

Gene expression profile, pathways, and transcriptional system regulation in indolent systemic mastocytosis

M. Niedozytko^{1,2}, J. N. G. Oude Elberink², M. Bruinenberg^{3,4}, B. Niedozytko⁵, J. G. R. de Monchy², G. J. te Meerman³, R. K. Weersma⁶, A. B. Mulder⁷, E. Jassem¹ & J. J. van Doormaal²

¹Department of Allergology Medical University of Gdansk, Gdansk, Poland; ²Department of Allergology, University Medical Center Groningen, University of Groningen; ³Department of Genetics University Medical Center Groningen, University of Groningen; ⁴Lifelines, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; ⁵Department of Dermatology Medical University of Gdansk, Gdansk, Poland; ⁶Department of Gastroenterology and Hepatology, University Medical Center Groningen, University of Groningen; ⁷Department of Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

To cite this article: Niedozytko M, Oude Elberink JNG, Bruinenberg M, Niedozytko B, de Monchy JGR, te Meerman GJ, Weersma RK, Mulder AB, Jassem E, van Doormaal JJ. Gene expression profile, pathways, and transcriptional system regulation in indolent systemic mastocytosis. *Allergy* 2010; DOI: 10.1111/j.1398-9995.2010.02477.x.

Keywords

gene expression; gene profiling; mastocytosis.

Correspondence

Marek Niedozytko, MD, PhD,
Department of Allergology Medical
University of Gdansk, Debinki 7 80-952
Gdansk, Poland.
Tel.: +48583491626
Fax: +48583491625
E-mail: mnied@amg.gda.pl

Accepted for publication 26 July 2010

DOI:10.1111/j.1398-9995.2010.02477.x

Edited by: Stephan Weidinger

Abstract

Background: Mastocytosis is an uncommon disease resulting from proliferation of abnormal mast cells infiltrating skin, bone marrow, liver, and other tissues. The aim of this study was to find differences in gene expression in peripheral blood cells of patients with indolent systemic mastocytosis compared to healthy controls. The second aim was to define a specific gene expression profile in patients with mastocytosis.

Methods: Twenty-two patients with indolent systemic mastocytosis and 43 healthy controls were studied. Whole genome gene expression analysis was performed on RNA samples isolated from the peripheral blood. For amplification and labelling of the RNA, the Illumina TotalPrep 96 RNA Amplification Kit was used. Human HT-12_V3_expression arrays were processed. Data analysis was performed using GENESPRING, Genecodis, and Transcriptional System Regulators.

Results: Comparison of gene expression between patients and controls revealed a significant difference ($P < 0.05$ corrected for multiple testing) and the fold change difference > 2 in gene expression in 2303 of the 48,794 analysed transcripts. Functional annotation indicated that the main pathways in which the differently expressed genes were involved are ubiquitin-mediated proteolysis, MAPK signalling pathway, pathways in cancer, and Jak-STAT signalling. The expression distributions for both groups did not overlap at all, indicating that many genes are highly differentially expressed in both groups.

Conclusion: We were able to find abnormalities in gene expression in peripheral blood cells of patients with indolent systemic mastocytosis and to construct a gene expression profile which may be useful in clinical practice to predict the presence of mastocytosis and in further research of novel drugs.

Mastocytosis is an uncommon disease resulting from proliferation of abnormal mast cells in different tissues including skin, bone marrow, liver, spleen, and lymph nodes (1). One of the key elements in the pathogenesis of the disease is the presence of a specific *KIT* mutation in mastocytes but also in other peripheral blood cells (1, 2). The clinical presentation of mast-

ocytosis is heterogenous, varying from solely skin presentation found in *urticaria pigmentosa* and *mastocytoma*, to different forms of systemic disease including indolent systemic mastocytosis, smouldering systemic mastocytosis, aggressive systemic mastocytosis, and mast cell leukaemia (1). Of the adult patients with systemic mastocytosis, the large majority (ca. 90%) have the indolent form of the disease. Most symptoms (like anaphylaxis, hypotension, urticaria, and diarrhoea) are related to mast cell infiltration and degranulation (1). However, other symptoms such as osteoporosis, hypertension,

Abbreviations

TSR, transcriptional system regulators; FA, factor analysis; NB, Naïve Bayes prediction; FC, fold change.

pain syndromes, and neurological symptoms are only partially understood and may involve other mechanisms (1).

The mechanism(s) involved in the development of mastocytosis are mainly unknown (1). Activation of kinase pathways (i.e. D816V mutation of *KIT*), and IL-13 and rIL-4 polymorphisms have previously been shown to be relevant in this respect (3, 4). It seems quite likely that more than one pathway maybe located in other cells than mast cells alone are involved in the development of the disease. The studies by Garcia-Montero (2) showed that the *KIT* mutation is present not only in mast cells but also in myeloid and lymphoid cell lineages. Furthermore, the presence of the *KIT* mutation in various cell lineages was related to the prognosis of mastocytosis (2).

The aim of this study was to find differences in gene expression in peripheral blood cells of patients with indolent systemic mastocytosis compared to healthy controls. The second aim was to define a specific gene expression profile in patients with mastocytosis.

