bjh research paper

Gene expression profiles of erythroid precursors characterise several mechanisms of the action of hydroxycarbamide in sickle cell anaemia

Flávia Chagas Costa, Anderson Ferreira da Cunha, André Fattori, Tarcísio de Sousa Peres, Gustavo Gilson Lacerda Costa, Tiago Ferraz Machado, Dulcinéia Martins de Albuquerque, Sheley Gambero, Carolina Lanaro, Sara T. Olala Saad and Fernando Ferreira Costa

The Haematology and Haemotherapy Centre, Faculty of Medical Sciences, State University of Campinas, Campinas, SP, Brazil

Received 3 August 2006; accepted for publication 11 October 2006; first published online 8 December 2006

Correspondence: Fernando F. Costa, Rua. Carlos Chagas, 480 – Cidade Universitária, Barão Geraldo, Campinas, SP 13083-970, Brazil. E-mail: ferreira@unicamp.br

Summary

Hydroxycarbamide (HC) (or hydroxyurea) has been reported to increase fetal haemoglobin levels and improve clinical symptoms in sickle cell anaemia (SCA) patients. However, the complete pathway by which HC acts remains unclear. To study the mechanisms involved in the action of HC, global gene expression profiles were obtained from the bone marrow cells of a SCA patient before and after HC treatment using serial analysis of gene expression. In the comparison of both profiles, 147 differentially expressed transcripts were identified. The functional classification of these transcripts revealed a group of gene categories associated with transcriptional and translational regulation, e.g. *EGR-1*, *CENTB1*, *ARHGAP4* and *RIN3*, suggesting a possible role for these pathways in the improvement of clinical symptoms of SCA patients. The genes involved in these mechanisms may represent potential tools for the identification of new targets for SCA therapy.

Keywords: sickle cell anaemia, fetal haemoglobin, hydroxycarbamide, gene expression, serial analysis of gene expression.

High levels of fetal haemoglobin (HbF) have long been recognised to ameliorate the clinical symptoms of sickle cell anaemia (SCA), decreasing painful events, transfusions and hospitalisations due to its ability to inhibit the polymerisation of HbS (Bunn, 1997). Consequently, the decades-long search for therapeutic modulators of HbF has continued to motivate basic and clinical investigators alike. As a result, several pharmacological agents have been shown to increase HbF production (Halsey & Roberts, 2003).

To date, the prototypic agent for increasing HbF expression in SCA patients is hydroxycarbamide (HC) (or Hydroxyurea), a ribonucleotide reductase inhibitor that exhibits an anti-tumour effect that arrests DNA synthesis (Halsey & Roberts, 2003). Recent multicentre studies with HC have shown that administration of this drug to SCA patients significantly increases HbF production and improves clinical symptoms by reducing the frequency of pain and vaso-occlusive crisis, acute chest syndrome, transfusion requirement and hospitalisations and also appears to reduce mortality. Moreover, some patients appear to benefit from HC even when HbF levels do not increase, suggesting that HC may act by additional mechanisms (Charache, 1997; Steinberg *et al*, 2003). The complete pathway by which HC acts remains unclear. HC has been shown to induce the expression of globin gene, possibly by induction of guanylate cyclase protein kinase G pathways (Ikuta *et al*, 2001; Cokic *et al*, 2003), or by enhancement of the expression of transcription factors, such as *early growth response 1* (*EGR1*), *GATA-1* and other genes with important roles in erythropoiesis (Wang *et al*, 2002).

The aim of this study was to evaluate the global gene expression pattern of human bone marrow cells from a SCA patient before and after the administration of HC, using serial analysis of gene expression (SAGE) (Velculescu *et al*, 1995; Hashimoto *et al*, 2003). The knowledge of which genes are activated or repressed by HC could help to clarify the mechanism by which HC acts and could distinguish potential gene targets for the development of new therapies for SCA and other haemoglobinopathies.

Materials and methods

Patients

A total bone marrow cell (HBMC) sample was obtained by puncture and aspiration of bone marrow from a 23-year-old female who presented a homozygous HbS genotype, and was

  First published online 8 December 2006
doi:10.1111/j.1365-2141.2006.06424.x

followed at the Haematology and Haemotherapy Centre (UNICAMP, Campinas, SP, Brazil). The cells were collected from the same sickle cell patient before and after 90 d of regular treatment with HC (25 mg/kg/d). Table I presents relevant haematological data.

Additionally, peripheral blood samples from eight patients with SCA were collected (four of these patients were receiving HC therapy). All patients enroled authorised the procedures and a consent form was signed only after local ethics committee approval.

K562 Cell culture

Human K562 erythroleukaemia cells were obtained from American Type Culture Collection (ATCC; catalogue number CCL-243) and maintained in RPMI1640 medium (Gibco BRL Life Technologies, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (Gibco BRL), 100 units/ml penicillin (Gibco BRL), and 100 mg/ml streptomycin (Gibco BRL) in a humidified 5% CO_2 atmosphere at 37°C. Cells were seeded at 5×10^5 cells per 75 cm^2 flask (Nunc, Naperville, IL, USA) prior to the addition of 80 BM HC (Sigma, St Louis, MO, USA). The medium was exchanged every third day.

Reticulocytes and leucocytes isolation

Reticulocytes were isolated as previously reported (Goossens & Kan, 1981). Peripheral blood cells were centrifuged and the plasma removed. Erythrocytes were then lysed with red blood cell lysis solution (0·144 mol/l NH_4Cl, 0·01 mol/l NH_4HCO_3) and the samples centrifuged. The remaining pellet containing leucocytes was resuspended in 1 ml of Trizol® (Invitrogen, Rockville, MS, USA) and the supernatant homogenised with a 1/10 dilution of a sucrose/KCl solution (1·5 mol/l $C_{12}H_{22}O_{11}$, 0·15 mol/l KCl). After centrifugation, supernatant containing reticulocytes was treated with 800 µl of 10% acetic acid and centrifuged again and the pellet was then resuspended in 1 ml of Trizol® (Invitrogen).

