478*	.068	.061	-.096	.224	-.216	-.506*	-.383	.131
UNC5D	-.589**	.707***	-.046	-.258	.000	-.081	-.141	-.491*	-.256	.030
MTDH	-.568**	.127	-.073	.425	-.226	.290	-.026	.138	-.009	.133
EMCN	-.534*	.719***	.024	.012	-.205	-.400	-.046	-.229	-.347	.021
HOOK1	-.534*	.544*	-.319	-.209	-.007	-.055	-.525*	-.341	-.343	-.266
MYO1B	-.514*	.711***	.301	-.103	.014	-.159	-.086	-.455	-.287	.245
GPM6A	-.510*	.597**	-.007	-.049	.191	-.021	-.092	-.256	-.131	-.251
CNOT6L	-.487*	.392	-.464	-.037	-.253	.093	-.633**	-.483*	-.473*	-.016
PRKACB	-.466*	.433	-.174	.157	-.321	-.195	-.356	-.358	-.501*	.031
ADGRB2	.386	-.061	.116	-.327	.034	-.086	.086	-.276	-.031	.015
MDK	.334	-.274	-.244	-.528	.150	-.583*	.073	.269	.426	-.297
PXN	.256	-.237	.011	-.219	-.014	-.069	-.110	.004	.063	-.056
MYH14	.241	-.192	.081	-.007	-.137	-.333	.003	-.125	.040	-.021
LGALSSBP	.227	.282	.024	-.160	-.103	-.552*	-.131	-.304	.131	-.203
FOS #	.223	-.299	.609*	.319	0.000	.200	.556*	.298	-.160	.335
WFDC2	.194	.147	-.029	.101	-.075	-.581*	-.044	-.160	.131	-.176
EML2	.144	-.176	.007	-.410	-.103	-.455	.108	.146	.283	0.000
CTSZ	.141	.102	.305	.351	-.253	.045	.251	.239	-.101	-.117
ANXA11	.053	-.188	.099	.831***	-.417	-.024	-.022	-.357	-.042	.176
FBLN1 #	.047	.288	-.121	.055	-.315	-.602*	.142	-.116	-.134	-.037
P3H3	-.024	.343	-.389	.057	-.609*	-.269	-.467	-.488*	-.297	.141
ATP1B2	-.029	.200	-.112	.120	-.479	-.350	-.185	-.187	-.199	-.181
RBP1	-.045	.495*	-.134	-.256	-.164	-.114	-.290	-.568*	-.280	-.052
BSG	-.131	.094	-.222	-.130	-.007	-.716**	.022	-.187	.129	-.338
CRABP1	-.183	.323	-.244	-.349	.123	-.331	-.288	-.388	.002	-.255
CD44	-.244	.168	.174	.447	-.212	.486	-.280	-.466	-.216	.175
EXPH5	-.281	.409	-.007	-.182	-.068	.459	-.327	-.493*	-.510*	.173
ARRDC3	-.295	.164	.103	-.437	.383	.264	.137	-.091	.231	.052
LRCH2	-.311	.135	.253	.381	.096	.790**	-.267	-.627**	-.244	.162
NRXN1	-.313	.233	.235	.076	-.082	.105	.061	-.213	-.092	.086
ROCK2	-.340	.290	-.288	.170	-.308	.188	-.201	-.015	.003	.205
KIF5B	-.355	.417	-.033	.069	-.219	.547*	-.325	-.507*	-.418	.402
IL6ST	-.367	.143	.235	.152	-.075	.276	.024	-.329	-.139	.046
RYR2 ×	-.377	.704**	-.068	.017	-.219	.217	.088	-.098	-.306	.407
MYO5A	-.388	.339	-.231	.197	-.444	.093	-.036	.151	-.250	.196
FNBP1L	-.395	.172	.169	.260	0.000	.638*	-.273	-.257	-.087	.317
ARHGEF12	-.427	.352	-.103	.120	-.356	.455	-.340	-.326	-.430	.344
UACA	-.432	.454*	.279	.037	-.123	-.071	.230	.066	-.289	.135

Data is presented as Spearmans's correlation coefficients, R; *, .05< P ≥ .01; **, .01 < P ≥ .001; ***, P < .001; #, n=19; ×, n=18; Grey area denotes significant correlations.