Methods

Patients

A total of 22 Caucasian patients with indolent systemic mastocytosis from the Department of Allergology, University Medical Center Groningen (UMCG) were studied [median age 50 range 35–73 years; 7 (31%) men and 15 (68%) women]. All patients underwent standard diagnostic procedures based on WHO guidelines for the workup of systemic mastocytosis including bone marrow histopathological, cytological, and flow cytometric (CD2, CD25) examinations of bone marrow. To detect the *KIT* D816V mutation, we used two different techniques in time. Initially, RNA was isolated from EDTA anti-coagulated bone marrow cells with the help of the QIAamp®RNA Blood MINI Kit (Qiagen, Westburg, Leusden, the Netherlands). The Promega Reverse Transcriptase kit (Promega Benelux, Leiden, the Netherlands) was used to synthesize c-DNA from approximately 500 ng RNA. The resulting c-DNA was amplified using previously described primers with the following PCR conditions: 30 cycles of denaturation (1 min at 95°C), annealing (1 min at 61°C), and extension (2 min at 72°C), followed by 7 min at 72°C and subsequent cooling (5). The resulting 346-bp PCR product was digested with the help of Hae III and Hinf I (BioLabs, Westburg, Leusden, the Netherlands), resulting in restriction fragments of 171, 127, and 48 (not detected) base pairs to detect the wild type and 157, 127, 48 (not detected), and 14 (not detected) base pairs to detect the Asp 816Val mutation. The restriction fragments were separated on a 6% agarose Multi purpose (Roche, Almere, the Netherlands) gel and visualized using ethidium bromide (patient no. 1, 3, 5, 7, 9, 14, 15, 16, 17, 20, and 22). From December 2007, detection of the *KIT* D816V mutation was performed with a real-time qPCR using previously published (6) primers 5'-TTGTGATTTTGGTCTAGCCAGACT-3' and 5'-GTGC-CATCCACTTCACAGGTAG-3' (patient no. 2, 4, 8, 11, 12, 18, and 19). Urinary histamine metabolites and serum

tryptase measurements were also taken (Table 1) (1). A group of 43 healthy Caucasian subjects [median age 50 range 19–73 years, 22 (51%) men and 21 (49%) women] were used as controls. They were nonrelated partners from patients with inflammatory bowel disease visiting the outpatient department of the inflammatory bowel disease unit of the UMCG.

The study was approved by the Medical Ethical Committee of the UMCG (METc 2008/340).

Collection of blood samples

PAXgene blood RNA tubes (Qiagen, Valencia, CA, USA) were used for RNA sampling. All tubes were immediately frozen and stored in –20°C till RNA isolation (maximal period 2 months). RNA was isolated using PAXgene blood RNA Kit CE (Qiagen, Venlo, the Netherlands). All RNA samples were stored in –80°C till labelling and hybridization.

The quality and concentration of RNA were determined using 2100 Bioanalyzer (Agilent, Amstelveen, the Netherlands) and the Agilent RNA 6000 Nano Kit. Samples with RNA integrity number > 7.5 were used for further analysis on expression arrays.

Gene expression

For amplification and labelling of the RNA with the Illumina TotalPrep 96 RNA Amplification Kit (Applied Biosystems, Nieuwerkerk ad IJssel, the Netherlands), 200 ng of RNA from each sample was used. The human HT-12_V3_expression arrays (Illumina, San Diego, CA, USA) were processed according to the manufacturer's protocol. Slides were scanned immediately using Illumina BeadStation iScan (Illumina).

Image and data analysis

First line check, background correction and quantile normalization of the data were carried out with GENOMESTUDIO Gene Expression Analysis module v 1.0.6 Statistics (San Diego, CA, USA). Entities of which at least 75% of the samples had a signal intensity value above 20th percentile in 100% of the samples of at least two groups were included for further analysis.

Data analysis was performed using GENESPRING package version 8.0.0 (Agilent Technologies Santa Clara, CA, USA). Genes of which expression was significantly different between the compared groups were chosen based on a log₂fold change > 2 in gene expression, *t*-test *P*-value < 0.05 and corrected for multiple testing by the Benjamin–Hochberg method. The naïve Bayes prediction model was used to build a prediction model which might be used in diagnosis of mastocytosis (7). Naïve Bayesian classifier assumes that the impact of single gene expression is unrelated to other genes in the prediction model. The method does not take into account the interactions of the genes composing the model or gene environmental interactions.

Functional annotation of genes was described using GeneCodis (8), functional annotation web-based tool using KEGG pathways (9) and GoSlim process analysis.

Table 1 Demographic and clinical data of the individual patients with indolent systemic mastocytosis