Table I. Haematological data of sickle cell anaemia patient participating in this study.

Haematological data	Before HC	After HC
Red blood cell count ($\times 10^{12}$/l)	2·41	2·33
Erythroblast in bone marrow (%)	55	65
Haematocrit (%)	23·9	24·3
Haemoglobin (g/l)	76	80
Mean corpuscular volume (fl)	99·2	104·3
Mean corpuscular haemoglobin (pg)	31·5	34·3
White blood cell count ($\times 10^9$/l)	10·85	5·16
Neutrophils ($\times 10^9$/l)	6·03	2·08
HbF (%)	3·40	8·80

RNA extraction and reverse transcription

Total RNA was extracted using TRIzol (Invitrogen), according to the manufacturer's protocol. RNA integrity was confirmed by electrophoresis in 1·2% denaturing agarose gel and the RNA concentration was quantified using a GeneQuant UV spectrophotometer (Amersham Pharmacia, Bucks, UK).

Five micrograms of each RNA sample was incubated with 1 U DNase I (Invitrogen) for 15 min at room temperature, and 2 mmol/l of EDTA was added to stop the reaction. The DNase I enzyme was subsequently inactivated for 5 min at 65°C. DNase I-treated RNA samples were then reverse transcribed with 30 U AMV Reverse Transcriptase (Promega, Madison, WI, USA) for 60 min at 42°C and subsequently for 15 min at 70°C. The cDNA samples were quantified using a GeneQuant UV spectrophotometer (Amersham Pharmacia).

Generation of SAGE profiles

Ten micrograms of total RNA was used to obtain the SAGE libraries using the I-SAGE kit (Invitrogen) as described by the manufacturer. Briefly, mRNA was converted into double-stranded cDNA using a biotinylated oligo (dT) primer. The cDNA was digested with NlaIII and 3′-cDNAs were isolated using streptavidin paramagnetic beads. 3′-cDNAs were split into two pools and SAGE linkers, A and B, were ligated to pools 1 and 2 respectively. SAGE tags were released with BsmfI and blunt ended with T4 polymerase. The tags from pools 1 and 2 were then ligated to each other. The ligation product was amplified with 27 cycles of polymerase chain reaction (PCR) and digested with NlaIII. Ditags were isolated using a 12% polyacrylamide gel, concatemerised and cloned into pZero digested with SphI. (Velculescu *et al*, 1995; Pollock, 2002).

Sequencing reactions were carried out using the Dynamic ET Terminator cycle sequencing (GE-Healthcare, San Francisco, CA, USA) and were loaded and run on a Mega-BACE automated DNA sequencer (Amersham Pharmacia). The sequences were analysed and the vector was trimmed using the PHRED/PHRAP software (Ewing and Green, 1998).

SAGE library data and statistical analysis

The SAGE data analysis involved the determination of tag abundance, identification of each tag, comparison of both profiles to obtain the statistically significant differences and functional classification.

To determine tag abundance, tag sequences and their counts in the libraries were organised and tallied via ESAGE software v1.2 (Margulies & Innis, 2000). For tag identification, a database was generated by comparing all potential tags of each library with tag sequences from the CGAP data bank, downloaded from the NCBI SAGE website (Lal *et al*, 1999).

The comparison of relative abundance of each tag between the HbSHC library (after HC administration) and control library (HbS) was perfomed using ESAGE software, producing a *P*-value

for the difference. Only tags that presented $P < 0{\cdot}01$ and that were more than fivefold higher were considered statistically significant (Audic & Claverie, 1997; Margulies & Innis, 2000; Chabardes-Garonne *et al*, 2003; de Chaldee *et al*, 2003).

The functional classification was performed according to the Gene Ontology Consortium (http://www.geneontology.org). The gene ontology categories were further grouped according to molecular functions using the classifications described by the Gene Ontology Consortium. Some transcripts bore several functions and were therefore counted in each category to which they belonged. The numbers of tags corresponding to a transcript found in each category were cumulated to represent the relative importance of each gene ontology class.

Validation of gene expression measurements using the semi-quantitative polymerase chain reaction (PCR)

Gene expression analysis of mRNA was performed using semi-quantitative reverse transcription (RT) PCR to validate *SAGE* data (Hashimoto *et al*, 2003; Schwering *et al*, 2003).

The validation was first performed using cDNA samples of the two bone marrow cell (HBMC) samples from the SCA patient. Additionally, cultured K562 cell lineages, which were treated or not with HC, were also used to confirm the data obtained in the BMC. The cDNA samples were quantified and the same amount (1 μg) were serially diluted to 1:2, 1:4, 1:8, 1:16, 1:32, 1:64, 1:128 and then used in the reaction. The RT-PCR reaction was performed in 25 μl of 1·25 mmol/l of each dNTP, 1x PCR Buffer (Minus $MgCl_2$), 2 mmol/l $MgCl_2$,1 U taq DNA polymerase (Invitrogen) and 15 pmol of each primer. The amplification program consisted of: 3 min at 94°C, followed by 40 cycles of 30 s at 94°C, 35 s at annealing temperature (changing according to the primer sequence) and finally 45 s at 72°C. The primers used in this study are listed in Table II.