B

PCDH18 — R = -0.669, P = 0.001; TVDT (months)
SPAG9 — R = -0.627, P = 0.003; TVDT (months)
SKIL — R = -0.594, P = 0.006; TVDT (months)
UNC5D — R = -0.589, P = 0.006; TVDT (months)
MTDH — R = -0.568, P = 0.009; TVDT (months)
EMCN — R = -0.534, P = 0.015; TVDT (months)
HOOK1 — R = -0.534, P = 0.015; TVDT (months)
MYO1B — R = -0.514, P = 0.020; TVDT (months)
GMP6A — R = -0.510, P = 0.022; TVDT (months)
CNOT6L — R = -0.487, P = 0.029; TVDT (months)
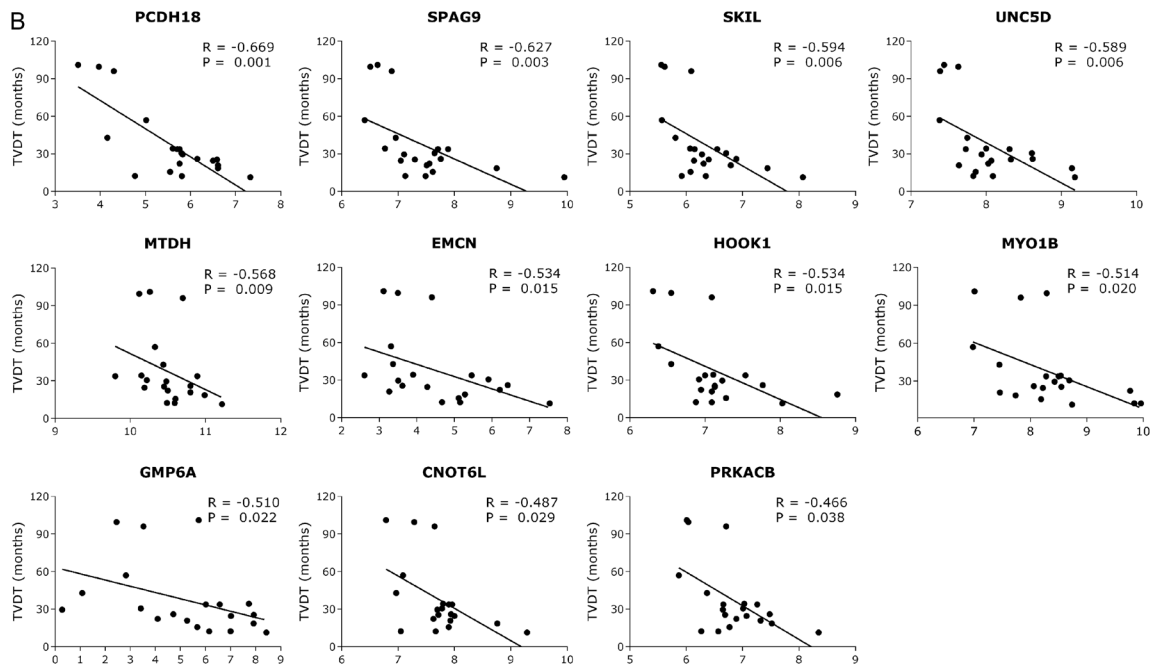**PRKACB** — R = -0.466, P = 0.038; TVDT (months)

Figure 2

RT-qPCR validation of 40 selected genes. (A) Associations between relative gene expression of the selected genes (n=40) and initial TVDT, growth models (exponential, linear and logistic), pre- and post- operative volume, and pre- and post- operative invasiveness in the non-functioning gonadotroph adenomas (n=20). Invasiveness was measured according to the Knosp–Steiner criteria (45), and also by the tumor's superior expansion (0=no superior growth, 1=upward convex bulging of the sella roof, 2=touching the optic chiasm, 3=lifting the optic chiasm, 4=blockage of interventricular foramina). (B) Scatter plots representing the associations between initial TVDT and significantly associated genes (n=11) in the non-functioning gonadotroph adenomas (n=20).

Silencing *MTDH* and *EMCN* in primary human adenoma cells

We chose two genes, *MTDH* and *EMCN*, for mechanistic *in vitro* studies. As depicted in Fig. 2B, both genes showed negative associations with TVDT, higher expression being associated with a lower TVDT (i.e. a more aggressive phenotype). *MTDH* mRNA gene expression was significantly reduced in *siMTDH* cells vs scramble controls (81%, $P < 0.001$) (Fig. 3A). A significant reduction in protein expression by WB (18%, $P = 0.001$) (Fig. 3C and D) and IHC (Supplementary file 3) was found. Similarly, a significant reduction in *EMCN* mRNA gene expression by 68% was observed in *siEMCN* cells ($P = 0.046$) (Fig. 3B). Due to the difficulties with the specific antibodies, we were unable to produce data on WB protein expression for EMCN. LDH cytotoxicity assay showed no significant difference between silenced and scramble control cells after two and three days (Supplementary file 5).

Migration assay

The wound-healing assay was performed to assess cell migration *in vitro*. It demonstrated that the *siMTDH* cells migrated at a slower rate in all three patients, even though there was a great difference in cells' behavior between patients (Fig. 4A). The only significant difference

Figure 3

Silencing of *MTDH* and *EMCN* in primary human adenoma cells. (A) *MTDH* mRNA levels in scramble control (scramble ctrl) vs silenced *MTDH* (*siMTDH*) cells. (B) *EMCN* mRNA levels in scramble control (scramble ctrl) vs silenced *EMCN* (*siEMCN*) cells. (C and D) *MTDH* protein expression by Western blot in scramble ctrl vs *siMTDH*. *GAPDH* was used as reference gene. Data are presented as mean ± s.e.m.; *n* = 3 patients, 3–4 technical replicates/patient; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

in cell migration for *siEMCN* cells was recorded after 6h in patient 73 (Fig. 4B). Lastly, we calculated the area under the curve (AUC) and found a significant difference between controls and *siMTDH* cells, but not for *siEMCN* cells (Fig. 4A and B).