Patient no.	Gender	Age at diagnosis (years)	UP	Serum tryptase ($\mu\text{g/l}$)	Urine MH ($\mu\text{mol/mol creat}$)	Urine MIMA (mmol/mol creat)	MCs in bone marrow aspirate (%)	CD2 immunophenotype	CD25 immunophenotype	D816V K/T mutation in bone marrow cells	≥ 2 aggregates of ≥ 15 MCs in bone marrow	Abnormal morphology of $\geq 25\%$ of MCs in bone marrow	Histological bone marrow cellularity
1	M	50	–	28.9	153	1.4	0.09	–	+	–	–	+	Normal
2	F	64	Adult UP	155	1046	10.9	0.49	+	+	+	+	+	Normal
3	F	43	Adult UP	34.1	380	4.8	0.29	–	+	+	+	+	Normal
4	F	64	Adult UP	21.7	166	3.1	0.11	+	+	+	–	+	Normal
5	F	47	–	48.3	604	6.5	0.17	+	+	+	+	+	Normal
6	F	25	Adult UP	27.4	1024	3.4	n.d.	n.d.	n.d.	n.d.	+	+	Normal
7	F	50	Adult UP	20.4	293	3.0	0.06	+	+	+	+	+	Normal
8	M	34	Adult UP	55.4	433	4.9	0.45	+	+	+	+	+	Normal
9	F	61	–	15.2	266	2.3	0.30	–	+	–	+	+	Normal
10	M	41	Juvenile UP	146	470	4.0	1.61	–	+	+	+	+	normal
11	F	64	Adult UP	4.62	102	1.6	0.10	+	+	+	–	+	Normal
12	M	54	–	28.2	77	1.4	0.09	+	+	+	–	+	Normal
13	M	50	Adult UP	52.3	530	6.4	n.d.	n.d.	n.d.	n.d.	+	+	Normal
14	F	43	Adult UP	36.5	851	5.9	0.34	+	+	+	+	+	Normal
15	F	51	–	74.5	194	2.7	0.10	+	+	+	+	+	Normal
16	F	38	Adult UP	5.13	190	1.7	0.13	+	+	–	+	+	Normal
17	M	38	–	44.4	310	3.4	n.d.	n.d.	n.d.	+	+	+	Normal
18	M	37	Adult UP	112	404	7.7	n.d.	n.d.	n.d.	+	+	+	Normal
19	F	36	–	31.3	542	6.5	0.17	+	+	+	+	+	Normal
20	F	70	–	19.7	289	2.7	0.10	+	+	+	–	+	Slightly hyperplastic
21	F	63	Adult UP	109	823	6.3	n.d.	n.d.	n.d.	n.d.	+	+	Normal
22	F	54	Adult UP	29.8	593	4.1	n.d.	n.d.	n.d.	+	+	+	Normal

M, male; F, female; n.d., not done; UP, urticaria pigmentosa; juvenile UP: juvenile-onset UP; adult UP, adult-onset UP; MH, methylhistamine; MIMA, methylimidazole acetic acid, MCs, mast cells; –, negative; +, positive.

Differences in gene expression between patients with mastocytosis and healthy controls were also analysed using transcriptional system regulators (TSR) and factor analysis (FA) described by Fehrmann et al. (10). This method uses principal components derived from the correlation between expressed genes in 15,000 Affymetrix expression arrays.

The gene specific weights were applied to normalized log transformed Illumina transcript data, using the average for every gene, if more than one transcript was available. This procedure was performed for the first 50 principal components identified by Fehrmann et al. (10), resulting in a new set of 50 data points for each person. Subsequently, FA was performed on the component scores, further reducing the set of data points per person to eight explaining 75% of the variance of the 50 original principal component scores. The correlation between the factor scores is caused by the much lower heterogeneity of the data in comparison with the 15,000 arrays used by Fehrmann et al. The aim of this method is to use the correlation structure between genes to find scores that have a higher reproducibility because the signal of many genes is added. The biological interpretation of the factors is derived from those genes that have the strongest contribution to the compound score. Gene related specificity is lost, but problems with overfitting and low reliability of individual gene signals are strongly reduced. Factor analysis was performed with SYSTAT 12.0 (San Jose, CA, USA) and component scores were computed with a computer program written in DELPHI 5.0 (Austin, TX, USA) available on demand from GTM.

Power calculation to find differences in expression is difficult to compute *a priori*, as we have no knowledge of the impact of systemic mastocytosis on expression. Considering that the phenotype has a substantial impact on health, we assumed that even with a small number of individuals compared, significant differences between cases and controls would be present even after correction for multiple testing. The analysis based on metagenes is more sensitive than single gene analysis as signals from many genes are combined and errors cancelled out.

Clinical data were analysed with STATISTICA 8.0 (StatSoft, Tulsa, OK, USA).

Results

Whole genome gene expression analysis was performed on RNA samples isolated from all blood cells in whole blood. From all 48,804 probes present in the array, 48,794 transcripts had sufficient data for further analysis.

Comparison of gene expression profiles between patients and controls revealed a significant difference in 5086 of the analysed transcripts. A fold change difference >2 in gene expression was found in 2330 of those transcripts among which 1951 (84%) were upregulated and 379 (16%) downregulated. Functional annotation indicated that the main pathways in which the differently expressed genes were involved are ubiquitin-mediated proteolysis, MAPK signalling pathway, pathways in cancer, Jak-STAT signalling, and p53 signalling pathway (Table 2). The most important processes influenced by mastocytosis are transcription, cell cycle, protein transport, and signal transduction (Table 3).

We matched 13,032 transcripts with Affymetrix genes, using official gene symbol agreement, and used in the factor score analysis. Split-half correlations were computed for each of the 50 factor scores as an indication of independence of factor scores of individual genes.

Among the 50 TSRs described by Fehrmann et al. (10), the TSRs 1, 2, 4, 5, 6, 7, 8, 10, 12, 13, 38, 46, 49, and 50 were most different between patients and controls. In a second FA, two uncorrelated (orthogonal) factors (nr 2 and 4) were identified that both differentiated between cases and controls. The factors 2 and 4 provided the best discriminative properties of predicting the presence of mastocytosis. The main function indicated by the TSRs and KEGG pathway (9) analysis are MAPK signalling pathway, focal and cell adhesion, calcium signalling pathway, neuroactive ligand-receptor interaction, ribosome, cytokine-cytokine receptor interactions, regulation of actin cytoskeleton, and oxidative phosphorylation.