Analysis of gene expression by real-time PCR

Quantitative real-time PCR was used to evaluate the gene expression pattern in the leucocytes and reticulocytes of a group of HC-untreated and -treated patients. Synthetic oligonucleotide primers for all selected genes were designed and synthesised by Invitrogen (see Table III for sequences). Quantification of *GAPDH* expression was used as an internal

Table II. Primer sequences for the semi-quantitative RT-PCR amplification of selected differentially expressed genes for corroboration of results obtained by SAGE.

Genes	Sequences
Beta-actin	F: 5′-AAGAGATGGCCACGGCTGCT-3′
	R: 5′-TCGCTCCAACCGACTGCTGT-3′
γ-globin *(HBG)*	F: 5′-CCATAAAGCACCTGGATGATC-3′
	R: 5′-ATCTGGAGGACAGGGCACTG-3′
Endothelial differentiation factor 1 *(EDF1)*	F: 5′-CCAAGAAATGGGCTGCTGG-3′
	R: 5′-GGGCTTTCCAATGTCCTTTCC-3′
Enhancer rudimentary homologue Drosophila *(ERH)*	F: 5′-GATGATCTGGCAGACCTCAGC-3′
	R: 5′-TCCGATTGAACAAGATCCTCAC-3′
Early growth response 1 *(EGR1)*	F: 5′-CCAGTCCCAGCTCATCAAAC-3′
	R: 5′-ATCTTGGTATGCCTCTTGCG-3′
Heat shock protein 1 *(HSPE1)*	F: 5′-CGAGTATTGGTTGAAAGGAGTG-3′
	R: 5′-TTTCAGAACTTCAGTGGAATGG-3′
Signal transducer and activator of transcription 5A *(STAT5A)*	F: 5′-GCTGTGGGACAATGCCTTTG-3′
	R: 5′-TTCTTCAACACCTCCATCACCC-3′
Peroxiredoxin 6 *(PRDX6)*	F: 5′-ACTGTCCTATCACGTCCTCTCC-3′
	R: 5′-ATTGCTCACACCTGCTACACC-3′

Table III. Primer sequences used for the quantitative real-time PCR amplification of selected differentially expressed genes.

Genes	Sequences
Early growth response 1 *(EGR1)*	F: 5′-GAGCATGTGTCAGAGTGTTGTT-3′
	R: 5′-CATGTGAGAGTACGGTCAAGC-3′
γ-globin *(HBG)*	F: 5′-CATGGCAAGAAGGTGCTGACT-3′
	R: 5′-GCAAAGGTGCCCTTGAGATC-3′
Glyceraldehyde-3-phosphate dehydrogenase *(GAPDH)*	F: 5′-GCACCGTCAAGGCTGAGAAC-3′
	R: 5′-CCACTTGATTTTGGAGGGATCT-3′
Peroxiredoxin 6 *(PRDX6)*	F: 5′-ATTAGTGCCATGTGCCTTTC-3′
	R: 5′-TCTCTTCTCTTCACATGGTGTAC-3′
Heat shock protein 1 *(HSPE1)*	F: 5′-CTGAAGTTCTGAAATCTTTCGTC-3′
	R: 5′- TTTGGAGACACTGGATGTCATT-3′

control to determine the amount and quality of cDNA and untreated samples were used as references. All samples were assayed in a 25 μl volume containing 100 ng cDNA, 12·5 μl SYBR Green Master Mix PCR (Applied Biosystems) and the concentration of each primer was 600 nmol/l for *EGR1*, *HSPE1* and *PRDX6;* 300 nmol/l for *GAPDH* and 150 nmol/l for *HBG* primers. Assays were performed in a MicroAmp Optical 96-well reaction plate (Applied Biosystems) using the 5700 Sequence Detection System (Applied Biosystems). To confirm the accuracy and reproducibility of real-time PCR, the intra-assay precision was calculated according to the equation $E^{(-1/slope)}$, to calculate the amplification efficiency (Pfaffl, 2001). The assayed transcripts demonstrated high real-time PCR amplification efficiency rates of 99% for *EGR1*, *HSPE1*, *PRDX6* and *TM7* and 100% for *GAPDH* and *HBG* (Pearson correlation coefficient $r > 0.95$).

To quantify gene expression, a mathematical model, using calibration data [$2^{(-\Delta\Delta CT)}$], was used. For this, the determination of the crossing threshold was necessary for each transcript. The crossing threshold was defined as the point at which the fluorescence rises appreciably above the background fluorescence. The dissociation protocol was performed at the end of each run to check for non-specific amplification. Two replicas were run on the plate for each sample and each sample was run twice, independently. Results are expressed as the fold change in gene expression when compared with the reference non-treated samples.

Results

Generation of SAGE data

A total of 45 014 tags for the HbS profile and 49 814 tags for the HbSHC profile were analysed, representing 15 358 and 16 507 unique tags respectively. Of these, approximately 76% of the unique tags from both profiles were matched to known genes or open reading frames (ORFs) in the human genome, and 24% corresponded to tags that did not match any of the published reference for the human genome. These sequences may represent novel transcripts (see Table IV for details).

Analysis of the relative distribution of gene expression levels showed that only 0·3% of the messengers were expressed at 100 copies/cell or more, 3% at more than 10 copies/cell, whereas most transcripts (97%) were present at 10 copies/cell or less (Fig 1). Linker contamination was approximately 0·7%, a similar level to those observed by other authors using the same procedure (Malago *et al*, 2005; Varela *et al*, 2005).

For subsequent analysis, only tags present at least four times in one of the libraries were considered (Chabardes-Garonne *et al*, 2003; de Chaldee *et al*, 2003; Hashimoto *et al*, 2003). Using these data, the expression profile of HbS and HbSHC libraries were compared using the Gene Ontology Consortium database. Some transcripts bore several functions and were therefore counted in each category to which they belonged, i.e. the weight of each tag expression was considered in the classification (Fig 2). The most abundant genes expressed in both libraries were the α- and β-globin genes. Excluding these genes, the 40 transcripts with the highest copy numbers in the HbS and HbSHC profiles are shown in Table V. A complete list of HbS and HbSHC analysis is available for download at http://www.hemocentro.unicamp.br/submission.

