Journal of Steroid Biochemistry & Molecular Biology 125 (2011) 231–242

Contents lists available at ScienceDirect

Journal of Steroid Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/jsbmb

Novel estrogen-related genes and potential biomarkers of ovarian endometriosis identified by differential expression analysis

Katja Vouk[a], Tina Šmuc[a], Christina Guggenberger[b], Martina Ribič-Pucelj[c], Jasna Šinkovec[c], Bettina Husen[d], Hubert Thole[d], Pieter Houba[d], Claudia Thaete[d], Jerzy Adamski[b,e], Tea Lanišnik Rižner[a,*]

[a] Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg, 2, 1000 Ljubljana, Slovenia
[b] Helmholtz Zentrum München, German Research Centre for Environment and Health, Institute of Experimental Genetics, Genome Analysis Centre, Neuherberg, Germany
[c] Department of Obstetrics and Gynecology, University Medical Centre, Ljubljana, Slovenia
[d] Solvay Pharmaceuticals Research Laboratories, Hannover, Germany, and Weesp, The Netherlands
[e] Institute of Experimental Genetics, Life and Food Science Center Weihenstephan, Technische Universität München, Freising-Weihenstephan, Germany

ARTICLE INFO

Article history:
Received 21 October 2010
Received in revised form 28 February 2011
Accepted 4 March 2011

Keywords:
Steroids
Biglycan
Low-density arrays
Gene expression
Diagnostic markers
Drug targets
Pathway analysis

ABSTRACT

In the search for novel biomarkers of endometriosis, we selected 152 genes from the GeneLogic database based on results of genome-wide expression analysis of ovarian endometriosis, plus 20 genes related to estrogen metabolism and action. We then performed low-density array analysis of these 172 genes on 11 ovarian endometriosis samples and 9 control endometrium samples. Principal component analysis of the gene expression levels showed clear separation between the endometriosis and control groups. We identified 78 genes as differentially expressed. Based on Ingenuity pathway analysis, these differentially expressed genes were arranged into groups according to biological function. These analyses revealed that 32 differentially expressed genes are estrogen related, 23 of which have not been reported previously in connection with endometriosis. Functional annotation showed that 25 and 22 genes are associated with the biological terms "secreted" and "extracellular region", respectively. Differential expression of 4 out of 5 genes related to estrogen metabolism and action (*ESR1*, *ESR2*, *PGR* and *BGN*) was also confirmed by immunohistochemistry. Our study thus reveals differential expression of several genes that have not previously been associated with endometriosis and that encode potential novel biomarkers and drug targets.

1. Introduction

Endometriosis is a very common gynecological disorder. It occurs in up to 15% of women of reproductive age, and increases to 30–50% in patients with infertility [1]. Since the symptoms of reduced fertility and the variety of associated pain are not specific for endometriosis, this disease is very difficult to diagnose without surgical assessment [2]. Extra-uterine endometrial tissue can be found on the ovaries, the pelvis, and the rectovaginal septum, thus forming three different entities: ovarian endometriosis, peritoneal endometriosis, and deep endometriotic nodules of the rectovaginal tract [3,4].

Previous studies have shown that changes in expression levels of a large number of genes are involved in the development of endometriosis. Most studies have compared gene-expression profiles between ectopic and eutopic endometrium of the same patients [5–11], while some have been performed on ectopic endometrium and eutopic endometrium of non-endometriosis patients [12]. Another group of studies examined expression profiles of eutopic endometrium from women with endometriosis in comparison with eutopic endometrium of disease-free women [13–17]. The majority of these studies were performed on whole tissue [8–13,15], while some used enriched endometriotic epithelial cells [16], and some others, laser-capture microdissection for separation of epithelial cells [7,14,18]. Using different methodologies, including DNA microarrays, subtractive hybridization, and serial analysis of gene expression, up to 5600 genes have been identified as differentially expressed, and belonging to different biological pathways [9,16].

In the present study, we used a different approach. We analyzed transcription-level differences between ovarian endometriosis and eutopic endometrium of non-endometriosis patients, for 152 genes

Abbreviations: BGN, biglycan; ECM, extracellular matrix; EMT, epithelial–mesenchymal transition; ER, estrogen receptor; LDA, low-density array; PCA, principal component analysis; PPAR, peroxisome proliferator-activated receptor; PR, progesterone receptor; RIN, RNA integrity number.

* Corresponding author. Tel.: +386 1 5437657; fax: +386 1 5437641.
E-mail address: Tea.Lanisnik-Rizner@mf.uni-lj.si (T.L. Rižner).

0960-0760/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsbmb.2011.03.010

that were selected based on Affymetrix chip analysis of ovarian endometriosis, which were combined with a selection of 20 genes related to estrogen metabolism and action which were not detected with this chip analysis. We validated these gene expression patterns by real-time PCR and selected five estrogen-related genes for validation at the protein level. Our results show that 78 genes from different biological processes and four of the five proteins examined are differentially expressed in endometriotic tissue. Our data thus suggest that proteins encoded by the aberrantly expressed genes identified can serve as potential diagnostic markers and as potential targets for pharmaceutical development.

2. Methods

2.1. Affymetrix chip analysis

The basis for the design of the TaqMan low-density array (LDA) was the analysis of genome-wide expression profiles from Affymetrix chip analyses of an original number of 129 endometriosis-related samples. These were obtained from the Genelogic ASCENTA® System (http://www.genelogic.com), a commercial database that provides microarray gene-expression profiles from clinically relevant normal and diseased human tissue sample sets. These sample sets are curated and annotated by pathologists and molecular biologists for use in target identification and discovery of biomarkers of disease. The hierarchical clustering of expression profiles (nearest-neighbor, the unweighted pair-group method, arithmetic means (UPGMA), Euclidean distance), and the comparison analysis were carried out using the Spotfire DecisionSite 7.0 software. Sample site (=tissue type), pathology, age, disease or medication were checked to see if they caused samples to cluster together (Fig. 1).

2.2. Subjects and tissue specimens for LDA analysis and conventional real-time PCR

Thirty-one premenopausal women were included: 11 patients with ovarian endometriosis stages II to IV (mean age, 30.4 ± 3.6 years), 9 patients with *Uterus myomatosus* or *Myoma uteri* (mean age, 45.2 ± 2.9 years) (Table 1); and 11 healthy women (mean age, 37.2 ± 1.8 years). Samples of ovarian endometriosis were obtained after surgical removal. Samples of endometrium were obtained either after hysterectomy due to myoma, or as biopsies during sterilization procedures. Cycle stage was estimated according to the date of the last menstrual phase or by histological evaluation (Table 1). The study was approved by the National Medical Ethics Committee of the Republic of Slovenia, and tissue samples were obtained with full and informed consent of the patients.

2.3. Reverse transcription and real-time PCR expression analysis

RNA isolation was performed using TriReagent (Sigma). RNA quality was checked on an Agilent Bioanalyser, and the RNA was reverse-transcribed to cDNA. The samples were not pooled; each of 20 cDNA samples (11 from ovarian endometriosis, nine from control endometrium of patients with myoma) was analyzed on a custom-made TaqMan LDA® (Applied Biosystems, Foster City, CA, USA) for the selection of the 172 genes of interest and 18 house-keeping genes. Expression of five genes (*CYP19A1*, *ESR1*, *ESR2*, *PGR* and *BGN*), and *PPIA* as a house-keeping gene, was examined separately by conventional real-time PCR, Taq Man PCR assay (Assay-on-Demand, Applied Biosystems) on the 11 samples of ovarian endometriosis, nine samples of control endometrium, and 11 samples of normal endometrium from disease-free women.

Table 1
Patient details for control endometrium (N) and ovarian endometriosis (E) groups.

Control endometrium	Age	Phase of menstrual cycle	Diagnosis
N1	42	Non-proliferative	Uterus myomatosus
N2	47	Late secretory	Myoma uteri
N3	43	Early secretory	Uterus myomatosus
N5	45	Proliferative	Uterus myomatosus
N6	50	Early secretory	Uterus myomatosus
N8	42	Late secretory	Myoma uteri
N9	47	Proliferative	Uterus myomatosus
N10	48	Proliferative	Myoma uteri
N11	43	Early secretory	Uterus myomatosus
Ovarian endometriosis	**Age**	**Phase of menstrual cycle**	**Disease stage[a]**
E14	24	Proliferative	III
E15	33	Late proliferative/early secretory	III
E16	30	Late proliferative/early secretory	IV
E19	27	Early secretory	II
E20	30	Late proliferative/early secretory	III
E21	34	Proliferative	III
E22	32	Secretory	III
E23	27	Secretory	III
E24	36	Proliferative	III
E25	33	Proliferative	IV
E26	28	Late proliferative/early secretory	III

[a] Stage of disease according to the revised American Fertility Society classification.

2.4. Immunohistochemistry

The paraffin-embedded specimens of ovarian endometriosis and control endometrium (11 and nine specimens, respectively, from the same patients as for LDA) were de-waxed in xylene and rehydrated. The sections were incubated in 0.9% H_2O_2 to block endogenous peroxidase. After antigen retrieval, the sections were incubated with the following mouse monoclonal antibodies: anti-aromatase (1:20, AbDSerotec), anti-estrogen receptor (ER)α (1:20, DakoCytomation, Denmark), anti-ERβ (1:200, GeneTex, USA), anti-progesterone receptor (PR)B (1:50, Santa Cruz, USA), and rabbit polyclonal antibodies anti-biglycan (BGN) (1:100, Sigma–Aldrich). The peroxidase–antiperoxidase complex was used with diaminobenzidine as substrate to reveal the bound antibodies. Staining intensities were scored as: 1, weak; 2, moderate; 3, very strong. The immunohistochemisry scores were calculated by multiplying the percentage of positive cells (P) by the intensity (I) (formula: $Q = P \times I$; maximum = 300). Two observers performed the scoring independently, showing good overall agreement at 68%, and the free marginal kappa was 0.62 [19].