Using leucocyte-specific transcripts described by Liu et al. (11), we analysed the expression profiles of the leucocyte-specific genes characteristic for dendritic cells, B cells, effector memory T cells, mast cells, and basophils. Statistically significant differences in expression between patients and controls

Table 2 Gene co-occurrence annotation found by Genecodis (KEGG pathways) for the genes differentially expressed ($FC > 2$, $P < 0.05$ corrected for multiple testing) between patients with indolent systemic mastocytosis and healthy controls. P -values have been obtained through hypergeometric analysis (Hyp) corrected by FDR method (Hyp*) NGR – number of annotated genes in the reference list, NG – number of annotated genes in the input list (14, 15)

Genes	NGR	NG	Hyp	Hyp*	Annotations
27 genes	129 (37435)	27 (1769)	5.93719e-11	1.82272e-08	(KEGG) 04120 :Ubiquitin-mediated proteolysis
35 genes	262 (37435)	35 (1769)	3.3285e-08	5.10924e-06	(KEGG) 04010 :MAPK signalling pathway
37 genes	320 (37435)	37 (1769)	5.6426e-07	5.77426e-05	(KEGG) 05200 :Pathways in cancer
22 genes	150 (37435)	22 (1769)	2.38024e-06	0.000182683	(KEGG) 04630 :Jak-STAT signalling pathway
14 genes	67 (37435)	14 (1769)	2.43652e-06	0.000149602	(KEGG) 04115 :p53 signalling pathway
18 genes	108 (37435)	18 (1769)	3.05782e-06	0.000156458	(KEGG) 04110 :Cell cycle
15 genes	87 (37435)	15 (1769)	1.3072e-05	0.000573302	(KEGG) 04210 :Apoptosis
20 genes	145 (37435)	20 (1769)	1.70886e-05	0.000655773	(KEGG) 00230 :Purine metabolism
20 genes	150 (37435)	20 (1769)	2.8221e-05	0.000962649	(KEGG) 04310 :Wnt signalling pathway

Table 3 Gene co-occurrence annotation found by Genecodis (14, 15) (GOSlim Process Function) for the genes differentially expressed ($FC > 2$, $P < 0.05$ corrected for multiple testing) between patients with indolent systemic mastocytosis and healthy controls. P -values have been obtained through hypergeometric analysis (Hyp) corrected by FDR method (Hyp*) NGR – number of annotated genes in the reference list, NG – number of annotated genes in the input list

Genes	NGR	NG	Hyp	Hyp*	Annotations
190 genes	1516 (37 435)	190 (1769)	3.22857e–35	1.51743e–33	GO:0006350 :transcription (BP)
61 genes	376 (37 435)	61 (1769)	3.51927e–17	8.27027e–16	GO:0007049 :cell cycle (BP)
53 genes	376 (37 435)	53 (1769)	1.34649e–12	2.10951e–11	GO:0015031 :protein transport (BP)
134 genes	1700 (37 435)	134 (1769)	4.64627e–09	5.45937e–08	GO:0007165 :signal transduction (BP)
51 genes	471 (37 435)	51 (1769)	3.90706e–08	3.67264e–07	GO:0008152 :metabolic process (BP)
18 genes	99 (37 435)	18 (1769)	8.26542e–07	6.47458e–06	GO:0006950 :response to stress (BP)
47 genes	505 (37 435)	47 (1769)	9.00404e–06	6.04557e–05	GO:0006810 :transport (BP)
21 genes	177 (37 435)	21 (1769)	0.00010213	0.000600016	GO:0006464 :protein modification process (BP)

were found for the following genes expressed in mast cells: *ATP6VOA1*, *LOC348262*, *RFESD*, *OSBPL6*, T cells: *GALK2*, *IL32*, *KLF12*, *IL12A*, *SOS1* dendritic cells: *C2orf64*, *CD1B*, *ZFP3* and B cells: *MEF2C*, and *MS4A1*. The genes identified by D'Ambrosio et al. (12) studying gene expression analysis of bone marrow mononuclear cells found similar changes in expression in four of 10 described genes (*CPA3*, *GATA2*, *KIT*, and *MAF*).

We subsequently went on to build the prediction model which could be used in diagnosing indolent systemic mastocytosis based on the gene expression in the peripheral blood cells. We built the prediction model using a Naïve Bayes classifier based on the most discriminative 29 genes with $P < 10^{-10}$ corrected for multiple testing (Table 4). The sensitivity of this predictive model was 100% with a specificity of 97%, a negative predictive value of 100%, and a positive predictive value of 96%. Clustering analysis divided those genes into two clusters based on the similarities in gene expression pattern (Fig. 1).

Discussion

The results of the present study show strong and very significant differences in gene expression in peripheral blood cells between patients with indolent systemic mastocytosis and healthy controls. Additionally, 29 gene expression profile differentiating patients from controls was created based on the most differently expressed transcripts.

We were able to show that gene expression differences are found in other cells than mast cells solely. It confirms the finding that expression effects of the specific *KIT* mutation of mastocytosis may be found in other cell lineages than mast cells. In addition, it shows a difference in expression of 2330 other transcripts (3).