E-cadherin and N-cadherin

We investigated if the gene expression of the EMT markers *CDH1* (E-cadherin) and *CDH2* (N-cadherin) would change between *siMTDH* cells and scramble controls (Fig. 5). In the *siMTDH* cells, the gene expression of *CDH1* was significantly higher ($P = 0.004$), no significant reduction in expression of *CDH2* was observed, but the *CDH1/CDH2* ratio was higher ($P = 0.039$). There was no significant difference between *siEMCN* cells and scramble controls in gene expression of the two EMT markers (data not showed).

Discussion

The present study demonstrates that fast- and slow-growing human GAs present different gene expression profiles. Of the 350 differently expressed genes, 40 were selected for further validation in a larger cohort of patients, but only 11 showed significant associations with initial TVDT and six of them were related to the EMT process. Some of the selected genes were also correlated to classical MRI invasion parameters, suggesting an additional role behind the involvement in the growth velocity. Furthermore, *MTDH*, but not *EMCN*, was involved in pituitary adenoma cell migration *in vitro*, making its coded protein, metadherin, a valuable biomarker for GAs tumor aggressiveness and regrowth, and a potential medical target.

Our study is the first to investigate genome-wide mRNA expression in fast- vs slow-growing GAs. The literature is sparse in this manner and to date two studies analyzed these tumors with some limitations. The first study presented differentially expressed long non-coding RNAs (lncRNA) and mRNAs in GAs vs normal pituitary tissues (28). In addition to the concern regarding the comparison of a tumor containing mainly of one cellular subtype (i.e. gonadotrophs) to the normal pituitary, comprising at least five different cell types, the authors chose to focus mostly on the lncRNA-mRNA co-expression network in the mTOR pathway, and the data were not evaluated in regards to markers of invasiveness and aggressiveness. The second study performed in multiple endocrine neoplasia X (MENX) affected rats by mRNA
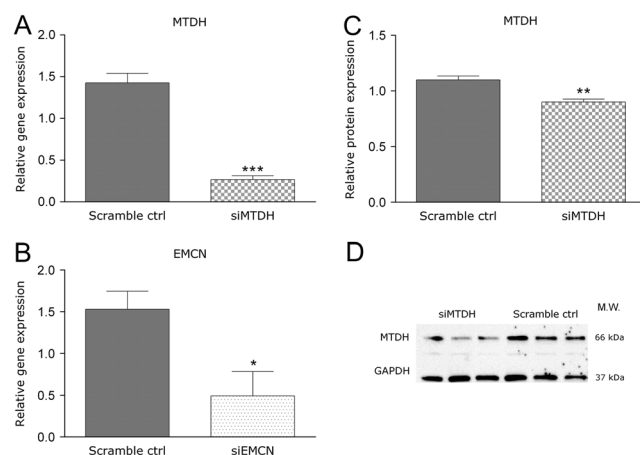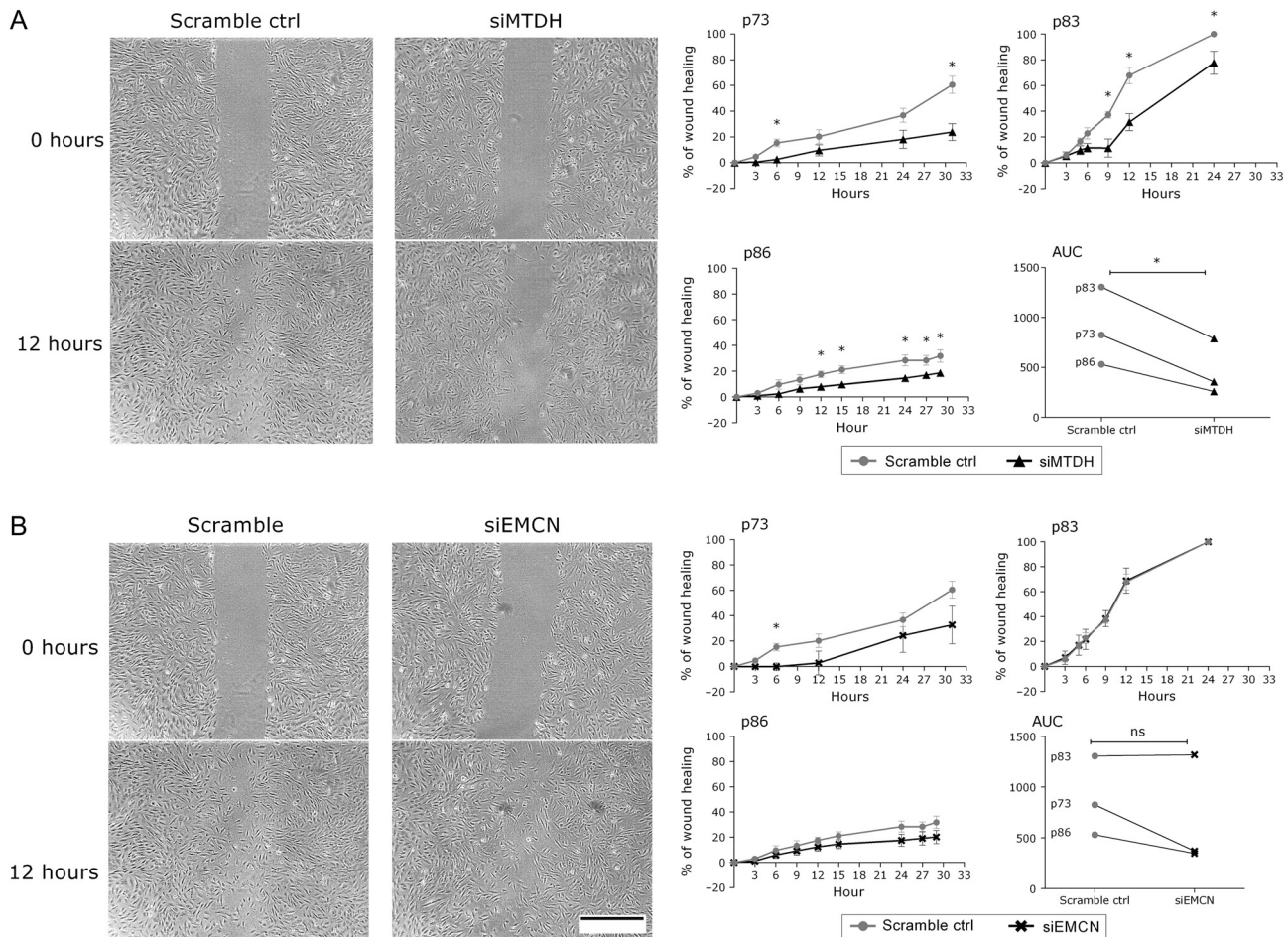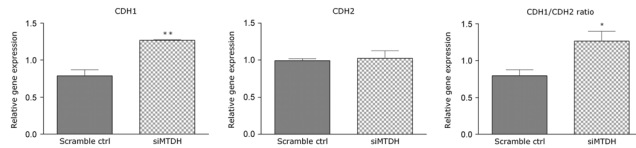
Figure 4