Table IV. Summary of the tag numbers in HbS and HbSHC profiles.

Classification	HbS	HbSHC
Total tags	45 014	49 814
Unique tags	15 358	16 507
Identified genes	8·143 (53%)	9·749 (59%)
ORFs	3·422 (22·3%)	2·685 (16·26%)
Non-identified tags	4·053 (24·7%)	4·074 (24·6%)

ORFs, open reading frames.

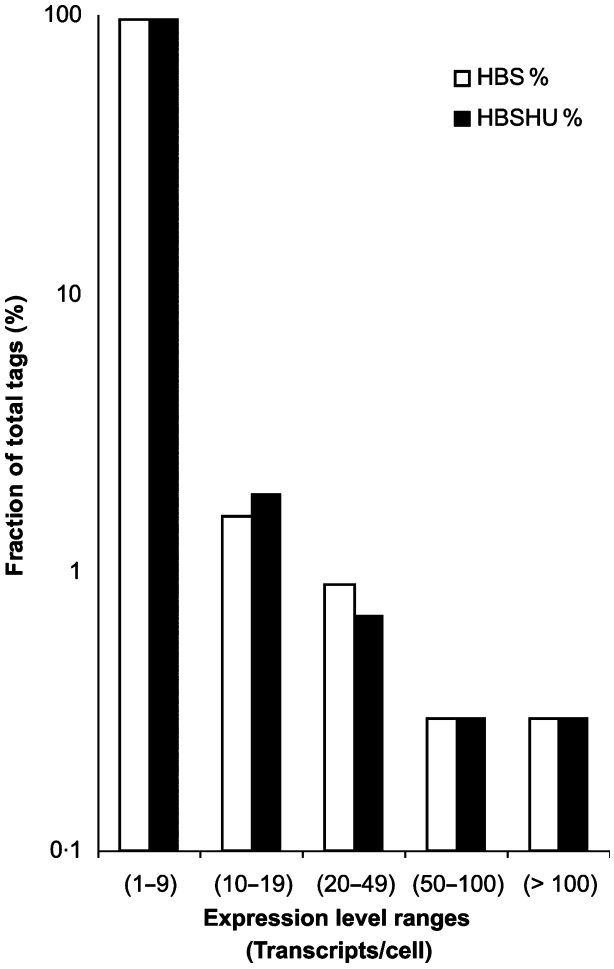

Fig 1. Relative distribution of gene expression pattern of the HbS and HbSHC profiles. Total tags obtained from the HbS profile was 45·014 tags and 49·814 tags for the HbSHC profile.

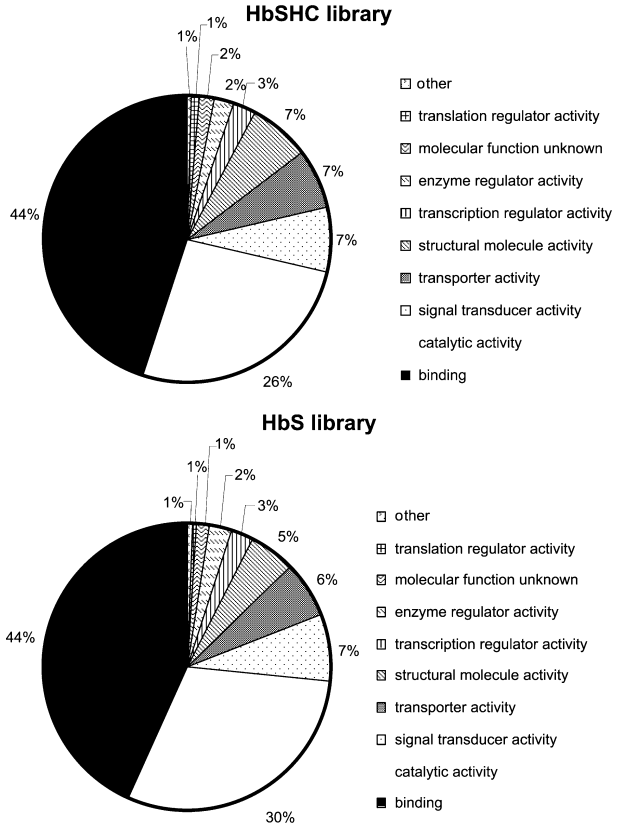

Fig 2. Functional classification of the HbSHC (upper panel) and HbS (lower panel) libraries. An increase in transporter activity, structural molecule activity and catalytic activity categories was observed after hydroxycarbamide (HC) treatment.

Gene expression profiles in HbS and HbSHC

Comparison of the 15·358 and 16·507 distinct tags from the HbS and HbSHC profiles showed that 147 tags presented a differential expression with a statistically significant level ($P < 0{\cdot}01$ and more than fivefold). These sequences corresponded to known genes, annotated sequences [like expressed sequence tags (ESTs), ORFs, cDNAs] and no matches. The comparison of the differentially expressed tags demonstrated that 82 genes were repressed and 65 are induced by HC treatment. Table VI details some genes found to be significantly differentially expressed and their classification, according to the Gene Ontology Consortium.