2.5. Data analysis

The optimal set of normalization controls for LDA analysis was identified using GeNorm, and the geometric mean of the best three (*YWHAZ*, *SDHA*, *RPLO*) was used as a normalization factor [20]. To overcome the necessity of assuming normal distributions for our 172 gene-expression variables, we used non-parametric methods. We first used the Mann–Whitney *U*-test, grouping participants according to disease status (case, control), and then the Kruskal–Wallis test, grouping participants according to phase of the menstrual cycle (proliferative, secretory, late proliferative/early secretory). To take account of multiple testing, we applied the Benjamini-Yekutieli false discovery rate correction procedure [21], and corrected our p-values for performing 344 tests. We rejected the null hypotheses in all tests where corrected p-values were <0.05. To examine the relationships between expression at the RNA and protein levels, we calculated the Spearman's rank correlation coefficients. For identification of sample groups by hierarchical

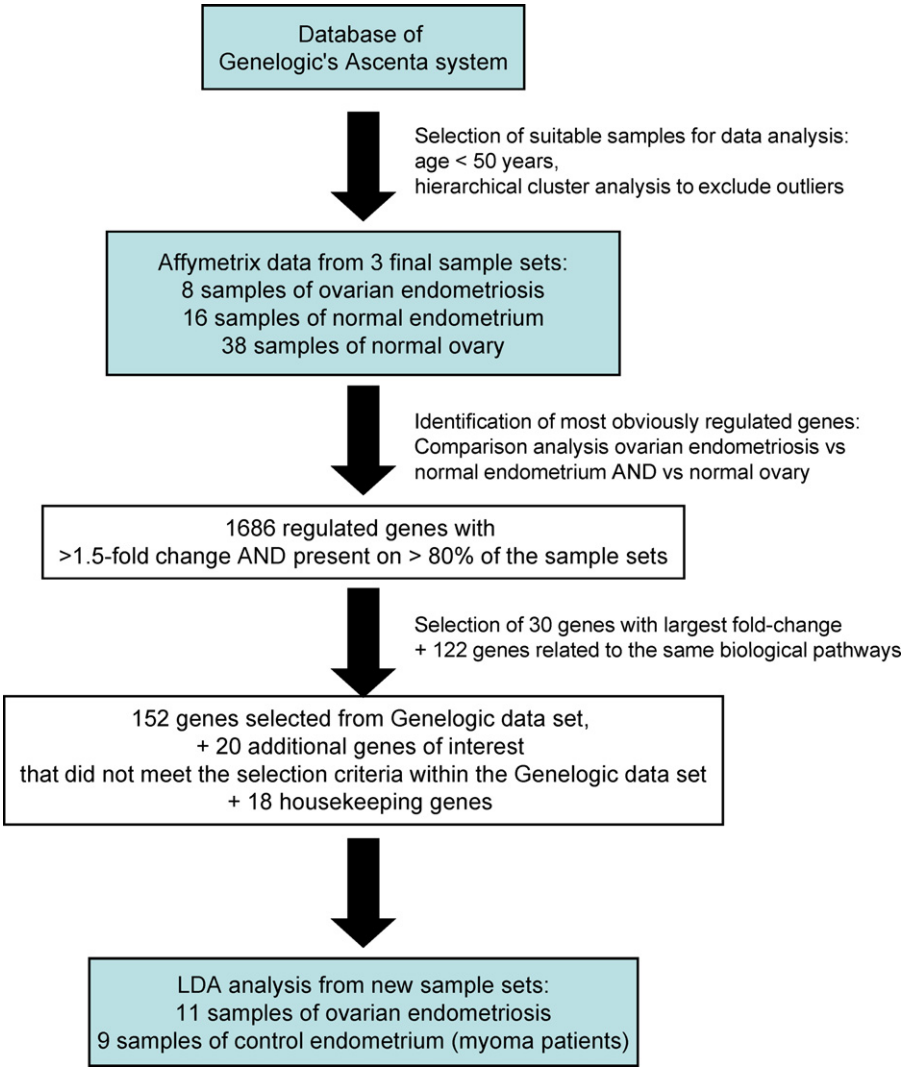

Fig. 1. Strategy for selection of potential biomarker genes for validation by LDA analysis.

clustering, we applied principal component analysis (PCA). The statistical calculations were performed using R-2.9.0 (R Development Core Team, http://www.R-project.org).

The networks and functional analyses were generated through Ingenuity pathways analysis (Ingenuity® Systems, www.ingenuity.com, version 8.0-2602). A corrected *p*-value cut-off of 0.044946 and fold-change cut-off of 1.73 were set, to identify genes showing significantly differentially regulated expression. The focus genes were overlaid onto a global molecular network that was developed from information from the Ingenuity pathways knowledge base. Networks of focus genes were algorithmically generated based on their connectivity, and they were also associated with the biological functions and/or diseases most significant for the dataset. The Fischer's exact test was used to calculate *p*-values relative to the probability that each biological function and/or disease assigned to that dataset was due to chance alone.

3. Results

3.1. Selection of potential biomarker genes for validation by LDA analysis

Our strategy for identification of potential biomarkers of ovarian endometriosis is illustrated in Fig. 1. First of all, we used *in silico* analysis of commercially available microarray expression data as a starting point for the validation of genes affected by ovarian endometriosis on a new sample set and with a different analytical technique (LDA). After hierarchical cluster analysis and restriction to samples from premenopausal women (aged < 50 years), the expression levels in ovarian endometriosis were compared to both normal ovary and normal endometrium, to identify endometrioma-specific genes that were not related to ovarian tissue or to normal endometrium. A list of 1686 genes with significantly different expression levels in ovarian endometriosis was obtained. The 30 genes with the largest differences were selected and were found to be assigned to the gene ontology terms "extracellular", "growth-factor activity", "angiogenesis" and "prostaglandin metabolism". The further 122 genes from the list were added if they were related either to the same gene ontology terms or if one of the following text words was included in their gene ontology description: extracellular, hormone, signal transduction, receptor, transcription, and G-protein. Confidence in the validity of the final gene list was provided by some of the genes already being mentioned in the scientific and patent literature, some of the genes from the same pathway sharing their patterns of differential gene expression, and some of the genes already being known to be expressed specifically in endometriosis tissue (see Section 4).

Finally, as especially ovarian endometriosis is regarded as an estrogen-dependent pathology, 20 genes were selected that are

known to be involved in estrogen metabolism and action, although they had not been detected in the Genelogic dataset.

3.2. LDA analysis identification of 78 genes that are differentially expressed

Among the 17 normalization controls, *YWHAZ* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide), *SDHA* (succinate dehydrogenase complex, subunit A) and *RPL0* (large ribosomal protein) showed the lowest variabilities; therefore, their geometric mean was used as the normalization factor. Out of 172 genes, 60 were significantly up-regulated and 18 were down-regulated in ovarian endometriosis compared to control endometrial tissue from patients with myomas (Table 2). The most highly up-regulated genes were *GATA4* (GATA binding protein 4), *PLA2G2A* (phospholipase A2, group IIA), *CYP19A1* (aromatase) and *ARHI* (Ras homolog gene family member I). These genes are a transcription factor that controls cell differentiation (*GATA4*), a putative tumor-suppressor gene (*ARHI*), and genes involved in immune responses and inflammation (*PLA2G2A*) and in steroid biosynthesis (*CYP19A1*). The most down-regulated genes were involved in remodeling of the extracellular matrix (ECM) (*MMP26*, matrix metalloproteinase 26), in fatty acid metabolism and in peroxisome proliferator-activated receptor (PPAR) signaling (*ACSL5*, acyl-CoA-synthetase long-chain family member 5), and nitric oxide signaling (*RASD1*, RAS dexamethasone induced 1). Functional annotation clustering by DAVID, a database for annotation, visualization and integrated discovery [22,23] revealed that "signal peptide", "secreted" and "extracellular region" are the most relevant biological terms associated with up to 36 out of the 78 genes (Supplemental Data).

3.3. Similar expression patterns in control endometrium of myoma patients and in normal endometrium

To validate the suitability of our control group of myoma patients, we separately performed conventional real-time PCR expression analysis of five exemplary genes (*CYP19A1*, *ESR2*, *ESR1*, *PGR*, *BGN*) in 11 samples of ovarian endometriosis, nine samples of control endometrium, and 11 samples of normal endometrium from healthy women. This real-time PCR supported the LDA analysis, showing up-regulation of *CYP19A1*, *ESR2* and *BGN*, and down-regulation of *ESR1* and *PGR* in ovarian endometriosis. Comparisons of the expression levels in ovarian endometriosis against control and normal endometrium showed very similar fold-deregulation of these genes (Fig. 2). The calculated pairwise correlation coefficients showed strong positive correlations between *ESR2*, *BGN* and *CYP19A1*, as well as between *PGR* and *ESR1*. The correlations were somewhat stronger when comparing ovarian endometriosis with normal endometrium from disease-free women, with positive Spearman's rank correlation coefficient Rho from 0.85 to 0.91, and negative Rho from −0.70 to −0.83. The ovarian endometriosis and control groups showed positive Rho from 0.78 to 0.84, and negative Rho from −0.46 to −0.70.

3.4. Immunohistochemistry analysis shows significant differences in expression of ESR1, ESR2, PGR and BGN, but not CYP19A1

Five estrogen-related genes were also selected to validate their expression at the protein level, using immunohistochemistry staining of paraffin sections: *CYP19A1* (aromatase), *ESR1* (ERα), *ESR1* (ERβ), *PGR* (PRB) and *BGN* (biglycan). Eleven samples of ovarian endometriosis and nine samples of control endometrium from the same patients as in the LDA analysis were stained (Fig. 3). For CYP19A1 and BGN, the scores represent cytoplasmic staining, and for ESR1, ESR2 and PGR, nuclear staining. The staining of these sections (Table 3) revealed significantly lower levels of ESR1 and PGR in epithelial cells of ovarian endometriosis, thus agreeing with the expression analysis at the mRNA level. Statistical analysis also showed strong positive correlations between expression at the mRNA and protein levels for *ESR1* (Rho = 0.5875, $p = 0.0104$) and *PGR* (Rho = 0.5111, $p = 0.0302$). In contrast to the real-time PCR data, the protein levels of *ESR2* and *BGN* were significantly lower in diseased endometrium. In stromal cells, there was a significant difference only in the expression of *ESR2*. *CYP19A1* mRNA levels were higher in epithelial and stromal cells of ovarian endometriosis, but these differences were not significant.