We also used TSR profiling in the data analysis. The regulation of transcription is a complex mechanism; however, the overlap between diseases and tissues analysed led to the conclusion that a large part of differences in transcription may be explained by a network of co-regulated gene clusters. The studies by Fehrmann et al. (10) showed that the number of orthogonal factors needed to explain most of the variability in expression may be limited to 50

TSRs. Studies made on 17,550 human microarray experiments led to identify 50 TSRs capturing 64% of the variability in gene expression (10). The TSR analysis in our study showed profound transcriptosome abnormalities in patients with mastocytosis. The results of the gene expression analysis made in GENESPRING and by the TSR method indicate similar processes and pathways involved in mastocytosis. The biological complexities of these systems suggest networks of co-regulated genes. The results of the TSR and FA analysis suggest, that the abnormalities in gene expression in indolent systemic mastocytosis are related to biological processes also found in other diseases. This approach, which analyses transcriptional mechanisms common across tissues and diseases, allows analysing of gene expression in whole blood without cell sorting and reduces the probability of finding gene expression patterns related to the experimental conditions and sample studied. The functional analysis reveals processes responsible for neoplastic cell transformation (pathways in cancer, MAPK, Jak-STAT signalling, p53 signalling pathway, cell cycle, and apoptosis). The finding of abnormally expressed genes and pathways may also lead to the application of novel drugs in systemic mastocytosis.

In the next step, we went on to create a gene expression profile which could be of help in diagnosing patients suffering from indolent systemic mastocytosis. We suggested a set of 29 most differently expressed genes divided in two clusters according to the pattern of expression (Fig. 1).

The genes composing cluster 1 were described previously in the pathogenesis of both solid tumours and haematological malignancies. We observed both upregulation of proto-oncogenes and downregulation of tumour suppressor genes. Three of the genes were described in lung cancer, namely *ZMAT3* (p53 target zinc finger protein) (13), arp2 actin-related protein 2 (*ACTR2*) (14), and cholinergic receptor, nicotinic alpha 5 (*CHRNA5*) (15), and two others in breast cancer, namely rho gtpase-activating protein 8 (*PRR5*) (16) and plectin 1 (*PLEC1*) (17). Plectin 1 was also described in ovarian cancer (17) and *PRR5* (rho gtpase-activating protein 8) (16) in colorectal cancer. Four other genes were also described in acute myeloid leukaemia, namely integrin beta 1 (*ITGB1*) (18), ataxia telangiectasia mutated (*ATM*) (19), and

Table 4 The list of the 29 most significantly different expressed genes ($FC > 2$, $P < 0.00000000001$ corrected for multiple testing by Benjamin Hochberg method) between patients with indolent systemic mastocytosis and healthy controls

Gene symbol	Gene name	FC M/C	FC C/M	Correlation	P	Gene function
RAB27A	rab27a, member ras oncogene family	0.39	2.58	3e-8	1.3e-11	Neutrophil secretion and shape (26, 27)
ETS1	v-ets erythroblastosis virus e26 oncogene homologue 1 (avian)	2.32	0.43	1.2e-8	3e-12	Carcinogenesis (20)
LOC730358		3.13	0.32	7.5e-8	6e-11	Unknown
ITGB1	Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen cd29 includes mdf2, msk12)	2.69	0.37	5.2e-8	3.4e-11	Carcinogenesis (18)
ARL16	adp-ribosylation factor-like 16	0.38	2.66	1.4e-9	1.7e-11	Signal transduction (8, 9)
LRAP	Leucocyte-derived arginine aminopeptidase	0.47	2.11	5.3e-11	3.3e-15	Carcinogenesis (22)
MLL3	Myeloid/lymphoid or mixed-lineage leukaemia 3	5.07	0.20	1.9e-8	6.2e-12	Carcinogenesis (28)
PLEC1	Plectin 1, intermediate filament binding protein 500 kDa	0.42	2.37	7.8e-8	6.6e-11	Carcinogenesis (17)
FAM39DP		0.48	2.09	1.2e-8	2.4e-12	Unknown
HSPC268	Hypothetical protein hspc268	0.37	2.69	5.2e-8	3.6e-11	Unknown
C3ORF34	Chromosome 3 open reading frame 34	0.41	2.44	3.4e-8	2e-11	Unknown
SERTAD2	Serta domain containing 2	2.02	0.49	1.2e-8	3.2e-12	Carcinogenesis (25)
ITGAV	Integrin, alpha v (vitronectin receptor, alpha polypeptide, antigen cd51)	4.29	0.23	9.2e-8	7.9e-11	Carcinogenesis (30)
SIAH1	Seven in absentia homologue 1 (drosophila)	2.06	0.49	1.2e-8	3.4e-12	Carcinogenesis (21)
CCR2	Chemokine (c-c motif) receptor 2	4.53	0.22	7.2e-8	5.6e-11	Carcinogenesis (29)
LRRFIP1	Leucine-rich repeat (in flil) interacting protein 1	0.41	2.42	2.9e-8	1.2e-11	Carcinogenesis (23)
C9ORF72	Hypothetical protein flj11109	5.23	0.19	5.2e-8	3.5e-11	Unknown
ATM	Ataxia telangiectasia mutated (includes complementation groups a, c and d)	2.74	0.37	9.7e-8	8.9e-11	Carcinogenesis (19)
PTP4A2	Protein tyrosine phosphatase type iva, member 2	0.47	2.14	2.6e-9	3.7e-13	Carcinogenesis (8, 9)
NCOA2	Nuclear receptor coactivator 2	7.36	0.14	6.6e-11	5.4e-15	Carcinogenesis (29)
LBR	Lamin B receptor	4.36	0.23	3e-8	1.4e-11	Granulopoiesis maturation of neutrophils (31)
MATR3	Matrin 3	4.80	0.21	9.7e-8	8.8e-11	Transcription (33)
ACTR2	arp2 actin-related protein 2 homologue (yeast)	2.14	0.47	6.8e-8	5.1e-11	Carcinogenesis (14)
PRR5	rho gtpase-activating protein 8	0.50	2.01	7.8e-8	6.6e-11	Carcinogenesis (16)
CHRNA5	Cholinergic receptor, nicotinic, alpha 5	0.41	2.42	3.9e-11	1.6e-15	Carcinogenesis (15)
ZMAT3	p53 target zinc finger protein	0.46	2.18	1e-11	7.2e-12	Carcinogenesis (13)
ZFAND5	Zinc finger, a20 domain containing 2	2.51	0.40	3.6e-8	2.2e-11	Carcinogenesis (8, 9)
SGPP1	Sphingosine-1-phosphate phosphatase 1	5.07	0.20	2.2e-8	8.8e-12	Cell proliferation, apoptosis (32)
C14ORF153	Chromosome 14 open reading frame 153	0.38	2.65	4.4e-10	4.5e-14	Unknown