Results showed that the dihydropyrimidine dehydrogenase and Microtubule associated monoxygenase genes, related to metal ion binding, and that the colony-stimulating factor 3 receptor (granulocyte), Interleukin 6 (IL-6) signal transducer (gp130 oncostatin M receptor), plexin B2 and Progesterone receptor membrane component 2, related to receptor activity categories, were downregulated after HC treatment. In contrast, some genes classified as structural constituents of ribosomes, nucleotide binding, transcription factor activity, enzyme activator activity and channel or pore class transporter activity categories were upregulated after treatment with this drug. The RNA binding category had some genes that were up- and downregulated (For details see Table VI).

Expression of the globin genes

According to the comparison with both profiles, our data showed that HC treatment induced the expression of γ-globin gene by 2.7-fold, as expected, while the gene expression patterns of the other globins were slightly decreased in abundance after HC administration. The expressions of all related globin tags are summarised in Table VII.

Validation of sage results using semi quantitative RT-PCR

To investigate the reliability of SAGE results, we arbitrarily selected a total of seven differentially expressed genes and analysed their expression levels by semi quantitative RT-PCR.

Initially, reactions for validation were performed using the remaining available cDNA obtained from the patient HBMC sample, before and after HC administration. The genes studied were *PRDX6*, γ-globin, *EGR1*, *HSPE1* and *EDF1*. Results are shown in Fig 3A.

Additionally, K562 cells that were treated or not with 80 μmol/l of HC were used as a template to corroborate the gene expression pattern of the previously analysed five genes (from the HBMC sample) and a further two genes (*ERH* and *STAT5A*). The expression levels of these genes, amplified in the K562 cell lines, were in agreement with the data achieved by SAGE (Fig 3B).

The γ-globin gene was slightly less expressed in un-treated than in treated HBMC, according to the SAGE data. This level of differential expression was difficult to detect by RT-PCR, however, at 1:16 dilution, a clear stronger amplification of γ-globin gene was observed in the treated cells compared with un-treated cells. The γ-globin gene was abundantly expressed in un-treated and treated K562 cells, but a slight enhancement in HC-treated samples was observed compared with untreated cells. *EDF1* was detected when in the HC-treated, but not the un-treated, HBMC at a 1:4 dilution. K562 cells treated with HC, slightly overexpressed *EDF1* even at the 1:8 dilution, compared with the expression at the same dilution in the untreated K562 cells. *EGR-1* was weakly expressed at a 1:16 dilution in treated HBMC and was not detected in the un-treated sample. This gene was weakly expressed at the 1:16 dilution and strongly expressed at the same dilution in treated K562 cells. RT-PCR for *HSPE1* was positive up to a 1:2 dilution of HC-treated and un-treated HBMC although *HSPE1* expression was weaker in the un-treated HBMC. In treated K562 cells, *HSPE1* was detectable up to a 1:8 dilution and showed weaker expression at the same dilution for untreated K562 cells. *PRDX6*, was downregulated in SAGE, but RT PCR showed slightly higher *PRDX6* expression in treated HBMC than in untreated HBMC (1:4), although the results were not clearly conclusive.

Additionally, two other genes were analysed only in HC-treated and untreated K562 cells and both results were in

Table V. The most abundant identified transcripts in the HbS and HbSHC SAGE profiles with their respective numbers of tags.