3.5. Differentially expressed genes belong to different biological processes

Bioinformatic pathway analysis defined the association of the differentially expressed genes to several biological functions (Table 4). Within the diseases-function category, "cancer", "reproductive system disease", "inflammatory disease" and "immunological disease" showed the best fits to our dataset (p from 7.01E−10 to 4.50E−07). "Cell death", "cell growth and proliferation", "cell movement", "cell-to-cell signaling and interaction" and "molecular transport", were the most highly ranked among the molecular and cellular functions (p from 7.39E−09 to 4.46E−06). "Tissue morphology", "organ development", "angiogenesis", "cardiovascular system development and function" and "skeletal and muscular system development and function", showed the best fits to our dataset within the physiological system functions (p from 2.26E−09 to 1.14E−05).

As endometriosis is an estrogen-dependent disease, we also examined which of these differentially expressed genes are related to estrogen action. Network analysis (Fig. 4) identified 32 deregulated genes from our dataset, among which nine had previously been related to endometriosis, while according to Ingenuity pathway analysis the other 23 genes had not been reported previously in connection with endometriosis: *BGN* (biglycan), *TIMP1* (tissue inhibitor of metalloproteinase I), *ITGB5* (integrin beta 5), *KAL1* (Kallmann syndrome 1 sequnce), *WISP2* (WNT1 inducible signaling pathway protein 2), *SEMA3C* (semaphoring 3C), *PTGIS* (prostaglandin I2 synthase), *APCDD1* (adenomatosis polyposis coli down-regulated 1), *ALDH1A2* (aldehyde dehydrogenase 1 family, member A2), *EDG1* (sphingolipid G-protein-coupled receptor 1, corresponds to S1PR1 in the Figure 4), *VCAM1* (vascular cell adhesion molecule 1), *CXCL1* (chemokine, ligand 1), *INHA* (inhibin, alpha), *GATA4* (GATA binding protein 4), *HOXC6* (homeo box C6), *ITGA2* (integrin, alpha 2), *SCARB1* (scavenger receptor class B, member 1), *SN* (sialoadhesin, corresponds to SIGLEC1 in the Figure 4), *SLC24A3* (solute carrier family 24, member 3), *FNDC1* (fibronectin type-III-domain-containing 1), *CYB5A* (cytochrome b-5), *ACSL5* (acyl-CoA synthase long-chain family member 5) and *ADAM23* (a disintegrin and metalloproteinase domain 23).

3.6. Menstrual phases affect the expression of only nine genes

As well as age, the stage in the menstrual cycle represents an important confounding variable in the analysis of our data. Therefore, the differentially expressed genes were separately analyzed, while the samples from the study participants were grouped according to the phase of the menstrual cycle (proliferative, late proliferative/early secretory and secretory). Out of the 172 genes tested, only nine revealed significant differences in expression levels between the menstrual phases in any of the tissues. Among these, in the ovarian endometriosis samples, two genes showed no significant differences (*SFRP4* [secreted frizzled-related protein 4] and *DVL2* [dishevelled, dsh homolog 2]), and seven genes showed significantly higher levels (*KIAA0427*,

Table 2
Genes differentially expressed in ovarian endometriosis.

	Gene ID[a]	Gene or its function	*p* value	FC
Up-regulated				
1	**GATA4**	GATA binding protein 4	0.005979	7066918.78
2	PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	0.006078	2469.32
3	CYP19A1	Aromatase	0.010386	509.66
4	ARHI	Ras homolog gene family, member I	0.013437	162.41
5	WISP2	WNT1 inducible signaling pathway protein 2	0.011775	60.85
6	COL8A1	Collagen, type VIII, alpha 1	0.006105	53.11
7	**TCF21**	Transcription factor 21	0.013437	51.38
8	CPXM2	Carboxypeptidase-like protein X2	0.001471	49.88
9	PTPRZ1	Protein tyrosine phosphatase, receptor-type, Z polypeptide 1	0.007985	45.94
10	ESR2	Estrogen receptor 2 (ER beta)	0.001977	42.30
11	HOXC6	Homeo box C6	0.007985	41.48
12	PTGIS	Prostaglandin I2 (prostacyclin) synthase	0.003512	41.05
13	INHA	Inhibin, alpha	0.001083	34.98
14	HF1	H factor 1 (complement)	0.004643	34.26
15	PRSS35	Protease, serine, 35	0.013437	24.03
16	C20orf103	Chromosome 20 open reading frame 103	0.022132	22.72
17	C3	Complement component 3	0.003512	18.80
18	KAL1	Kallmann syndrome 1 sequence	0.001089	12.26
19	VCAM1	Vascular cell adhesion molecule 1	0.003512	11.95
20	CTHRC1	Collagen triple helix repeat containing 1	0.010386	11.35
21	TNFSF13B	Tumor necrosis factor (ligand) superfamily, member 13b	0.002642	10.09
22	GATA6	GATA binding protein 6	0.006105	10.06
23	CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	0.044946	9.70
24	**FZD7**	Frizzled homolog 7	0.013437	9.36
25	ADAM23	A disintegrin and metalloproteinase domain 23	0.001977	9.17
26	FIBL-6	Hemicentin	0.022132	9.10
27	AEBP1	AE binding protein 1	0.004643	8.77
28	SN	Sialoadhesin	0.001977	8.49
29	BGN	Biglycan	0.007985	8.22
30	LAMA1	Laminin, alpha 1	0.013437	8.10
31	C4B	Complement component 4A (Rodgers blood group)	0.010386	7.26
32	TIMP1	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	0.004643	7.05
33	FNDC1	Fibronectin type III domain containing 1	0.035682	6.88
34	ARHGAP25	EGF-containing fibulin-like extracellular matrix protein 1	0.002642	6.32
35	EFEMP1	Rho GTPase activating protein 25	0.001089	6.14
36	DKK3	Dickkopf homolog 3 (Xenopus laevis)	0.017292	6.03
37	DOK5	Docking protein 5	0.007985	5.98
38	SYNPO	Synaptopodin	0.007985	5.42
39	ICAM1	Intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	0.002642	5.30
40	**SCARB1**	Scavenger receptor class B, member 1	0.003512	5.24
41	SEMA3C	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C	0.013437	4.78
42	MRC1	Mannose receptor, C type 1	0.002642	4.53
43	TIMP2	Tissue inhibitor of metalloproteinase 2	0.013437	4.32
44	CYB5	Cytochrome b-5	0.035682	4.02
45	ING1L	Inhibitor of growth family, member 2	0.017292	3.99
46	ANKRD6	Ankyrin repeat domain 6	0.001977	3.99
47	TM6SF1	Transmembrane 6 superfamily member 1	0.006105	3.94
48	**KIAA0427**	KIAA0427	0.004643	3.85
49	CALD1	Caldesmon 1	0.004643	3.76
50	EDIL3	EGF-like repeats and discoidin I-like domains 3	0.017292	3.44
51	**COL14A1**	Collagen, type XIV, alpha 1 (undulin)	0.007985	3.42
52	RNASE1	Ribonuclease, RNase A family, 1 (pancreatic)	0.001471	3.41
53	EDG1	Endothelial differentiation, sphingolipid G-protein-coupled receptor, 1	0.006105	3.15
54	RAB7L1	RAB7, member RAS oncogene family-like 1	0.001089	2.70
55	RGS10	Regulator of G-protein signaling 10	0.010386	2.51
56	ERG	v-ets erythroblastosis virus E26 oncogene homolog	0.007985	2.49
57	ITGB5	Integrin, beta 5	0.022132	2.29
58	KCTD2	Potassium channel tetramerisation domain containing 2	0.010386	1.88
59	CHN2	Chimerin	0.028177	1.83
60	NFE2L3	Nuclear factor (erythroid-derived 2)-like 3	0.017292	1.73
Down-regulated				
1	MMP26	Matrix metalloproteinase 26	0.006856	−173491.72
2	ACSL5	Acyl-CoA synthetase long-chain family member 5	0.003512	−11.11
3	RASD1	RAS, dexamethasone-induced 1	0.001977	−10.62
4	**LGR7**	Relaxin/insulin-like family peptide receptor 1	0.002642	−8.69
5	APCDD1	Adenomatosis polyposis coli down-regulated 1	0.017292	−8.63
6	ALDH1A2	Aldehyde dehydrogenase 1 family, member A2	0.001471	−8.21
7	SLC24A3	Solute carrier family 24 (sodium/potassium/calcium exchanger), member 3	0.004643	−6.21

Table 2 (*Continued*)

	Gene ID[a]	Gene or its function	*p* value	FC
8	PPM1H	Protein phosphatase 1H (PP2C domain containing)	0.001471	−4.56
9	ESR1	Estrogen receptor 1	0.028177	−4.03
10	SLC7A1	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 1	0.002642	−3.93
11	PGR	Progesterone receptor	0.001089	−3.71
12	ITGA2	Integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	0.028177	−3.14
13	VEGF	Vascular endothelial growth factor	0.044946	−2.97
14	FLJ11539	Hypothetical protein FLJ11539	0.022132	−2.86
15	ANK3	ankyrin 3, node of Ranvier (ankyrin G)	0.001977	−2.67
16	WNT2	Wingless-type MMTV integration site family member 2	0.044946	−1.98
17	MGC4172	DHRS11 dehydrogenase/reductase (SDR family) member 11	0.010386	−1.95
18	PSME4	Proteasome (prosome, macropain) activator subunit 4	0.013437	−1.83

Bold text, the seven genes differentially expressed in different phases of the menstrual cycle. FC, fold-change.
p-values are corrected for multiple testing.

[a] The 78 differentially expressed genes are listed according to descending fold-changes.