v-ets erythroblastosis virus e26 oncogene homologue 1 (*ETS1*) (20) and seven in absentia homologue (*SIAH1*) (21). Leucocyte-derived arginine aminopeptidase (*LRAP*) plays a role in the development of lymphoma (22). Leucine-rich repeat interacting protein (*LRRFIP1*) contributes to the pathology of myelodysplastic syndrome (23) and glioblastoma (24). Multiple tumours including lymphomas and solid tumours are related to overexpression of *SERTAD2* (serta domain containing) (25). *RAB27A* gene product is a protein member of the ras oncogene family involved in neutrophil secretion (26) and melanocyte shape (27).

All genes in cluster 2 were upregulated in mastocytosis. Myeloid/lymphoid or mixed-lineage leukaemia (3*MLL3*) (28), nuclear receptor coactivator 2 (*NCOA2*), and eosinophilic

leukaemia *CCR2* [chemokine (c-c motif) receptor 2] (29) also play a role in the pathology of myeloid leukaemia. Integrin alpha v (*ITGAV*) involvement was described in laryngeal and hypopharyngeal carcinomas (30). For the three genes lamin B receptor (*LBR*), *SGPP1*, and *MATR3* involvement in cancer was not described, but their function may contribute to carcinogenesis. Lamin B receptor plays a role in the morphological maturation of neutrophils and granulopoiesis (31). Sphingosine-1 phosphate phosphatase (*SGPP1*) is important in the regulation of cell proliferation, angiogenesis and apoptosis (32). *MATR3* plays a role in the regulation of transcription (33).

The analysis of gene expression is becoming a popular diagnostic method in neoplastic and inflammatory diseases

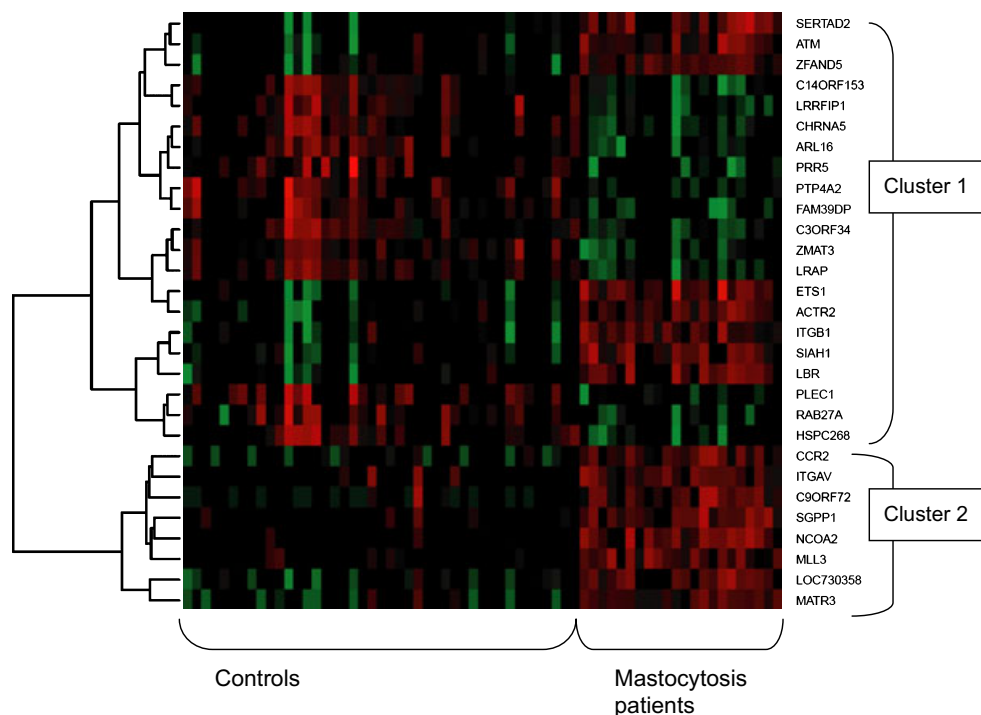


Figure 1 Hierarchical clustering dendrogram of 29 most differentially expressed genes between patients with mastocytosis and controls. Each column represents a patient sample, each row an

individual gene. For each gene, green colour represents underexpression, red colour overexpression, and a black signal denotes missing data.