Wound-healing assay. (A) Wound-healing assay on silenced *MTDH* (*siMTDH*) cells. The pictures (cells are from p.83) illustrate *siMTDH* cells migrating at a slower rate than scramble control (scramble ctrl) cells. A significant difference was observed after 6 and 31h (*P*=0.03 and 0.021) in patient 73 (p73), after 9, 12 and 24h (*P*=0.013, 0.008 and 0.048) in patient 83 (p83) and after 12, 15, 24, 27 and 29h (*P*=0.04, 0.029, 0.025, 0.019 and 0.019) in patient 86 (p86). AUC shows a significant difference between *siMTDH* cells and scramble ctrl cells. (B) Wound-healing assay on silenced *EMCN* (*siEMCN*) cells. The pictures show *siEMCN* cells migrating at the same rate as scramble ctrl cells. The only significant difference in cell migration for silenced *EMCN* (*siEMCN*) cells was recorded after 6h (*P*=0.014) in p73. No significant difference was observed between *siEMCN* cells and scramble ctrl cells when calculating AUC. Data are presented as mean±s.e.m.; *n*=3 patients, 4 technical replicates/patient; *P≤0.05; **P≤0.01; ***P≤0.001; AUC, area under the curve. Scale bar, 500µm.

microarrays affymetrix, again comparing gonadotroph tumors and normal pituitary, identified *CYP11A1*, a gene regulated by the SF-1, as a promoter of proliferation/survival of gonadotroph cells both in rats and humans (29). However, we were unable to find *CYP11A1* as one of differentially expressed genes in our study. In addition, two recent studies described invasive-related candidate biomarkers by mRNA microarray in NFPAs, but no data on the histological subtype was presented making it difficult to directly compare the results to our study (5, 30).

The only study performing whole-exome DNA sequencing in seven sporadic GAs identified 24 somatic mutations in independent genes with no recurrent mutations. Moreover, DNA sequence analysis of these in the validation set did not reveal any mutations, thus indicating that these genes were unlikely to contribute in the etiology of sporadic GAs (31). Indeed, none of the somatic mutated genes were found to be differently expressed in our study, although our study was not designed to identify genes involved in tumorigenesis, but genes associated with aggressive growth behavior.

Figure 5

Epithelial-mesenchymal transition genes. mRNA levels of *CDH1* (E-cadherin) and *CDH2* (N-cadherin) and the ratio of *CDH1/CDH2* in silenced *MTDH* (*siMTDH*) and scramble control (scramble ctrl) cells. *GAPDH* was used as reference gene. Data are presented as mean ± s.e.m.; $n=3$ patients, 4 technical replicates/patient; *$P \le 0.05$; **$P \le 0.01$; ***$P \le 0.001$.

We mainly looked at the initial TVDT to distinguish between more and less aggressive GAs, based on the fact that initial TVDT might serve as an effective parameter of estimating early growth velocity that can predict the course of tumor remnant (13). Earlier studies compared invasive and non-invasive tumors based on their cavernous sinus invasion. However, differentiating radiologically between these two categories can be challenging, especially in the beginning of a tumors' invasive course. The growth rate as estimated by TVDT and the threatening of optic nerve and optic chiasm, rather than the lateral invasiveness, represents clinical criteria for tumor aggressiveness and the need for surgery. Despite their limitations, Ki-67 and p53 might be used as markers of aggressiveness and recurrence in pituitary adenomas (32, 33). However, our study did not find any associations between gene expression of *MKI67* and *TP53* and TVDT, *MTDH* and *EMCN*.