Tag	HbS	HbSHC	*P*-value	Fold	Unigene	Symbol	Description
TACCTGCAGA	1131	743	6·7763E-21	1·65	Hs.416073	S100A8	S100 calcium binding protein A8 (calgranulin A)
GTGGCCACGG	418	743	0·27332849	1·8	Hs.112405	S100A9	S100 calcium binding protein A9 (calgranulin B)
GTTGTGGTTA	361	428	0·2053277	−1·09	Hs.534255	B2M	Beta-2-microglobulin
TCCCCGTACA	326	29	6·7763E-21	12·2	Hs.530970	BCL7A	B-cell CLL/lymphoma 7A
TCCCTATTAA	304	31	6·7763E-21	10·6	Hs.518618	CORIN	Corin, serine protease
GCCTGCTATT	304	236	0·00011687	1·39	Hs.380781	DEFA1	Defensin, alpha 1, myeloid-related sequence
TAGGTTGTCT	267	327	0·13594007	−1·13	Hs.374596	TPT1	Tumour protein, translationally controlled 1
GACTGCTATT	235	196	0·00683077	1·3	Hs.294176	DEFA3	Defensin, alpha 3, neutrophil-specific
ATAATTCTTT	145	206	0·01195174	−1·31	Hs.156367	RPS29	Ribosomal protein S29
TTGGTGAAGG	143	187	0·08952269	−1·21	Hs.522584	TMSB4X	Thymosin, beta 4, X-linked
ATGTAAAAAA	133	177	0·07228109	−1·23	Hs.524579	LYZ	Lysozyme (renal amyloidosis)
TGTGTTGAGA	124	116	0·25682612	1·16	Hs.549262	EEF1A1	Eukaryotic translation elongation factor 1 alpha 1
TTGGTCCTCT	123	164	0·08106035	−1·23	Hs.242947	DGKI	Diacylglycerol kinase, iota
CACCTAATTG	122	284	1·9913E-13	−2,15	Hs.461412		CDNA clone IMAGE:5268125, partial cds
CCTGTAATCC	122	99	0·03251547	1·33	Hs.158209	NT5C2	5′-nucleotidase, cytosolic II
CTGACCTGTG	119	157	0·10432859	−1·22	Hs.181244		M4–50 mRNA for HLA class I antigen
GAAATAAAGC	118	163	0·04366961	−1·28	Hs.525648	IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)
CCACTGCACT	116	86	0·00745563	1·46	Hs.107003	CCNB1IP1	Cyclin B1 interacting protein 1
CCCATCGTCC	115	248	2·9655E-10	−1·99	Hs.421608	EEF1B2	Eukaryotic translation elongation factor 1 beta 2
GTGAAACCCC	114	131	0·64606032	−1·06	Hs.477083	PAFAH2	Platelet-activating factor acetylhydrolase 2, 40kDa
GCATAATAGG	108	234	7·5669E-10	−2	Hs.381123	RPL21	Ribosomal protein L21
CACAAACGGT	108	204	1·84E-06	−1·74	Hs.129826	NET-5	Tetraspanin 9
TCAGATCTTT	106	134	0·2351239	−1·17	Hs.446628	RPS4X	Ribosomal protein S4, X-linked
TTCATACACC	105	189	2·1931E-05	−1·66	Hs.527213	DCOHM	6-pyruvoyl-tetrahydropterin synthase (TCF1)
CTGGGTTAAT	105	150	0·02901554	−1·32	Hs.438429	RPS19	Ribosomal protein S19
TTCAATAAAA	104	109	1	−0·97	Hs.356502	RPLP1	Ribosomal protein, large, P1
CTAAGACTTC	102	160	0·00321007	−1·45	Hs.533150		Moderately similar to NP_536843·1 (Homo sapiens)
GTTCACATTA	101	219	2·5578E-09	−2	Hs.436568	CD74	CD74 antigen
GTAAAAAAAA	100	25	2·1551E-13	4·33	Hs.356190	UBB	Ubiquitin B
TTGGGGTTTC	96	96	1	1·08	Hs.448738	FTH1	Ferritin, heavy polypeptide 1
TGCACGTTTT	96	115	0·46816946	−1·11	Hs.265174	RPL32	Ribosomal protein L32
CAATAAATGT	95	110	0·63593158	−1·07	Hs.80545	RPL37	Ribosomal protein L37
GAGGGAGTTT	88	229	2·023E-13	−2·4	Hs.523463	RPL27A	Hypothetical protein MGC10850
GGATTTGGCC	87	309	6·7763E-21	−3·28	Hs.437594	RPLP2	Ribosomal protein, large P2
GGGCTGGGGT	86	164	1·4433E-05	−1·76	Hs.425125	RPL29	Ribosomal protein L29
CAAGCATCCC	86	65	0·02807653	1·43	Hs.470205		Moderately similar to XP_531272·1 (Pan troglodytes)
ACACAGCAAG	85	82	0·45435255	1·12	Hs.476965		Full-length cDNA clone CS0DI084YH04
TGTACCTGTA	83	54	0·00324525	1·66	Hs.524390	K-ALPHA-1	Tubulin, alpha, ubiquitous
TACCATCAAT	82	94	0·71005987	−1·06	Hs.544577	GAPD	Glyceraldehyde-3-phosphate dehydrogenase
GCTTTCTCAC	56	310	6·7763E-21	−5·11	Hs.446484	CSNK2A1	Casein kinase 2, alpha 1 polypeptide

agreement with SAGE data. The *STAT5A* gene presented a higher expression when treated K562 cells were diluted to 1:2 and had almost no detectable expression in the un-treated sample at the same dilution. Treated, but not un-treated, K562 cells expressed *ERH* at a 1:4 dilution.

Real-time PCR

In order to study some of these genes *(EGR-1, HBG, PRDX-6, TM7* and *HSPE-1*) in a number of patients with sickle cell anaemia, we analysed reticulocyte and leucocyte mRNA from the peripheral blood of four SCA patients not treated with HC and four SCA patients on HC therapy. Real-time PCR showed that, in the treated patients, the production of the γ-globin protein in reticulocytes was twice that observed in untreated patients, as was already observed in the SAGE data.

For *EGR-1*, the results showed an increased expression in both cellular types after treatment with HC and the same result was observed for the *HSPE-1* and *PRDX-6* (Fig 4). For *PRDX-6*, in contrast to *EGR-1, HSPE-1* and γ-globin gene, the PCR

Table VI. Some of the genes found to be significantly induced (I) or repressed (R) and their classification according to the functional analysis of both profiles at level three of molecular function using the Gene Ontology Consortium.

Functional category	Gene list	HbS	HbSHC
Metal ion binding	Dihydropyrimidine dehydrogenase (R)	8	0
	Microtubule associated monoxygenase (R)	6	0
Nucleotide binding	Casein Kinase2, alpha 1 polypeptide (I)	56	310
	Eukaryotic translation initiation factor 3, subunit 4 delta (I)	0	9
	Eukaryotic translation initiation factor 3, subunit 6 48 kDa (I)	1	13
	Eukaryotic translation initiation factor 5A (I)	3	20
	Eukaryotic translation initiation factor 5A (I)	1	12
	Eukaryotic translation initiation factor 3, subunit 3 gamma, 40 kDa (I)	1	11
Receptor activity	Colony stimulating factor 3 receptor (granulocyte) (R)	25	7
	Interleukin 6 signal transducer (gp130 oncostatin M receptor) (R)	20	0
	Plexin B2 (R)	9	0
	Progesterone receptor membrane component 2 (R)	8	0
RNA binding	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3 X-linked (R)	27	11
	Fusion [involved in t (12–16) in malignant liposarcoma] (I)	2	12
	HIV-1 Rev-binding protein (R)	8	1
	PAI-1 mRNA-binding protein (R)	21	2
	RNA-binding protein (autoantigenic hnRNP-associated) (I)	1	12
	Small nuclear ribonucleoprotein 70 KDa polypeptide (RNP antigen) (I)	0	8
	Small nuclear ribonucleoprotein polypeptides B and B1 (I)	2	12
Structural constituent of ribosome	Ribosomal protein L10a (I)	0	7
	Ribosomal protein, large, P0 (I)	2	13
	Ribosomal protein large P2 (I)	87	309
Transcription factor actitvity	Early growth response 1 (I)	1	25
	Kruppel-like factor 2 (lung) (I)	2	14
	Prefoldin 5 (I)	2	18
	Pre-B-cell leukaemia transcription factor interacting protein 1 (I)	0	7
	Activating transcription factor 4 (tax-responsive enhancer element) (I)	0	9
Enzyme activator actitvity	Centaurin, beta 1 (I)	1	12
	Rho GTPase activating protein 4 (I)	0	6
	Ras and Rab interactor 3 (I)	0	9
Channel or pore class transporter actitvity	FXYD domain containing ion transport regulator 5 (I)	1	10

Table VII. Globin gene distributions in the HbS and HbSHC profiles.