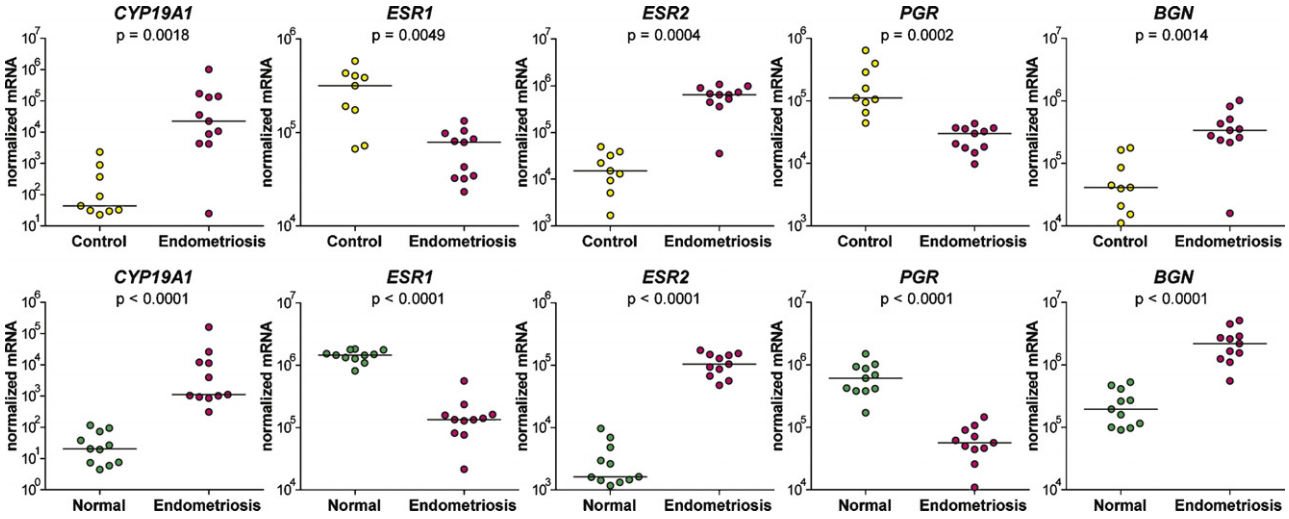

Fig. 2. Real-time PCR expression analysis of *CYP19A1*, *ESR1*, *ESR2*, *PGR* and *BGN*. Top panels: comparisons between control endometrium of *Uterus myomatosus* or *Myoma uteri* patients (Control) and ovarian endometriosis specimens (Endometriosis). Bottom panels: comparisons between normal endometrium of healthy women (Normal) and ovarian endometriosis specimens (Endometriosis).

GATA4, *SCARB1*, *LGR7* [relaxin/insulin-like family peptide receptor 1], *COL14A1* [collagen, type XIV, alpha 1], *FZD7* [frizzled homolog 7] and *TCF21* [transcription factor 21]) (Table 2). For these seven genes, we cannot rule out potential interactions between disease status and phase of the menstrual cycle.

3.7. Cluster analysis shows clear separation between endometriosis and control endometrium groups

When PCA was applied to a matrix of expression levels of the 172 genes or five proteins and 20 subjects (11 specimens of ovarian endometriosis and nine specimens of control endometrium), a separation of subjects was seen (Fig. 5). The deregulated genes and proteins were clustered into two groups, the endometriosis and control endometrium groups. This demonstrates that the samples can be distinguished based on their expression profiles. Interestingly, one of the endometriosis samples clustered together with the samples of control endometrium. This sample was obtained from the only endometriosis patient with stage II disease. To a certain extent, PCA also showed clustering according to the phases of the menstrual cycle.

4. Discussion

Early diagnosis of endometriosis would be beneficial for the quality of life of these patients. However, as the pathogenesis of this disease is still not well understood, more information at the molecular level is needed. We have here contributed to this challenge through this analysis of differentially expressed genes.

4.1. Introduction of an additional control group

Our LDA analysis was performed on samples of ovarian endometriosis and control endometrium of myoma patients, which have often been used as a control group. However, this can be a potential source of bias due to considerable age differences between the case and the control groups, and as *Uterus myomatosus* and *Myoma uteri* do not represent healthy states. To address this issue, we additionally examined the expression of five genes in ovarian endometriosis compared with normal endometrium from healthy women, as well as with control endometrium of myoma patients. In both cases, we saw similar expression patterns and fold deregulation (Fig. 2), which supports the results of Bukulmez et al. [24]. Indeed, these data represent a good basis for the assumption that our myoma patients are a suitable control group for our study. Of note, due to the minimum age requirement for sterilization procedures, it was not possible to obtain samples from younger women, and therefore the mean age of our healthy control group was still seven years more than that of our ovarian endometriosis group. We therefore cannot entirely rule out potential bias due to the age differences between our cases and these healthy controls.

Table 3
Scoring of immunohistochemistry staining for the CYP19A1, ERα, ERβ, PRB and BGN proteins in control endometrium (N) and ovarian endometriosis (E).

Sample	Immunohistochemistry staining score[a]									
	CYP19A1		ERα		ERβ		PRB		BGN	
	e	s	e	s	e	s	e	s	e	s
N1	14.00	9.60	295.80	196.00	195.00	184.74	283.00	240.00	100.00	111.12
N2	0.00	100.00	200.70	154.00	205.93	184.00	210.14	250.00	145.40	108.38
N3	92.70	100.00	196.10	177.00	235.70	260.02	242.80	227.00	230.00	55.04
N5	6.10	114.48	287.50	265.00	5.00	93.64	196.60	142.00	246.90	114.60
N6	195.30	192.50	42.80	155.00	2.56	206.06	156.00	196.00	209.00	42.11
N8	8.60	0.00	209.95	123.00	55.30	100.00	86.60	155.00	289.80	16.70
N9	0.00	0.00	287.50	259.00	111.90	72.14	103.30	127.00	100.00	135.29
N10	0.00	0.00	295.40	46.00	70.00	0.00	300.00	107.00	140.40	110.18
N11	1.96	0.00	288.90	56.00	114.50	131.10	209.00	242.00	255.40	104.90
Median	6.10	9.60	287.50	155.00	111.90	131.10	209.00	196.00	209.00	108.38
E14	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	25.14	148.40
E15	5.40	0.00	211.80	211.88	0.00	0.00	27.30	56.04	128.00	147.24
E16	3.40	82.20	64.60	131.40	6.80	48.40	73.20	60.00	8.10	99.40
E19	34.65	51.90	174.72	158.30	140.57	102.28	211.00	223.33	275.00	117.00
E20	0.00	63.32	36.50	125.00	0.00	3.75	0.00	188.50	13.30	89.23
E21	17.90	50.82	64.90	78.40	5.95	95.14			216.60	45.24
E22	3.57	15.00	75.90	103.50	2.20	4.82	109.60	191.70	77.50	82.34
E23	22.20	0.00	198.60	129.89	0.00	0.00			40.80	159.90
E24	205.90	21.65	158.60	82.50	45.90	6.25	146.80	202.00	87.90	105.00
E25	76.40	26.66	148.30	172.10	31.20	30.72	101.15	182.50	24.70	136.30
E26	171.50	0.00	57.43	0.00	124.99	0.00	56.15	198.30	152.90	56.99
Median	17.90	21.65	75.90	125.00	5.95	4.82	73.20	188.50	77.50	105.00
Fold-change	2.93	2.26	-3.79	-1.24	-18.81	-27.20	-2.86	-1.04	-2.70	-1.03
p-value	0.42	0.61	0.01	0.18	0.02	0.01	0.01	0.23	0.02	0.34

e, scoring for epithelial cells; s, scoring for stromal cells.
[a] Maximum score, 300 (see Section 2).

4.2. Selection and characterization of samples of ovarian endometriosis and control endometrium

In our mRNA expression analysis (LDA and conventional real-time PCR), we decided to use the complex samples of ovarian endometriosis rather than laser-captured microdissected tissue. This decision was made as all cell types within ovarian endometriosis tissue can contribute to the pathogenesis of this disease [5,9], and as laser-capture microdissection may produce RNA of a lower quality. Matsuzaki et al. [18] published a protocol for the production of high quality RNA from endometriotic tissue after laser capture microdissection; however, their approach is laborious, and it is therefore only suitable for small numbers of samples. As the low quality of RNA can also be a serious problem often [25], we instead opted to work with high-quality RNA isolated from ovarian endometriosis that had RNA Integrity Number (RIN) values close to, or above 8.

The cycle stages of the control endometrium were determined by histological evaluation. We did not have access to the eutopic endometrium of the endometriosis patients, and therefore for this group the cycle stage was estimated according to the date of the last menstruation. Classification of the cycle stage based on menstrual dating is usually considered not to be satisfactory. Interestingly, Burney and colleagues [15] recently reported that the histological classification of menstrual phases by Noyes et al. [26] is not always appropriate, since gene expression profiles show more relevant information. For three endometrial specimens, they showed that rather than the histological classification, the dating based on the last menstrual period confirmed their molecular-based dating [15].

4.3. Validation of LDA data by immunohistochemistry

Expression analysis at the protein level was performed with antibodies described in the literature [27,28] and characterized as specific by Western blotting. The staining for ERα (*ESR1*) and PRB (*PGR*) supports the LDA data. In contrast to the LDA results, we found lower levels of the ERβ (*ESR2*) and BGN proteins in ovarian endometriosis. Increased expression of *ESR2* at the mRNA level has been explained by deficient DNA methylation of the *ESR2* promotor [29]. Our data support this hypothesis at present, in terms of its up-regulation at the mRNA level. One possible explanation for lower protein levels of *ESR2* in epithelial and stromal cells might be different ratios of *ESR2* splice variants in ovarian endometriosis *versus* control endometrium, and failure of the monoclonal antibody to recognize these variants [30]. Increased mRNA levels of *BGN* have been reported previously in ovarian endometriosis [10]; however, expression at the protein level has not been examined before. Differences in protein levels of BGN in stromal and epithelial cells might partially explain the contrast between up-regulation at the mRNA level and down-regulation at the protein level, while the significantly lower protein levels seen in epithelial cells support previous reports that estrogens decrease *BGN* expression [31]. Interestingly, in the epithelial and stromal cells of ovarian endometriosis, we did not see significant differences in expression of *CYP19A1*. Interestingly, Colette et al. [32] noted no staining for CYP19 in sections of ovarian, peritoneal and deep infiltrating endometriosis. Using real-time PCR, they showed low levels of the CYP19A1 transcript in ovarian endometriosis, and barely detectable levels in peritoneal and rectovaginal endometriosis. They explained the presence of the *CYP19A1* transcript in ovarian endometriosis by follicular expression [32]. Our data, along with further published data, thus show that immunohistochemistry analysis is indispensable when validating differentially expressed genes, especially with samples of heterogenous tissue, such as ovarian endometriosis.

4.4. Information obtained from cluster analysis

PCA of expression profiles shows a clear separation of the subjects into endometriosis and control groups, with the exception

Table 4
Gene categories according to biological processes and pathways identified.