EMT is shown to be of significant importance in the aggressiveness and progression of somatotroph and corticotroph adenomas (15, 17, 18). However, using our selection criteria, we cannot conclude that EMT is the major component of the tumor growth in gonadotroph adenomas. For the *in vitro* study, we chose genes that code for proteins possible to target by medical treatment. *MTDH* is extensively studied in other cancers and its role as an oncogene is recognized (34). It has recently been shown that Bortezomib treatment can suppress *MTDH* in multiple myeloma (35). EMCN is a membrane protein that could, in theory, be easy to approach if proved to be a major actor in pituitary tumorigenesis (36).

MTDH, a known regulator of EMT (37), was significantly higher expressed in the fast-growing tumors, and the wound-healing assay demonstrated *MTDH* to be a contributor in cell migration, an important characteristic of mesenchymal cells. Moreover, silencing of *MTDH* led to a more epithelial and less aggressive profile, reassuring its role. Similarly, a study in human glioblastoma cells showed

that knockdown of *MTDH* increased the expression of E-cadherin and decreased the expression of N-cadherin and also when upregulating *MTDH* expression the cells showed overexpression of mesenchymal markers (38). Recent studies have shown that *MTDH* promotes tumor progression in many cancers types, including breast, prostate, glioma, esophageal and hepatocellular, through the integration of multiple signaling pathways, suggesting that *MTDH* may represent a potentially valuable target in cancer treatments (34).

EMCN did not show involvement in cell migration, nor association with EMT markers. It has recently been shown that *EMCN* prevents leukocyte-endothelial cell adhesion and has a critical role under resting and inflammatory conditions (36). The lack of effect on cell migration by *EMCN* could be explained by different means, such as: (1) *EMCN* does not have a role in cell migration of pituitary adenoma cells and our premises were incorrect, (2) technical challenges related to *in vitro* culturing did not allow us to adequately silence the *EMCN* gene such as its effect could be recorded, (3) other cells than pituitary adenoma cells (e.g. endothelial cells) may drive the pituitary adenoma aggressiveness and be responsible for *EMCN* expression at RNA-seq and RT-qPCR levels, (4) although there is an association with early TVDT, *EMCN* is not involved in tumor aggressiveness.

The remaining nine genes correlating significantly to the initial TVDT should be further studied, as they might represent useful markers for aggressiveness and tumor growth or targets for potential medical therapy. Three of these genes (*SKIL*, *CNOT6L* and *HOOK1*) code for proteins related to EMT suppression (39, 40, 41). However, our results somehow contradict their described functions, as the genes were higher expressed in the fast-growing tumors. Two other genes that definitely need further characterization are *GPM6A* (the highest positive fold of change on RNA-seq) and *PCDH18* (the highest association with TVDT). *GPM6A* is highly expressed in brain tissue and is involved in neuronal differentiation, including differentiation and migration of neuronal stem cells, neuronal plasticity and filopodia outgrowth and motility and probably synapse formation (42). *PCDH18* belongs to the protocadherin gene family and is thought to play a role in the establishment and function of specific cell–cell connections in the brain (43).

We observed that tumors with high *LRCH2* and *KIF5B* were more invasive, but also easier to remove since they have a negative association with the postoperative invasiveness. Although this is an interesting observation, it should be regarded with caution since other factors

such as perioperative conditions, the operative intention (debulking vs total removal) and patients' age, might have interfered.

When performing RNA-seq studies, a large amount of data is generated, and further investigations with mechanistic *in vitro* studies are of great value to gain a broader understanding of the importance of the identified genes in the tumor pathology. Accordingly, we carefully selected genes based on various relevant criteria to further test our hypothesis in a larger cohort of adenomas. Interestingly, just one-fourth of our selected genes showed associations with initial TVDT, supporting the RNA-seq data. Another fourth of the selected genes were close to reach statistical significance, suggesting that perhaps a larger cohort of GAs in the validation study would have resulted in a higher number of genes that were in agreement with the hypothesis. However, the rest of the selected genes did not show any association with initial TVDT, being in disagreement with RNA-seq data and emphasizing the need for validation of small RNA-seq studies in larger cohorts of patients. Furthermore, when two of the genes were tested *in vitro*, in a functional assay, in primary human adenoma cells, just one gene (i.e. *MTDH*) showed to have relevance.

The strengths of the present study are: (1) the thorough clinical, image and immunohistochemical characterization of the cohort, (2) the translational approach and (3) the use of human primary adenoma cells.