Gene	HbS	HbSHC	Fold
α-globin	3951	2210	1·95
β-globin	2499	1486	1·84
Delta	56	45	1·36
ε-globin	5	2	2·73
$^{A}\gamma$-globin	119	321	−2·47
ζ-globin	57	18	3·46

results disagreed with the SAGE data. This gene showed increased expression in reticulocytes and leucocytes after treatment with HC, but the SAGE data showed that *PRDX-6* was repressed after treatment.

Discussion

Gene expression patterns were determined in the bone marrow cells of a SCA patient before and after HC treatment. The bone marrow cells included cells predominantly of the erythroid lineage, but also included other haematopoietic progenitors (Table I). The results of gene expression for *HSPE-1*, *EGR-1*, *PRDX-6* and *HBG* by real-time PCR in reticulocytes and leucocytes obtained in different SCA patients with or without HC treatment strongly indicated that the SAGE pattern observed in bone marrow cells was predominantly representative of erythroid cell lines but also, to a lesser degree, of granulocyte cell lines. The discrepancy observed for *PRDX-6* needs further clarification but may be a consequence of the sum of gene expression in other cell lines, or an individual effect.

The study of gene expression in reticulocytes, which present a large number of different mRNA (Bonafoux *et al*, 2004), provided the opportunity to study just the erythroid lineage, however these mature erythroid cells may present a limitation when compared with the erythroid precursors studied in bone marrow samples. In spite of this limitation, our results indicate that most of the effects observed in bone marrow were also present in reticulocyte.

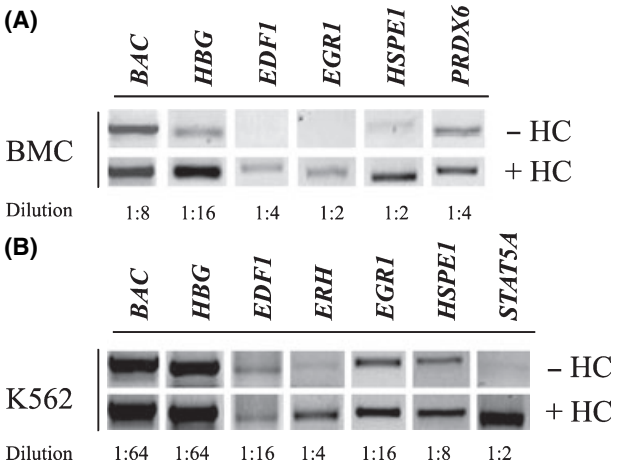

Fig 3. Semi quantitative RT-PCR analysis of the genes in bone marrow cells (BMC) from a SCA patient before and after hydroxycarbamide (HC) treatment (A) and in K562 cells lines treated or not with HC (B). The cDNAs were quantified and equal amounts of cDNA were used in RT-PCR with seven different dilutions for both samples. The dilution that showed a difference between samples was loaded on 2% agarose gels and stained with ethidium bromide. The dilutions are indicated at the bottom of each figure.

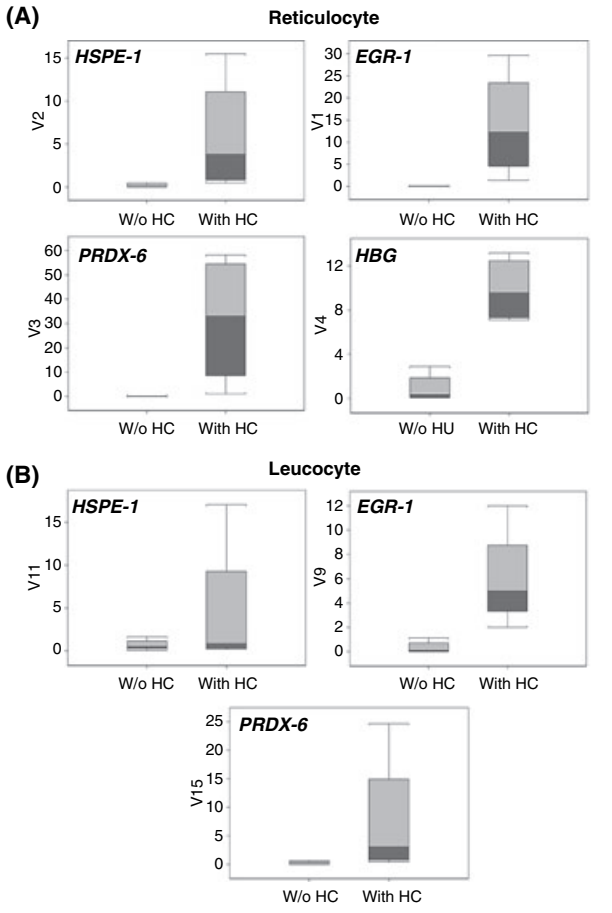

Fig 4. Expression of *EGR1*, *PRDX-6*, *HSPE1* and *HBG* in reticulocytes (A) and *EGR1*, *PRDX-6* and *HSPE1* in leucocytes (B) from the peripheral blood of sickle cell patients on ($n = 4$) and off ($n = 4$) HC therapy.