(34). Also, gene profiling is a standard procedure in assessing the need for chemotherapy in breast cancer (35). To date, the diagnosis of systemic mastocytosis is based on WHO criteria (3–5). The large differences we observed between cases and controls suggest that a gene expression based test could be developed that would improve the reliability of current diagnostic methods. A potential role for the described gene profile may be in the differential diagnosis of patients with myeloproliferative or myelodysplastic disorders co-existing or masking mastocytosis, or in patients refusing bone marrow examination.

Furthermore, in contrast to the study by D'Ambrosio et al. (21), we analysed RNA isolated from whole blood without prior cell separation. This approach (1) reduces the effect of sample handling and (2) is a simple and standardized method which may be used in the clinical practice in the future, and furthermore (3) peripheral blood sampling is less of a burden to patients in comparison with a bone marrow biopsy. Additionally, RNA isolation and gene expression analysis used in this protocol have become standardized methods which avoid human laboratory errors and may be further adapted to clinical practice in the future.

Some aspects of our study warrant comment. A necessary next step is the validation of the abnormal gene profile in an independent group of patients with indolent systemic mastocytosis to confirm our findings, and in cases with other haematological diseases to see whether such RNA profiles are specific for mastocytosis. Further studies may also answer

the questions whether analysis of gene expression profiles may be used in clinical practice to (1) establish the diagnosis of mastocytosis and its clinical variants, (2) assess the risk of anaphylaxis in patients with mastocytosis and (3) assess the risk of progressive disease or to develop a non-mast cell haematological malignancy in these patients. The results of the present study, although limited, may open a new area of research.

In conclusion, we were able to find abnormalities in gene expression in peripheral blood cells of patients with indolent systemic mastocytosis and to construct a specific gene expression profile which may be useful in further research and possibly in clinical practice.

Acknowledgments

The authors would like to thank Professor Cisca Wijmenga Head of the Department of Genetics UMCG and Professor Dirkje Postma from the Department of Pneumology UMCG for help in all steps of the scientific work, the nurses from the Department of Allergology for help in collecting the blood samples, Pieter van der Vlies and all colleagues from the Department of Genetics UMCG for help in laboratory work, Agata Somla from the Medical University of Gdansk for help in financial logistics.

The research was supported by the Foundation for Polish Science, and grant of the Polish Ministry of Science and Higher Education, no. N402085934 and N40201031.