Function category	Category	*p* value	Number of genes	Genes[a]
Diseases	Cancer	7.01E−10	36	ACSL5, **ADAM23**, ALDH1A2, ANK3, **BGN**, **CALD1**, **HF1**, **CXCL1**, **CYP19A1**, **ARHI**, **EFEMP1**, **ERG**, ESR1, **ESR2**, **FZD7**, **GATA4**, **GATA6**, **HOXC6**, **ICAM1**, **INHA**, **ITGB5**, PGR, **PLA2G2A**, PSME4, **PTGIS**, **RNASE1**, **EDG1**, **SEMA3C**, SLC24A3, SLC7A1, **TCF21**, **TIMP1**, **TIMP2**, **VCAM1**, VEGFA, WNT2
	Reproductive system disease	4.56E−08	12	ALDH1A2, **HF1**, **CYP19A1**, **ARHI**, ESR1, **ESR2**, **GATA4**, **GATA6**, **INHA**, PGR, **SEMA3C**, VEGFA
	Inflammatory disease	4.50E−07	7	**C4B**, **HF1**, ESR1, **ICAM1**, **TNFSF13B**, **VCAM1**, VEGFA
	Immunological disease	6.18E−07	31	**ADAM23**, ANK3, **BGN**, **C4B**, **CALD1**, **HF1**, **CHN2**, **COL14A1**, **COL8A1**, **CXCL1**, **CYB5A**, **CYP19A1**, **DOK5**, **EDIL3**, **EFEMP1**, ESR1, **ESR2**, **ICAM1**, **INHA**, **KIAA0427**, **LAMA1**, **MRC1**, PGR, PPM1H, **EDG1**, **SN**, SLC24A3, SLC7A1, **TNFSF13B**, **VCAM1**, VEGFA
Molecular and cellular function	Cell death	2.89E−09	32	ALDH1A2, **BGN**, **HF1**, **CXCL1**, **CYB5A**, **CYP19A1**, **ARHI**, **DKK3**, **ERG**, ESR1, **ESR2**, **GATA4**, **GATA6**, **HOXC6**, **ICAM1**, ING1L, **INHA**, ITGA2, **ITGB5**, PGR, **PTGIS**, **PTPRZ1**, RASD1, **RGS10**, **RNASE1**, **EDG1**, **TIMP1**, **TIMP2**, **TNFSF13B**, VEGFA, WNT2
	Cell growth and proliferation	7.39E−09	24	**AEBP1**, **BGN**, **COL14A1**, **COL8A1**, **CXCL1**, ESR1, **ESR2**, **GATA4**, **GATA6**, **ICAM1**, **INHA**, ITGA2, **LAMA1**, PGR, **PLA2G2A**, **EDG1**, **SCARB1**, **SN**, **TIMP2**, **TNFSF13B**, **VCAM1**, VEGFA, **WISP2**
	Cell movement	1.24E−07	19	**BGN**, **C4B**, **CALD1**, **CXCL1**, **CYP19A1**, **DKK3**, ESR1, **ESR2**, **ICAM1**, ITGA2, **ITGB5**, **KAL1**, **PTPRZ1**, LGR7, **EDG1**, **TIMP1**, **VCAM1**, VEGFA, **WISP2**
	Cell-to-cell signaling and interaction	4.46E−06	10	**BGN**, **CXCL1**, **ICAM1**, ITGA2, **ITGB5**, **MRC1**, **EDG1**, **SCARB1**, **VCAM1**, VEGFA
	Molecular transport	5.78E−06	6	**CYP19A1**, **ARHI**, ESR1, **ESR2**, **INHA**, **TIMP1**
Physiological system development and function	Tissue morphology	2.26E−09	22	ANK3,, **BGN**, **C4B**, **HF1**, **CXCL1**, **CYP19A1**, **DKK3**, ESR1, **ESR2**, **GATA4**, **GATA6**, **ICAM1**, **INHA**, PGR, **RGS10**, **EDG1**, **SCARB1**, **SN**, **TNFSF13B**, **VCAM1**, VEGFA
	Organ development	1.35E−08	18	**AEBP1**, ALDH1A2, **COL8A1**, **CYP19A1**, **ARHI**, ESR1, **ESR2**, **GATA4**, **GATA6**, **INHA**, ITGA2, **EDG1**, **SEMA3C**, **TCF21**, **TIMP1**, **TIMP2**, **VCAM1**, VEGFA
	Cardiovascular system development and function	2.96E−07	13	ALDH1A2, **EDIL3**, **GATA4**, **GATA6**, ITGA2, **ITGB5**, **LAMA1**, **EDG1**, **SEMA3C**, **TIMP1**, **TIMP2**, **VCAM1**, VEGFA
	Angiogenesis	1.14E−05	10	**EDIL3**, **GATA4**, ITGA2, **ITGB5**, **EDG1**, **SEMA3C**, **TIMP1**, **TIMP2**, **VCAM1**, VEGFA
	Skeletal and muscular system development and function	3.80E−07	9	ESR1, **GATA4**, **GATA6**, ITGA2, **LAMA1**, PGR, **EDG1**, VEGFA, **WISP2**

Some of the genes are further discussed in the main text: "Disturbed cellular processes as a source of biological markers and novel drug targets".
[a] Genes in bold are up-regulated.

of the patient with endometriosis stage II. Our data appear to be in contrast to those of Burney et al. [15], who showed that the clustering of gene expression in three types of endometriosis was more dependent on the cycle phase than on the endometriosis status. However, expression analysis by Wu et al. [33] clearly showed separation between ovarian endometriosis and non-ovarian endometriosis, which thus supports our data. Although further studies on larger numbers of samples are needed, our data suggest that the expression profiles in this stage-II patient did not differ significantly from the expression profiles of eutopic endometrium of non-endometriosis patients.

4.5. Disturbed cellular processes as a source of biological markers and novel drug targets

Our LDA gene expression analysis revealed that 78 putative endometriosis-related genes are involved in different cellular processes. Most of these genes have roles in more than one of the biological processes identified. We have chosen here to summarize the four processes that were identified on the basis of pathway analysis (Table 4), supplemented with information from the literature.

4.5.1. Cell-to-cell signaling, interaction and cellular movement

The specific alterations of eutopic and ectopic cell-adhesion molecules facilitate the attachment of refluxed endometrium to ectopic sites [34]. Four genes in our study identified as up-regulated: a disintegrin and metalloproteinase domain 23 (*ADAM23*), laminin, alpha 1 (*LAMA1*), integrin beta-5 (*ITGB5*), and fibronectin type III domain-containing protein 1 (*FNDC1*) are related to cell-adhesion processes [34–36]. Additionally, cancer-cell migration promoting collagen triple helix repeat containing 1 (*CTHRC1*) [37] and Rho GTPase activating protein (*ARHGAP*) family genes, which are implicated in actin remodeling, cell polarity and migration [38], were up-regulated in the endometriosis group. We also observed a 3-fold down-regulation of integrin alpha 1 subunit (*ITGA2*), which changes cell–cell and cell–matrix interactions that are important for the development of endometriosis [39]. We thus found aberrant mRNA levels of several novel genes associated with cell migration and cell adhesion: *ADAM23*, *ITGB5*, *FNDC1*, *CTHRC1* and *ARHGAP25*.

Once an ovarian cyst has developed and its invasion through degradation of the ECM is complete, the proteolytic activity decreases [40]. Among the nine matrix metalloproteinases (MMPs) included in the present study, only *MMP26* was differentially expressed. As expected, we observed up-regulation of genes encod-

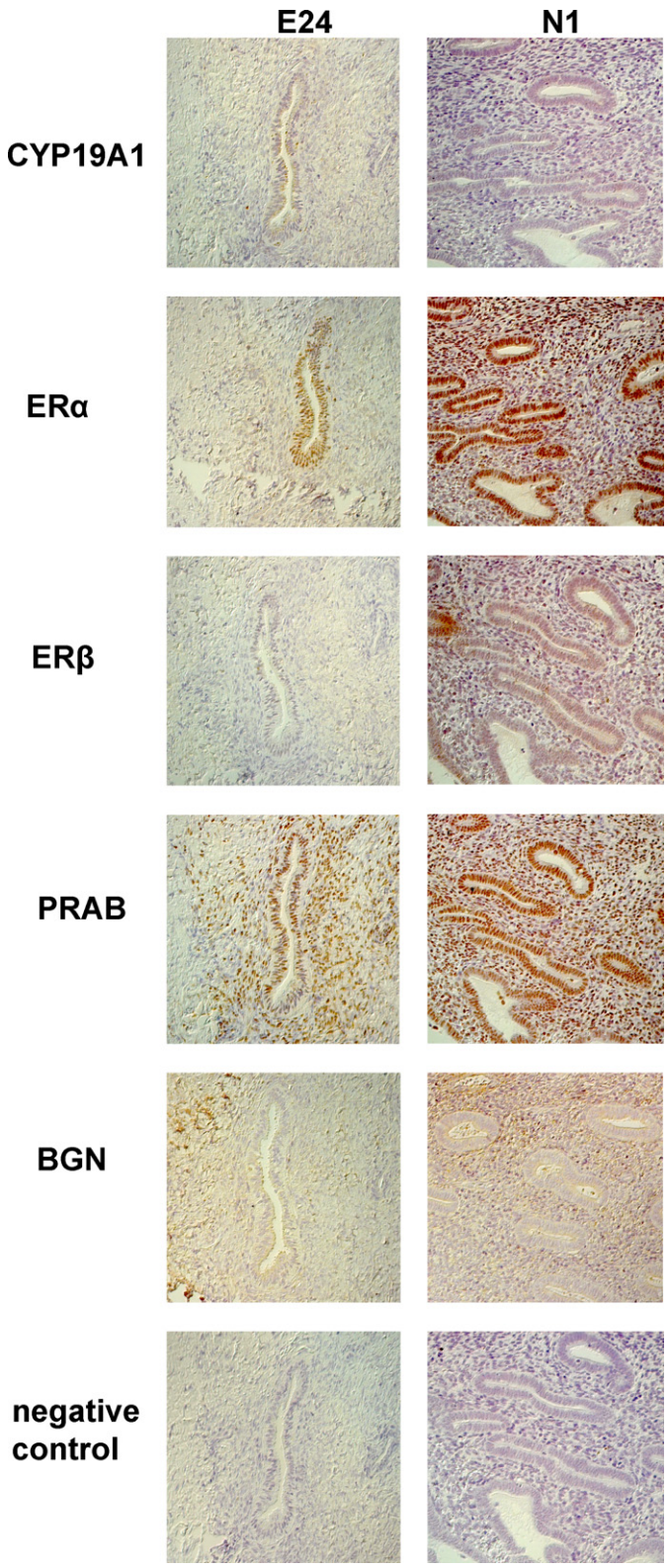

Fig. 3. Immunohistochemistry staining of the CYP19A1, ERα, ERβ, PRAB and BGN proteins in ovarian endometriosis and control endometrium. E24: ectopic endometrium from a patient with ovarian endometriosis. NM1: control endometrium from a patient with *Myoma uteri* (magnification, 200×).

ing tissue inhibitors of metalloproteinase, TIMP-1 and TIMP-2, the latter being a potent inhibitor of MMP26 [41]. We have contributed to new knowledge here also by the up-regulation of the gene encoding EGF-containing fibulin-like extracellular matrix protein 1 (*EFEMP1*) in the endometriosis group. This protein targets endothelial cell expression of the *MMP* and *TIMP* genes and reduces proteolysis and cell remodeling [42]. Our study thus confirms previously reported differential expression of the *MMP* and *TIMP* genes [40], and suggests that *EFEM1* contributes to the lower proteolytic activity. We also observed higher mRNA levels of hemicentin (*FIBL-6*), undulin (*COL14A1*) and biglycan (*BGN*, discussed in Section 4.6). The last two of these proteins have special organizing functions in the assembly of the ECM [43,44].