Our study has limitations. First, we might have chosen a selected set of tumors. The slowest growing tumors might have ended follow-up before four MRI exams or a five year follow-up was obtained, and the fastest growing tumors might have required re-intervention before the sufficient follow-up. However, these criteria were necessary to compute growth models and initial TVDT (13). Second, a limited number of samples were sequenced, and the *in vitro* studies were performed only in three patients. However, the validation study was performed in a larger cohort of GAs. Third, we have deliberately chosen to perform the present study in a well-characterized cohort of GAs and did not include other types of NFPAs. Clear evidence shows that the different NFPA subtypes have different clinical behavior and molecular profiles (44). Consequently, before extrapolating, further validation of our data should be performed in other subtypes of NFPAs. Fourth, it is not known to date, if the GAs molecular profile, and subsequently the tumor behavior, is similar before and after operation. In this study, initial TVDT was measured retrospectively in postoperative growing

remnants and might not accurately reflect the growth potential of the primary tumor. Although this may be regarded as a limitation, we consider that measuring TVDT, even on postoperative recordings, is a more precise approach in characterizing the growth potential of GA, than by the simpler, wider used criteria as macro/micro- or invasive/non-invasive adenomas. Lastly, although the present study shows *in vitro* mechanistic data for two selected proteins, we did not acquire protein data in the entire cohort of differentially expressed genes.

In conclusion, fast- and slow-growing tumors have different molecular profiles, but genes identified by RNA-seq need to be further validated in a larger cohort of patients, and with mechanistic studies. *MTDH* is a promising biomarker in fast-growing GAs and may be considered as a potential drug target.

References

1 Zhou Y, Zhang X & Klibanski A. Genetic and epigenetic mutations of tumor suppressive genes in sporadic pituitary adenoma. *Molecular and Cellular Endocrinology* 2014 **386** 16–33. (https://doi.org/10.1016/j.mce.2013.09.006)

2 Nishioka H, Inoshita N, Mete O, Asa SL, Hayashi K, Takeshita A, Fukuhara N, Yamaguchi-Okada M, Takeuchi Y & Yamada S. The complementary role of transcription factors in the accurate diagnosis of clinically nonfunctioning pituitary adenomas. *Endocrine Pathology* 2015 **26** 349–355. (https://doi.org/10.1007/s12022-015-9398-z)

3 Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, Leiserson MD, Niu B, McLellan MD, Uzunangelov V *et al*. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* 2014 **158** 929–944. (https://doi.org/10.1016/j.cell.2014.06.049)

4 Galland F, Lacroix L, Saulnier P, Dessen P, Meduri G, Bernier M, Gaillard S, Guibourdenche J, Fournier T, Evain-Brion D *et al*. Differential gene expression profiles of invasive and non-invasive non-functioning pituitary adenomas based on microarray analysis. *Endocrine-Related Cancer* 2010 **17** 361–371. (https://doi.org/10.1677/ERC-10-0018)

5 Yu SY, Hong LC, Feng J, Wu YT & Zhang YZ. Integrative proteomics and transcriptomics identify novel invasive-related biomarkers of non-functioning pituitary adenomas. *Tumor Biology* 2016 **37** 8923–8930. (https://doi.org/10.1007/s13277-015-4767-2)

6 Marko NF, Coughlan C & Weil RJ. Towards an integrated molecular and clinical strategy to predict early recurrence in surgically resected non-functional pituitary adenomas. *Journal of Clinical Neuroscience* 2012 **19** 1535–1540. (https://doi.org/10.1016/j.jocn.2012.01.038)

7 Honegger J, Zimmermann S, Psaras T, Petrick M, Mittelbronn M, Ernemann U, Reincke M & Dietz K. Growth modelling of non-functioning pituitary adenomas in patients referred for surgery. *European Journal of Endocrinology* 2008 **158** 287–294. (https://doi.org/10.1530/EJE-07-0502)

8 Balogun JA, Monsalves E, Juraschka K, Parvez K, Kucharczyk W, Mete O, Gentili F & Zadeh G. Null cell adenomas of the pituitary gland: an institutional review of their clinical imaging and behavioral characteristics. *Endocrine Pathology* 2015 **26** 63–70. (https://doi.org/10.1007/s12022-014-9347-2)

9 Hsu CY, Guo WY, Chien CP & Ho DM. MIB-1 labeling index correlated with magnetic resonance imaging detected tumor volume doubling time in pituitary adenoma. *European Journal of Endocrinology* 2010 **162** 1027–1033. (https://doi.org/10.1530/EJE-09-1100)

10 Monsalves E, Larjani S, Loyola Godoy B, Juraschka K, Carvalho F, Kucharczyk W, Kulkarni A, Mete O, Gentili F, Ezzat S *et al*. Growth patterns of pituitary adenomas and histopathological correlates. *Journal of Clinical Endocrinology and Metabolism* 2014 **99** 1330–1338. (https://doi.org/10.1210/jc.2013-3054)

11 Tanaka Y, Hongo K, Tada T, Sakai K, Kakizawa Y & Kobayashi S. Growth pattern and rate in residual nonfunctioning pituitary adenomas: correlations among tumor volume doubling time, patient age, and MIB-1 index. *Journal of Neurosurgery* 2003 **98** 359–365. (https://doi.org/10.3171/jns.2003.98.2.0359)

12 Chen Y, Wang CD, Su ZP, Chen YX, Cai L, Zhuge QC & Wu ZB. Natural history of postoperative nonfunctioning pituitary adenomas: a systematic review and meta-analysis. *Neuroendocrinology* 2012 **96** 333–342. (https://doi.org/10.1159/000339823)