References

- Valent P, Sperr W, Schwartz L, Horny H. Diagnosis and classification of mast cell proliferative disorders: delineation from immunologic diseases and non-mast cell hematopoietic neoplasms. *J Allergy Clin Immunol* 2004;**114**:3–11.
- Garcia-Montero AC, Jara-Acevedo M, Teodosio C, Sanchez ML, Nunez R, Prados A et al. KIT mutation in mast cells and other bone marrow hematopoietic cell lineages in systemic mast cell disorders: a prospective study of the Spanish Network on Mastocytosis (REMA) in a series of 113 patients. *Blood* 2006;**108**:2366–2372.
- Niedoszytko B, Niedoszytko M, Lange M, van Doormaal J, Gleis J, Zabotna M et al. Interleukin-13 promoter gene polymorphism -1112C/T is associated with the systemic form of mastocytosis. *Allergy* 2009;**64**:287–294.
- Daley T, Metcalfe D, Akin C. Association of the Q576R polymorphism in the interleukin-4 receptor alpha chain with indolent mastocytosis limited to the skin. *Blood* 2001;**98**:880–882.
- Nagata H, Worobec AS, Oh CK, Chowdhury BA, Tannenbaum S, Suzuki Y et al. Identification of a point mutation in the catalytic domain of the protooncogene c-kit in peripheral blood mononuclear cells of patients who have mastocytosis with an associated hematologic disorder. *Proc Natl Acad Sci USA* 1995;**92**:10560–10564.
- Lawley W, Hird H, Mallinder P, McKenna S, Hargadon B, Murray A et al. Detection of an activating c-kit mutation by real-time PCR in patients with anaphylaxis. *Mutat Res* 2005;**572**:1–13.
- Kazmierska J, Malicki J. Application of the Naïve Bayesian Classifier to optimize treatment decisions. *Radiation Oncol* 2008;**3**:211–216.
- Nogales-Cadenas R, Carmona-Saez P, Vazquez M, Vicente C, Yang X, Tirado F et al. GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res* 2009;**37**:514–521.
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2008;**36**:480–484.
- Fehrmann RS, de Jonge HJ, Ter Elst A, de Vries A, Crijns AG, Weidenaar AC et al. A new perspective on transcriptional system regulation (TSR): towards TSR profiling. *PLoS ONE* 2008;**3**:e1656.
- Liu SM, Xavier R, Good KL, Chtanova T, Newton R, Sisavanh M et al. Immune cell transcriptome datasets reveal novel leukocyte subset-specific genes and genes associated with allergic processes. *J Allergy Clin Immunol* 2006;**118**:496–503.
- D'Ambrosio C, Akin C, Wu Y, Magnusson MK, Metcalfe DD. Gene expression analysis in mastocytosis reveals a highly consistent profile with candidate molecular markers. *J Allergy Clin Immunol* 2003;**112**:1162–1170.
- Varmeh-Ziaie S, Ichimura K, Yang F, Rabbits P, Collins VP. Cloning and chromosomal localization of human WIG-1/PAG608 and demonstration of amplification with increased expression in primary squamous cell carcinoma of the lung. *Cancer Lett* 2001;**174**:179–187.
- Semba S, Iwaya K, Matsubayashi J, Serizawa H, Kataba H, Hirano T et al. Coexpression of actin-related protein 2 and Wiskott-Aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung. *Clin Cancer Res* 2006;**12**:2449–2454.
- Wang JC, Cruchaga C, Saccone NL, Bertelsen S, Liu P, Budde JP et al. Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. *Hum Mol Genet* 2009;**18**:3125–3135.
- Johnstone CN, Castellví-Bel S, Chang LM, Sung RK, Bowser MJ, Piqué JM et al. PRR5 encodes a conserved proline-rich protein predominant in kidney: analysis of genomic organization, expression, and mutation status in breast and colorectal carcinomas. *Genomics* 2005;**85**:338–351.
- Puiffe ML, Le Page C, Filali-Mouhim A, Zietarska M, Ouellet V, Tonin PN et al. Characterization of ovarian cancer ascites on cell invasion, proliferation, spheroid formation, and gene expression in an in vitro model of epithelial ovarian cancer. *Neoplasia* 2007;**9**:820–829.
- Pillozzi S, Brizzi MF, Bernabei PA, Bartolozzi B, Caporale R, Basile V et al. VEGFR-1 (FLT-1), beta1 integrin, and hERG K⁺ channel for a macromolecular signaling complex in acute myeloid leukemia: role in cell migration and clinical outcome. *Blood* 2007;**110**:1238–1250.
- Grosjean-Raillard J, Tailler M, Adès L, Perfettini JL, Fabre C, Braun T et al. ATM mediates constitutive NF-kappaB activation in high-risk myelodysplastic syndrome and acute myeloid leukemia. *Oncogene* 2009;**28**:1099–1109.
- Matsuo Y, Drexler HG, Harashima A, Okochi A, Shimizu N, Orita K. Transcription factor expression in cell lines derived from natural killer-cell and natural killer-like T-cell leukemia-lymphoma. *Hum Cell* 2004;**17**:85–92.
- Bursen A, Moritz S, Gaussmann A, Moritz S, Dingermann T, Marschalek R. Interaction of AF4 wild-type and AF4.MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? *Oncogene* 2004;**23**:6237–6249.
- Pinyol M, Bea S, Plà L, Ribrag V, Bosq J, Rosenwald A et al. Inactivation of RB1 in mantle-cell lymphoma detected by nonsense-mediated mRNA decay pathway inhibition and microarray analysis. *Blood* 2007;**109**:5422–5429.
- Soler G, Nusbaum S, Varet B, Macintyre EA, Vekemans M, Romana SP et al. LRRFIP1, a new FGFR1 partner gene associated with 8p11 myeloproliferative syndrome. *Leukemia* 2009;**23**:1359–1361.
- Li Y, Li W, Yang Y, Lu Y, He C, Hu G et al. MicroRNA-21 targets LRRFIP1 and contributes to VM-26 resistance in glioblastoma multiforme. *Brain Res* 2009;**1286**:13–18.
- Cheong JK, Gunaratnam L, Zang ZJ, Yang CM, Sun X, Nasr SL et al. TRIP-Br2 promotes oncogenesis in nude mice and is frequently overexpressed in multiple human tumors. *J Transl Med* 2009;**7**:8.
- Goishi K, Mizuno K, Nakanishi H, Sasaki T. Involvement of Rab27 in antigen-induced histamine release from rat basophilic leukemia 2H3 cells. *Biochem Biophys Res Commun* 2004;**324**:294–301.
- Kuroda TS, Fukuda M. Rab27A-binding protein Slp2-a is required for peripheral melanosome distribution and elongated cell shape in melanocytes. *Nat Cell Biol* 2004;**6**:1195–1203.
- Ruault M, Brun ME, Ventura M, Roizès G, De Sario A. MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a chromosome region frequently deleted in myeloid leukaemia. *Gene* 2002;**284**:73–81.
- Lee JS, Yang EJ, Kim IS. The roles of MCP-1 and protein kinase Cdelta activation in human eosinophilic leukemia EoL-1 cells. *Cytokine* 2009;**48**:186–195.
- Lu JG, Sun YN, Wang C, Jin de J, Liu M. Role of the alpha v-integrin subunit in cell proliferation, apoptosis and tumor metastasis of laryngeal and hypopharyngeal squamous cell carcinomas: a clinical and in vitro investigation. *Eur Arch Otorhinolaryngol* 2009;**266**:89–96.
- Zwergler M, Herrmann H, Gaines P, Olins AL, Olins DE. Granulocytic nuclear differentiation of lamin B receptor-deficient mouse EPRO cells. *Exp Hematol* 2008;**36**:977–987.
- Johnson KR, Johnson KY, Becker KP, Bielawski J, Mao C, Obeid LM. Role of human sphingosine-1-phosphate phosphatase 1 in the regulation of intra- and extracellular

- sphingosine-1-phosphate levels and cell viability. *J Biol Chem* 2003;**278**:34541–34547.
33. Valencia CA, Ju W, Liu R. Matrin 3 is a Ca²⁺/calmodulin-binding protein cleaved by caspases. *Biochem Biophys Res Commun* 2007;**361**:281–286.
34. Crijns AP, Fehrmann RS, de Jong S, Gerbens F, Meersma GJ, Klip HG et al. Survival-related profile, pathways, and transcription factors in ovarian cancer. *PLoS Med* 2009;**6**:e24.
35. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;**415**:530–536.