4.5.2. Cancer-related genes

Endometriosis is a benign proliferative disorder where ectopic lesions resemble metastasis [1,3]. Initiation of metastasis comprises steps similar to epithelial–mesenchymal transition (EMT) (including loss of cell adhesion and cell mobility) [45]. We have shown here that in contrast to cancerous tissue, endometriotic tissue also expresses genes acting against EMT and proliferation. Several tumor-suppressor genes are up-regulated in ovarian endometriosis: inhibitor of growth family, member 2 (*ING1L*), dickkopf homolog 3 (*DKK3*), inhibin, alpha (*INHA*), and Ras homolog gene family, member I (*ARHI*); homeo box C6 (*HOXC6*) is also up-regulated [46]. Frizzled homolog 7 (*FZDZ7*), which is involved in the activation of the Wnt/β-catenin pathway was up-regulated, and the wingless-type MMTV integration site family member 2 (WNT2), a ligand of the frizzled family of seven transmembrane receptors, was down-regulated in our study. These proteins are involved in tissue morphogenesis and/or differentiation, oncogenesis and several development processes. Also mRNA levels of transcription factor 21 (*TCF21*) and protein tyrosine phosphatase, receptor-type, Z polypeptide 1 (*PTPRZ1*) were higher in ectopic endometrium, and both *TCF21* and *PTPRZ1* act against EMT [47,48]. We have thus newly identified the deregulated genes: *ING1L*, *DKK3*, *INHA*, *ARHI*, *HOXC6*, *TCF21* and *PTPRZ1*.

4.5.3. Angiogenesis

In the endometrium, vascular endothelial growth factor (VEGF) is fundamental for angiogenesis. As previously described, and also seen in our study, mRNA levels of *VEGF* were lower in ovarian endometriosis compared to eutopic endometrium [49]. The two up-regulated genes encoding EGF-like repeats, discoidin I-like domains 3 (*EDIL3*) and sphingolipid G-protein-coupled receptor 1 (*EDG1*) identified in our study appear to be involved in this process. While EDIL3 is an integrin ligand and acts as an angiogenic factor in solid-tumor formation [50] and endothelial differentiation, EDG1 has an important role in the morphogenetic differentiation of vascular endothelial cells [51]. Our results thus point towards different mechanisms of vascularization in ovarian endometriosis.

4.5.4. The immune response and inflammation

The complement system is deregulated and increased inflammation is seen in peritoneal fluid of patients with ovarian endometriosis [52]. Expression of genes that encode the alternative pathway activation complement component 3 (C3), regulatory H factor 1 (HF1) [53] and the classical pathway component 4A (C4B) [53] were up-regulated in the ovarian endometriosis. We also observed higher mRNA levels of phospholipase A2, group IIA (*PLA2G2A*), vascular cell adhesion molecule 1 (*VCAM1*), sialoadhesin (*SN*), human rhinovirus receptor (*ICAM1*), and chemokine (C-X-C motif) ligand 1 (*CXCL1*) in endometriotic tissue. While high levels of *PLA2G2A*, *VCAM1* and *SN* have been previously associated with different inflammatory diseases [54–56], the CXC family members

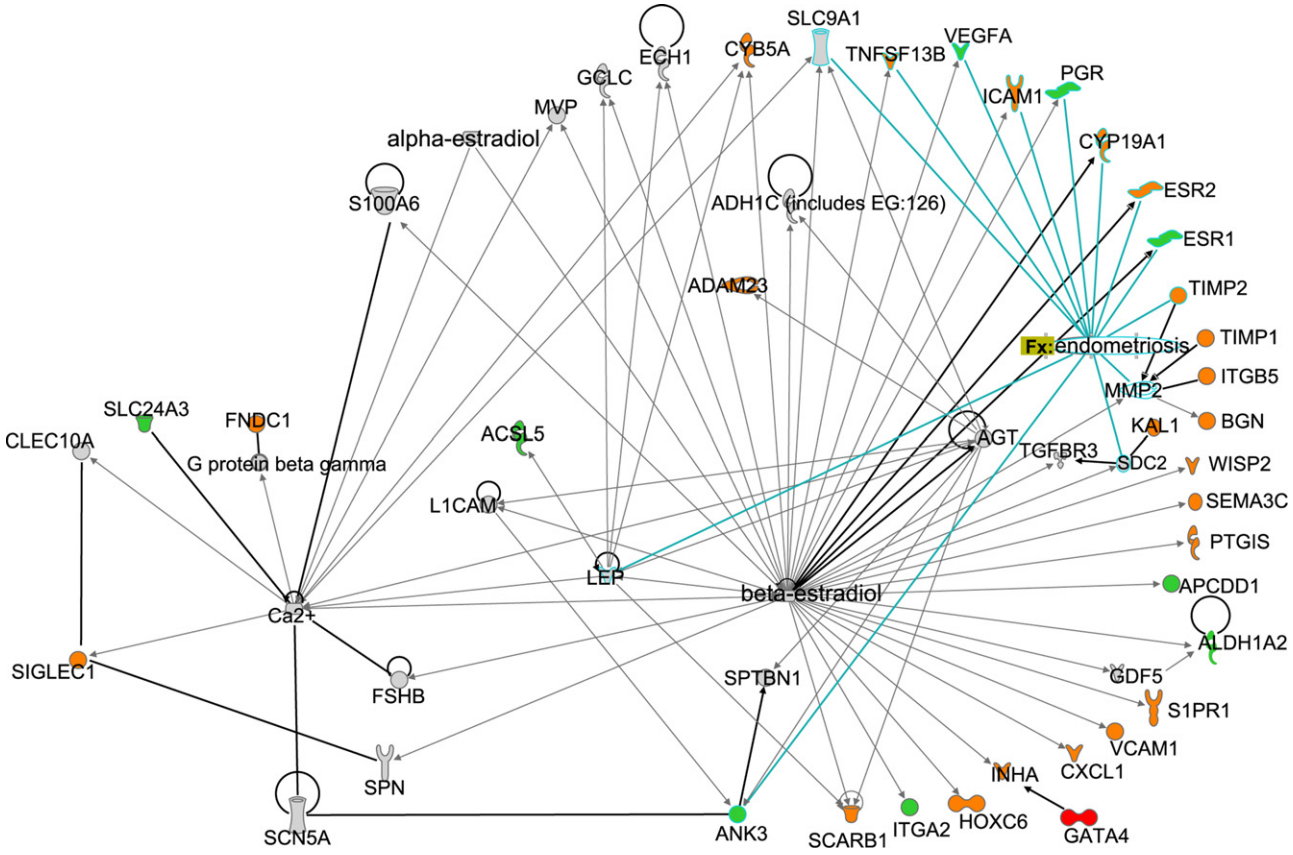

Fig. 4. Estrogen-related genes from the dataset that are differentially expressed in ovarian endometriosis. Graphical representation of the network of the molecular relationships between genes/gene products. The genes and gene products are shown as nodes, and the biological relationships between two nodes are shown as gray (indirect interaction) and black (direct interaction) lines. All gray lines are supported by at least one reference from the literature, a textbook, or canonical information stored in the Ingenuity pathways knowledge base. Human, mouse, and rat orthologs of a gene are stored as separate objects in the Ingenuity pathways knowledge base, but are designated as a single node in this network. The node color indicates the degree of regulation: green, down-regulation; orange, up-regulation; red, strong up-regulation. Genes previously reported as related to endometriosis are connected to the endometriosis node by solid blue lines. SIGLEC1 corresponds to SN and S1PR1 to EDG1 in the Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

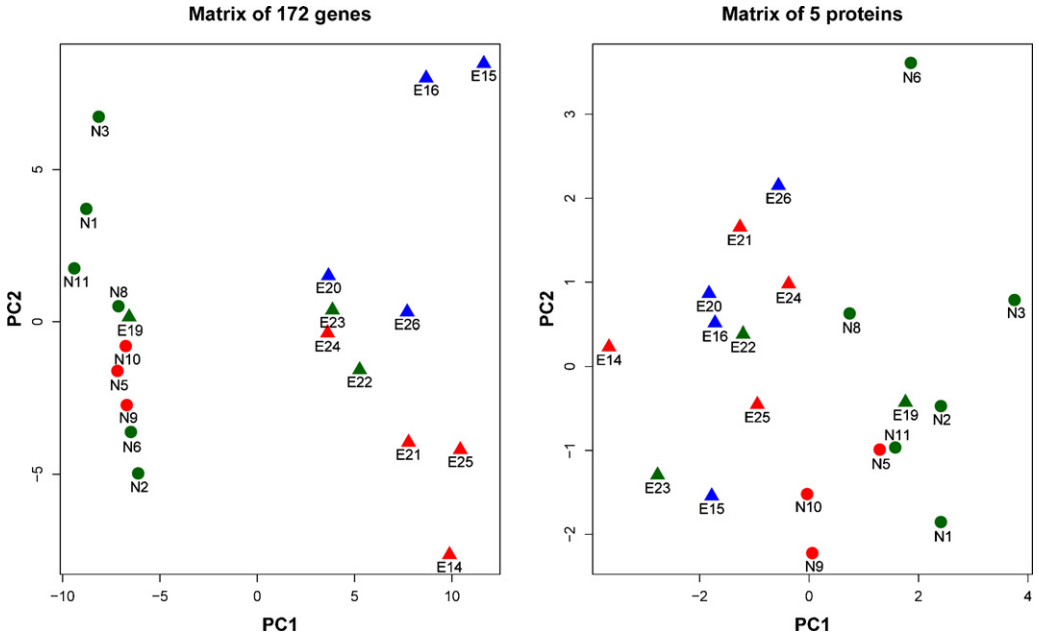

Fig. 5. Principal component analysis of the expression profiles. PCA was applied to a matrix of expression levels of 172 genes (left) or 5 proteins (right) in 20 subjects (11 ovarian endometriosis specimens and nine samples of control endometrium). Circles, control endometrium; triangles, endometriomas; red, proliferative phase; blue, late proliferative/early secretory phase; green, secretory phase. Along the PC1 axes, there is good clustering between the endometriosis group and the control endomerium group. There is also a certain level of clustering according to phases of the menstrual cycle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

have been reported to act as angiogenic factors [57]. Our results indicate that in ovarian endometriosis, the complement system is triggered through classical and alternative pathways, and is regulated via HF1.