13 Oystese KA, Zucknick M, Casar-Borota O, Ringstad G & Bollerslev J. Early postoperative growth in non-functioning pituitary adenomas; a tool to tailor safe follow-up. *Endocrine* 2017 **57** 35–45. (https://doi.org/10.1007/s12020-017-1314-5)

14 Guarino M, Rubino B & Ballabio G. The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* 2007 **39** 305–318. (https://doi.org/10.1080/00313020701329914)

15 Lekva T, Berg JP, Fougner SL, Olstad OK, Ueland T & Bollerslev J. Gene expression profiling identifies ESRP1 as a potential regulator of epithelial mesenchymal transition in somatotroph adenomas from a large cohort of patients with acromegaly. *Journal of Clinical Endocrinology and Metabolism* 2012 **97** E1506–E1514. (https://doi.org/10.1210/jc.2012-1760)

16 Thiery JP, Acloque H, Huang RY & Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell* 2009 **139** 871–890. (https://doi.org/10.1016/j.cell.2009.11.007)

17 Evang JA, Berg JP, Casar-Borota O, Lekva T, Kringen MK, Ramm-Pettersen J & Bollerslev J. Reduced levels of E-cadherin correlate with progression of corticotroph pituitary tumours. *Clinical Endocrinology* 2011 **75** 811–818. (https://doi.org/10.1111/j.1365-2265.2011.04109.x)

18 Fougner SL, Lekva T, Borota OC, Hald JK, Bollerslev J & Berg JP. The expression of E-cadherin in somatotroph pituitary adenomas is related to tumor size, invasiveness, and somatostatin analog response. *Journal of Clinical Endocrinology and Metabolism* 2010 **95** 2334–2342. (https://doi.org/10.1210/jc.2009-2197)

19 Lekva T, Berg JP, Lyle R, Heck A, Ringstad G, Olstad OK, Michelsen AE, Casar-Borota O, Bollerslev J & Ueland T. Epithelial splicing regulator protein 1 and alternative splicing in somatotroph adenomas. *Endocrinology* 2013 **154** 3331–3343. (https://doi.org/10.1210/en.2013-1051)

20 Lekva T, Berg JP, Heck A, Lyngvi Fougner S, Olstad OK, Ringstad G, Bollerslev J & Ueland T. Attenuated RORC expression in the presence of EMT progression in somatotroph adenomas following treatment with somatostatin analogs is associated with poor clinical recovery. *PLoS ONE* 2013 **8** e66927. (https://doi.org/10.1371/journal.pone.0066927)

21 Bolger AM, Lohse M & Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014 **30** 2114–2120. (https://doi.org/10.1093/bioinformatics/btu170)

22 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R & Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 2013 **14** R36. (https://doi.org/10.1186/gb-2013-14-4-r36)

23 Langmead B & Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 2012 **9** 357–359. (https://doi.org/10.1038/nmeth.1923)

24 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL & Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 2012 **7** 562–578. (https://doi.org/10.1038/nprot.2012.016)

25 Goff LA, Trapnell C & Kelley D. CummeRbund: visualization and exploration of cufflinks high-throughput sequencing data. 2014. (available at: http://bioconductor.org/packages/release/bioc/vignettes/cummeRbund/inst/doc/cummeRbund-manual.pdf).

26 Normann KR, Oystese KA, Berg JP, Lekva T, Berg-Johnsen J, Bollerslev J & Olarescu NC. Selection and validation of reliable reference genes for RT-qPCR analysis in a large cohort of pituitary adenomas. *Molecular and Cellular Endocrinology* 2016 **437** 183–189. (https://doi.org/10.1016/j.mce.2016.08.030)

27 Murrell W, Palmero E, Bianco J, Stangeland B, Joel M, Paulson L, Thiede B, Grieg Z, Ramsnes I, Skjellegrind HK *et al*. Expansion of multipotent stem cells from the adult human brain. *PLoS ONE* 2013 **8** e71334. (https://doi.org/10.1371/journal.pone.0071334)

28 Li J, Li C, Wang J, Song G, Zhao Z, Wang H, Wang W, Li H, Li Z, Miao Y *et al*. Genome-wide analysis of differentially expressed lncRNAs and mRNAs in primary gonadotrophin adenomas by RNA-seq. *Oncotarget* 2017 **8** 4585–4606. (https://doi.org/10.18632/oncotarget.13948)

29 Lee M, Marinoni I, Irmler M, Psaras T, Honegger JB, Beschorner R, Anastasov N, Beckers J, Theodoropoulou M, Roncaroli F *et al*. Transcriptome analysis of MENX-associated rat pituitary adenomas identifies novel molecular mechanisms involved in the pathogenesis of human pituitary gonadotroph adenomas. *Acta Neuropathologica* 2013 **126** 137–150. (https://doi.org/10.1007/s00401-013-1132-7)