4.6. Identification of novel biomarkers and drug targets

We have here confirmed aberrant expression of several genes that have been identified previously as differentially expressed in ovarian endometriosis, including *PLA2G2A*, *CPXM2*, *ESR2*, *C3*, *TNFSF13B*, *CXCL5*, *LAMA1*, *C4B*, *TIMP2*, *MMP26*, *PGR*, *ESR1*, *VEGF*, *ANK3*, *CALD1*, *DKK3*, *KLHDC8A*, *ICAM 1*, *WNT*, *SYNPO*, *TCF21* [9,15,29]. Proteins encoded by previously known and newly identified deregulated genes can serve as potential diagnostic markers or as targets for pharmaceutical development. It is important to note that our LDA analysis was a validation of microarray data; thus, the genes identified here have been selected by two methods. Literature searches have revealed that several of these proteins are found in peritoneal fluid or serum. For instance: ICAM1 has been detected in peritoneal fluid and serum, and elevated levels have been reported in women with endometriosis [33]; C3 and CXCL1 have been detected in peritoneal fluid of endometriosis patients [58]; and higher CXCL5 and VEGFA levels have been shown in serum of endometriosis patients, with the levels of VEGFA decreasing after surgical removal of the lesions [59].

Other extracellular proteins might serve as novel biomarkers as well, such as BGN, an ECM proteoglycan. BGN has a role in several processes, including cell death, growth, proliferation, movement and molecular transport, and in diseases, such as cancer and inflammatory disease (Table 4). BGN acts as an inhibitor of the activation of the classical pathway of complement activation, and of proinflammatory cytokine production [60]. However, in macrophages, BGN also acts as a proinflammatory, and signals through Toll-like receptors 4 and 2 [61]. BGN modulates inflammation in several ways; its effects appear to be dependant on the context, the site and the phase of the inflammatory process [61]. The expression of BGN at the protein level has not been reported previously in endometriosis, and despite the differences in BGN expression at the mRNA and protein levels, BGN might still be beneficial as a biomarker. In addition to potential novel biomarkers, our study also identified several genes encoding potential drug targets that have not been previously associated with endometriosis: up-regulated genes encoding enzymes (*PTPRZ1*, *PTGIS*, *PRSS35*, *RNASE1*), and an aberrantly expressed gene encoding a receptor (*MRC1*).

5. Conclusions

The LDA analysis in the present study represents the first large-scale validation of microarray data of ovarian endometriosis. The majority of expression analyses have validated the expression of several (up to 22) genes, while here we examined the expression of 152 genes previously identified by genome-wide microarray expression analysis. The 78 deregulated genes identified in our study, and especially the genes encoding secreted proteins or proteins of the ECM, thus represent a list of potential biomarkers and drug targets for diagnosis and treatment of ovarian endometriosis. Further studies are warranted to examine the expression of these genes at the protein level in peritoneal fluid and blood, where protein chip analysis or ELISA assays would be particularly useful. A selection of biomarkers based on serum analysis of patients and healthy subjects should eventually lead to the development of non-invasive diagnostic tools for patients with endometriosis.

Acknowledgements

This work was supported by grant J3-9448 to T.L.R. and a Young Researcher Grant to T.Š., both from the Slovenian Research Agency, and by Solvay Pharmaceuticals GmbH, Hannover, Germany. The authors are grateful to Neli Hevir for performing real-time PCR experiments and for her help with preparation of Fig. 4, and Dr. Guido Haarpaintner for his help with programming in R. The authors also thank Dr. Chris Berrie for critical reading of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsbmb.2011.03.010.

References

[1] V.M. Rice, Conventional medical therapies for endometriosis, Ann. N. Y. Acad. Sci. 955 (2002) 343–352.
[2] K.J. Berkley, A.J. Rapkin, R.E. Papka, The pains of endometriosis, Science 308 (2005) 1587–1589.
[3] L.C. Giudice, L.C. Kao, Endometriosis, Lancet 364 (2004) 1789–1799.
[4] A.W. Nap, P.G. Groothuis, A.Y. Demir, J.L.H. Evers, G.A.J. Dunselman, Pathogenesis of endometriosis, Best. Pract. Res. Clin. Obstet. Gynaecol. 18 (2004) 233–244.
[5] K.M. Eyster, A.L. Boles, J.D. Brannian, K.A. Hansen, DNA microarray analysis of gene expression markers of endometriosis, Fertil. Steril. 77 (2002) 38–42.
[6] T. Arimoto, T. Katagiri, K. Oda, T. Tsunoda, T. Yasugi, Y. Osuga, H. Yoshikawa, O. Nishii, T. Yano, Y. Takenati, Y. Nakamura, Genome-wide cDNA microarray analysis of gene-expression profiles involved in ovarian endometriosis, Int. J. Oncol. 22 (2003) 551–560.
[7] Q. Wu, S. Sucheta, S. Azhar, K.M.J. Menon, Lipoprotein enhancement of ovarian theca-interstitial cell steroidogenesis: relative contribution of scavenger receptor class B (type I) and adenosine 5′-triphosphate-binding cassette (type A1) transporter in high-density lipoprotein-cholesterol transport and androgen synthesis, Endocrinology 144 (2003) 2437–2445.
[8] W.P. Hu, S.K. Tay, Y. Zhao, Endometriosis-specific genes identified by real-time reverse transcription-polymerase chain reaction expression profiling of endometriosis versus autologous uterine endometrium, J. Clin. Endocrinol. Metab. 91 (2006) 228–238.
[9] B. Borghese, F. Mondon, J.C. Noël, I. Fayt, T.M. Mignot, D. Vainman, C. Chapron, Gene expression profile of ectopic vs. eutopic endometrium provides new insgights into endometriosis oncogenic potential, Mol. Endocrinol. 22 (2008) 2657–2662.
[10] M. Zafrakas, B.C. Tarlatzis, T. Streichert, F. Pournaropoulos, U. Wölfle, S.J. Smeets, B. Wittek, G. Grimbizis, R.H. Brakenhoff, K. Pantel, J. Bontis, C. Guenes, Genome-wide microarray gene expression, array-CGH analysis, and telomerase activity in advanced ovarian endometriosis: a high degree of differentiation rather than malignant potential, Int. J. Mol. Med. 21 (2008) 335–344.
[11] J. Meola, E. Rosa, J.C. Silva, D.B. Dentillo, W.A. da Silva Jr., L.C. Veiga-Castelli, L.A. de Souza Bernardes, R.A. Ferriani, C.C. Paro de Paz, S. Giuliatti, L. Martelli, Differentially expressed genes in eutopic and ectopic endometrium of women with endometriosis, Fertil. Steril. 93 (2009) 1750–1773.
[12] L. Mettler, A. Salmassi, T. Schollmeyer, A.G. Schmutzler, F. Puengel, W. Jonat, Comparison of c-DNA microarray analysis of gene expression between eutopic endometrium and ectopic endometrium (endometriosis), J. Assist. Reprod. Genet. 24 (2007) 249–258.
[13] L.C. Kao, A. Germeyer, S. Tulac, S. Lobo, J.P. Yang, R.N. Taylor, K. Osteen, B.A. Lessey, L.C. Giudice, Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility, Endocrinology 144 (2003) 2870–2881.
[14] S. Matsuzaki, M. Canis, C. Vaurs-Barrière, O. Boespflug-Tanguy, B. Dastugue, B. Mage, DNA microarray analysis of gene expression in eutopic endometrium from patients with deep endometriosis using laser capture microdissection, Fertil. Steril. 84 (2005) 1180–1190.
[15] R.O. Burney, S. Talbi, A.E. Hamilton, K.C. Vo, M. Nyegaard, C.R. Nezhat, B.A. Lessey, L.C. Giudice, Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis, Endocrinology 148 (2007) 3814–3826.
[16] H. Honda, F. Barrueto, J. Gogusev, D.D. Im, P.J. Morin, Serial analysis of gene expression reveals differential expression between endometriosis and normal endometrium. Possible roles for AXL and SHC1 in the pathogenesis of endometriosis, Reprod. Biol. Endocrinol. 6 (2008) 59.
[17] J.R.A. Sherwin, A.M. Sharkey, A. Mihalyi, P. Simsa, R.D. Catalano, T.M. D'Hooghe, Global gene analysis of late secretory phase, eutopic endometrium does not provide the basis for a minimally invasive test of endometriosis, Hum. Reprod. 23 (2008) 1063–1068.
[18] S. Matsuzaki, M. Canis, G. Mage, Use of laser capture microdissection in studying hormone-dependent diseases: endometriosis, Method. Mol. Biol. 590 (2009) 295–306.

[19] J.J. Randolph, Online Kappa Calculator. 2008. Retrieved November 12, 2009. http://justus.randolph.name/kappa.
[20] J. Vandesompele, K. DaPreter, F. Pattyn, B. Poppe, N. Van Roy, A. Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, Genome Biol. 3 (2002) 1–12.
[21] Y. Benjamini, D. Yekutieli, The control of the false discovery rate in multiple testing under dependency, Ann. Stat. 4 (2001) 1165–1188.
[22] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources, Nat. Protoc. 4 (2009) 44–57.
[23] G. Dennis Jr., B.T. Sherman, D.A. Hosack, J. Yang, W. Gao, H.C. Lane, R.A. Lempicki, DAVI.D. Database for annotation, visualization, and itegrated discovery, Genome Biol. 4 (2003) P3.
[24] O Bukulmez, D.B. Hardy, R.A. Carr, C.R. Mendelson, Inflammatory status influences aromatase and steroid receptor expression in endometriosis, Endocrinology 149 (2008) 1190–1204.
[25] P. Laudanski, J.O. Szamatowicz, O. Kowalczuk, M. Kuzmicki, M. Grabowicz, L. Chyczewski, Expression of selected tumor suppressor and oncogenes in endometrium of women with endometriosis, Hum. Reprod. 24 (2009) 1880–1890.
[26] R.H. Noyes, A.T. Hertig, J. Rock, Dating the endometrial biopsy, Am. J. Obstet. Gynecol. 122 (1975) 262–263.
[27] G.P. Skliris, A.T. Parkes, J.L. Limer, S.E. Burdall, P.J. Carder, V. Speirs, Evaluation of seven oestrogen receptor β antibodies for immunohistochemistry, western blotting, and flow cytometry in human breast tissue, J. Pathol. 197 (2002) 155–162.
[28] M. Izawa, M. Inoue, M. Osaki, H. Ito, T. Harada, N. Terakawa, M. Ikeguchi, Cytochrome, P450 aromatase gene (*CYP19*) expression in gastric cancer, Gastric Cancer 11 (2008) 103–110.
[29] Q. Xue, Z. Lin, Y.H. Cheng, C.C. Huang, E. Marsh, P. Yin, M.P. Milad, E. Confino, S. Reierstad, J. Innes, S.E. Bulun, Promoter methylation regulates estrogen receptor 2 in human endometrium and endometriosis, Biol. Reprod. 7 (2007) 681–687.
[30] S.E. Taylor, P.L. Martin-Hirsch, F.L. Martin, Oestrogen receptor splice variants in the pathogenesis of disease, Cancer Lett. 288 (2009) 133–148.
[31] M.C. Rodrigo, D.S. Martin, K.M. Eyster, Estrogen decreases biglycan mRNA expression in resistance blood vessels, Am. J. Physiol. Regul. Integr. Comp. Physiol. 285 (2003) R754–R761.
[32] S. Colette, J.C. Lousse, S. Defrère, M. Curaba, J.F. Heilier, A. Van Langendonckt, M. Mestdagt, J.M. Foidart, E. Loumaye, J. Donnez, Absence of aromatase protein and mRNA expression in endometriosis, Hum. Reprod. 24 (2009) 2133–2141.
[33] Y. Wu, A. Kajdacsy-Balla, E. Strawn, Z. Basir, G. Halverson, P. Jailwala, Y. Wang, X. Wang, S. Ghosh, S.W. Guo, Transcriptional characterizations of differences between eutopic and ectopic endometrium, Endocrinology 147 (2006) 232–246.
[34] C.A. Witz, Cell adhesion molecules and endometriosis, Semin. Reprod. Med. 21 (2003) 173–182.
[35] S. Cal, J.M.P. Freije, J.M. Lopez, Y. Takada, C. Lopez-Otin, ADAM23/MDC3, a human disintegrin that promotes cell adhesion via interaction with the αvβ3 integrin through an RGD-independent mechanism, Mol, Biol. Cell. 11 (2000) 1457–1469.
[36] H. Colognato, P.D. Yurchenco, Form and function: the laminin family of hetero trimers, Dev. Dyn. 218 (2000) 213–234.
[37] L. Tang, D.L. Dai, M. Su, M. Martinka, G. Li, Y. Zhou, Aberrant expression of collagen triple helix repeat containing 1 in human solid cancers, Clin. Cancer Res. 12 (2006) 3716–3722.
[38] S.Y. Moon, Y. Zheng, Rho GTPase-activating proteins in cell regulation, Trends Cell Biol. 13 (2003) 13–22.
[39] P.A. Regidor, C. Vogel, M. Redigor, A.E. Schindler, E. Winterhager, Expression pattern of integrin adhesion molecules in endometriosis and human endometrium, Hum. Reprod. Update 4 (1998) 710–718.
[40] L. Ramón, J. Gilabert-Estellés, R. Castelló, J. Gilabert, F. España, A. Romeu, M. Chirivella, J. Aznar, A. Estellés, mRNA analysis of several components of the plasminogen activator and matrix metalloproteinase system in endometriosis using real-time quantitative RT-PCR assay, Hum. Reprod. 20 (2005) 272–278.
[41] Y.G. Zhao, A.Z. Xiao, H.I. Park, R.G. Newcomer, M. Yan, Y.G. Man, Endometase/matrilysin-2 in human breast ductal carcinoma *in Situ* and its inhibition by tissue inhibitors of metalloproteinases-2 and -4: a putative role in initiation of breast cancer invasion, Cancer Res. 64 (2004) 590–598.
[42] A.R. Albig, J.R. Neil, W.P. Schiemann, Fibulins 3 and 5 antagonise tumor angiogenesis *In vivo*, Cancer Res. 66 (2006) 2621–2629.
[43] E. Schönherr, P. Witsch-Prehm, B. Harrach, H. Robenek, J. Rauterberg, H. Kresse, Interaction of biglycan with type I collagen, J. Biol. Chem. 270 (1995) 2776–2783.
[44] D. Schuppan, C.M. Cantaluppi, J. Becker, A. Veit, T. Buntel, D. Troyer, Undulin, an extracellular matrix glycoprotein associated with collagen fibrils, J. Biol. Chem. 265 (1990) 8823–8832.
[45] E.W. Thompson, D.F. Newgreen, Carcinoma invasion and metastasis: a role for epithelial–mesenchymal transition, Cancer Res. 65 (2005) 5991–5995.
[46] C. Abate-Shen, Deregulated homeobox gene expression in cancer: cause or consequence? Nat. Rev. Cancer 2 (2002) 777–785.
[47] O.P. Smith, S. Battersby, K.J. Sales, H.O.D. Critchley, H.N. Jabbour, Prostacyclin receptor up-regulates the expression of angiogenic genes in human endometrium via cross talk with epidermal growth factor receptor and the extracellular signalling receptor kinase ½ pathway, Endocrinology 147 (2006) 1697–1705.
[48] P. Perez-Pinera, S. Alcantara, T. Dimitrov, J.A. Vega, T.F. Deuel, Pleiotrophin disrupts calcium-dependent homophilic cell–cell adhesion and initiates an epithelial–mesenchymal transition, in: Proc. Natl. Acad. Sci. U.S.A., 103, 2006, pp. 17795–17800.
[49] X.J. Tan, J.H. Lang, D.Y. Liu, K. Shen, J.H. Leng, L. Zhu, Expression of vascular endothelial growth factor and trombospondin-1 mRNA in patients with endometriosis, Fertil. Steril. 78 (2002) 148–153.
[50] Y. Aoka, F.L. Johnson, K. Penta, K. Hirata, C. Hidai, R. Schatzman, J.A. Varner, T. Quertermous, The embryonic angiogenic factor Del1 accelerate tumor growth by enhancing vascular formation, Microvasc. Res. 64 (2002) 148–161.
[51] T. Hla, T. Maciag, An abundant transcript induced in differentiating human endothelial cells encodes a polypeptide with structural similarities to G-protein-coupled receptors, J. Biol. Chem. 265 (1990) 9308–9313.
[52] C.M. Kyama, S. Debrock, J.M. Mwenda, T.M. D'Hooghe, Potential involment of the immune system in the development of endometriosis, Reprod. Biol. Endocrinol. 1 (2003) 123.
[53] J.M. Thurman, V.M. Holers, The central role of the alternative complement pathway in human disease, J. Immunol. 176 (2006) 1305–1310.
[54] L. Touqui, M. Alaoui-El-Azher, Mammalian secreted phospholipase A_2 and their pathophysiological significance in inflammatory disease, Curr. Mol. Med. 1 (2001) 739–754.
[55] M.P. Bevilacqua, R.M. Nelson, G. Mannori, O. Cecconi, Endothelial-leukocyte adhesion molecules in human disease, Ann. Rev. Med. 45 (1994) 361–378.
[56] P.R. Crocker, A. Hartnell, J. Munday, D. Nath, The potential role of sialoadhesin as a macrophage recognition molecule in health and disease, Glycoconj. J. 14 (1997) 601–609.
[57] J.A. Belperio, M.P. Keane, D.A. Arenberg, C.L. Addison, J.E. Ehlert, M.D. Burdick, R.M. Strieter, CXC chemokines in angiogenesis, J. Leukoc. Biol. 68 (2000) 1–8.
[58] S. Ferrero, D.J. Gillott, V. Remorgida, P. Anserini, N. Ragni, J.G. Grudzinskas, Proteomic analysis of peritoneal fluid in fertile and infertile women with endometriosis, J. Reprod. Med. 51 (2009) 32–40.
[59] V. Bourlev, N. Iljasova, L. Adamyan, A. Larsson, M. Olovsson, Signs of reduced angiogenic activity after surgical removal of deeply infiltrating endomemtriosis, Fertil. Steril. 94 (2009) 52–57.
[60] T.W. Groeneveld, M. Oroszlan, R.T. Owens, M.C. Faber-Krol, A.C. Bakker, R.T. Owens, G.J. Arlaud, D.J. McQuillan, U. Kishore, M.R. Daha, A. Roos, Interactions of the extracellular matrix proteoglycan decorin and biglycan with C1q and collectins, J. Immunol. 175 (2005) 4715–4723.
[61] A. Babelova, K. Moreth, W. Tsalastra-Greul, J. Zeng-Brouwers, O. Eickelberg, M.F. Young, P. Bruckner, J. Pfeilschifter, R.M. Schaefer, H.J. Gröne, L. Schaefer, Biglycan a danger signal that activates the NLRP3 inflammasome via toll-like and P2X receptors, J. Biol. Chem. 284 (2009) 24035–24048.