30 Feng J, Yu SY, Li CZ, Li ZY & Zhang YZ. Integrative proteomics and transcriptomics revealed that activation of the IL-6R/JAK2/STAT3/MMP9 signaling pathway is correlated with invasion of pituitary

null cell adenomas. *Molecular and Cellular Endocrinology* 2016 **436** 195–203. (https://doi.org/10.1016/j.mce.2016.07.025)

31 Newey PJ, Nesbit MA, Rimmer AJ, Head RA, Gorvin CM, Attar M, Gregory L, Wass JA, Buck D, Karavitaki N *et al*. Whole-exome sequencing studies of nonfunctioning pituitary adenomas. *Journal of Clinical Endocrinology and Metabolism* 2013 **98** E796–E800. (https://doi.org/10.1210/jc.2012-4028)

32 Di Ieva A, Rotondo F, Syro LV, Cusimano MD & Kovacs K. Aggressive pituitary adenomas – diagnosis and emerging treatments. *Nature Reviews Endocrinology* 2014 **10** 423–435. (https://doi.org/10.1038/nrendo.2014.64)

33 Raverot G, Burman P, McCormack AI, Heaney AP, Petersenn S, Popovic V, Trouillas J & Dekkers O. European Society of Endocrinology clinical practice guidelines for the management of aggressive pituitary tumours and carcinomas. *European Journal of Endocrinology* 2018 **178** G1–G24. (https://doi.org/10.1530/EJE-17-0796)

34 Hu G, Wei Y & Kang Y. The multifaceted role of MTDH/AEG-1 in cancer progression. *Clinical Cancer Research* 2009 **15** 5615–5620. (https://doi.org/10.1158/1078-0432.CCR-09-0049)

35 Gu C, Feng L, Peng H, Yang H, Feng Z & Yang Y. MTDH is an oncogene in multiple myeloma, which is suppressed by Bortezomib treatment. *Oncotarget* 2016 **7** 4559–4569. (https://doi.org/10.18632/oncotarget.6610)

36 Zahr A, Alcaide P, Yang J, Jones A, Gregory M, dela Paz NG, Patel-Hett S, Nevers T, Koirala A, Luscinskas FW *et al*. Endomucin prevents leukocyte-endothelial cell adhesion and has a critical role under resting and inflammatory conditions. *Nature Communications* 2016 **7** 10363. (https://doi.org/10.1038/ncomms10363)

37 Wang Z, Tang ZY, Yin Z, Wei YB, Liu LF, Yan B, Zhou KQ, Nian YQ, Gao YL & Yang JR. Metadherin regulates epithelial-mesenchymal transition in carcinoma. *OncoTargets and Therapy* 2016 **9** 2429–2436. (https://doi.org/10.2147/OTT.S104556)

38 Park SY, Choi M, Park D, Jeong M, Ahn KS, Lee J, Fisher PB, Yun M & Lee SG. AEG-1 promotes mesenchymal transition through the activation of Rho GTPases in human glioblastoma cells. *Oncology Reports* 2016 **36** 2641–2646. (https://doi.org/10.3892/or.2016.5106)

39 Ikeuchi Y, Dadakhujaev S, Chandhoke AS, Huynh MA, Oldenborg A, Ikeuchi M, Deng L, Bennett EJ, Harper JW, Bonni A *et al*. TIF1gamma protein regulates epithelial-mesenchymal transition by operating as a small ubiquitin-like modifier (SUMO) E3 ligase for the transcriptional regulator SnoN1. *Journal of Biological Chemistry* 2014 **289** 25067–25078. (https://doi.org/10.1074/jbc.M114.575878)

40 Li S, Wang L, Zhao Q, Liu Y, He L, Xu Q, Sun X, Teng L, Cheng H & Ke Y. SHP2 positively regulates TGFbeta1-induced epithelial-mesenchymal transition modulated by its novel interacting protein Hook1. *Journal of Biological Chemistry* 2014 **289** 34152–34160. (https://doi.org/10.1074/jbc.M113.546077)

41 Zhou W & Thiery JP. Loss of Git2 induces epithelial-mesenchymal transition by miR146a-Cnot6L-controlled expression of Zeb1. *Journal of Cell Science* 2013 **126** 2740–2746. (https://doi.org/10.1242/jcs.126367)

42 Alvarez Julia A, Frasch AC & Fuchsova B. Neuronal filopodium formation induced by the membrane glycoprotein M6a (Gpm6a) is facilitated by coronin-1a, Rac1, and p21-activated kinase 1 (Pak1). *Journal of Neurochemistry* 2016 **137** 46–61. (https://doi.org/10.1111/jnc.13552)

43 Kim SY, Yasuda S, Tanaka H, Yamagata K & Kim H. Non-clustered protocadherin. *Cell Adhesion and Migration* 2011 **5** 97–105. (https://doi.org/10.4161/cam.5.2.14374)

44 Raverot G, Jouanneau E & Trouillas J. Management of endocrine disease: clinicopathological classification and molecular markers of pituitary tumours for personalized therapeutic strategies. *European Journal of Endocrinology* 2014 **170** R121–R132. (https://doi.org/10.1530/EJE-13-1031)

45 Knosp E, Steiner E, Kitz K & Matula C. Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings. *Neurosurgery* 1993 **33** 610–617. (https://doi.org/10.1227/00006123-199310000-00008)