A comparison of the functional classification of the molecular function ontology of both profiles revealed a number of genes that are induced after HC treatment; One of these is *FXYD5*. FXYD proteins are type I membrane proteins with an extracellular N terminus (sometimes including a signal peptide), a single transmembrane domain and an intracellular C terminus. *FXYD5* has been cloned as a tissue-specific and developmentally regulated gene, induced by the oncoprotein E2a-Pbx1 and termed 'related to ion channel' (Fu & Kamps, 1997). *FXYD5* interacts with the Na, K-ATPase and modulates its properties and at least one of the functional effects of *FXYD5* is to increase maximal pump activity (Lubarski *et al*, 2005). Adragna *et al* (1994) suggested that HC altered the Na and K composition of cultured vascular endothelial cells (VECs) when added after cell attachment, but not during cell seeding. Interestingly, VECs are also involved in the vaso-occlusive mechanism (Adragna *et al*, 1994). A hypothesis that may be suggested is that the induction in expression of these genes associated with channel or pore class transporter activity, after HC administration, may change the intracellular concentration of Na and K in the precursor red cells.

The gene categories associated with the structural constituent of the ribosome, RNA polymerase activity, transcription factor activity and enzyme activator activity also showed increased expression after HC treatment (see details in Table VI). These data suggest that a recruitment of the transcriptional/ translational machinery by HC may occur. Some genes that were induced in these categories included a group of ribosomal proteins, such as ribosomal protein L10a, ribosomal protein large P0 and P2, that are known to play important roles in controlling differentiation, cellular transformation, tumour growth and metastasis. Transcription of a set of ribosomal proteins may be involved in enhancing the transcription of specific genes (Hashimoto *et al*, 2003).

In this study, the transcription factor, *EGR1*, was one of the most upregulated genes following HC treatment. Wang *et al* (2002), using cultured $AC133^+$ human bone marrow stem cells, characterised the gene expression profile in response to HC treatment by cDNA microarray. The treatment of the $AC133^+$ cells with low doses of HC caused a twofold induction of *EGR1* mRNA expression. *EGR1*, also known as *TIS28*, *Krox-24* and *NGFI-A*, is a member of the zinc finger family that encodes for zinc-finger transcription factors that have specificity to related, but not identical, guanine-cytosine-rich DNA binding motifs (Gashler & Sukhatme, 1995; Krishnaraju *et al*, 2001). These were found to be induced in various cell types in response to diverse stimuli. *EGR1* has been reported to be induced by the mitogen-activated protein kinase cascade via phosporylation of Elk-1 or another member of the ternary complex factor (TCF) family (Schaefer *et al*, 2004). Another study suggested that HC may function as a growth factor to upregulate *EGR1* expression and to promote the maturation of erytrhoid progenitors by increasing the total haemoglobin level (Wang *et al*, 2002).

Another group of genes, represented by *CENTB1* (*centaurin, beta 1*), *ARHGAP4* (*Rho GTPase activating protein 4)* and *RIN3* (*Ras and Rab interactor 3*), were found to be induced after HC administration. These genes are members of the GTP-binding protein family involved in signal transduction and the control of an extraordinarily wide variety of cellular processes, including transcriptional activity. *ARHGAP4* has been found to be highly expressed in cells of haematopoieitic origin (including the K562 cell line), suggesting that it may be a haematopoietic-specific rhoGAP (Tribioli *et al*, 1996; Demura *et al*, 2002; Katoh, 2003; Katoh & Katoh, 2004).

The receptor activity was one of the categories that were downregulated after HC treatment, with genes such as colony-stimulating factor 3 receptor, IL-6 signal transducer and other genes (see Table VI for details). Interestingly, *IL6* expression was significantly repressed after HC treatment (21:0). IL-6 is a cytokine with diverse biological functions and involved in triggering inflammation. Pathways, such as JAK-STAT, have been shown to be critical for IL-6 signalling, as the sets of *JAKs* (*JAK1*, *JAK2* and *Tyk2*) and *STATs* (*STAT1*, *STAT3* and *STAT5*) that are activated by this cytokine are also involved in the signal transduction of other growth factors and cytokines (May *et al*, 2003). High levels of circulating IL-6 are found in SCA patients when compared with controls (Taylor *et al*, 1995). IL-6 was also demonstrated to reduce steady-state HbF synthesis in fetal erythroid progenitors and inhibit γ-globin expression via the STAT3β protein (Ferry *et al*, 1997; Foley *et al*, 2002). Levels of IL-6 are reported to decrease after treatment of K562 cell lines with γ-globin inducers such as hemin and sodium butyrate (Ferry *et al*, 1997; Foley *et al*, 2002). These data may suggest that the induction of γ-globin would be associated with the suppression of *IL6* gene expression.

Jison *et al* (2004) analysed the peripheral blood mononuclear cells of a group of sickle cell patients on and off HC treatment using the microarray technique to evaluate the gene expression pattern in SCA (Jison *et al*, 2004). Notably, SCA patients off HC therapy presented very similar gene expression patterns to SCA patients on HC and were clearly distinguished from controls; the data suggested that HC does not have a direct effect on the gene expression of leucocytes in SCA patients. The data from bone marrow cells demonstrated herein indicates a different expression pattern. This may reflect the fact that the cell type used for the gene expression analysis were different, and may indicate that HC could mediate its effect mainly in a population of primitive erythroid progenitor cells in the human bone marrow.

Taken together, these results strongly suggest that HC produces significant changes in the gene expression pattern of human bone marrow cells in a SCA patient. Our data suggest a recruitment of transcriptional machinery and pathways involved in signal transduction during HC treatment. These pathways may have a direct effect, inducing globin gene expression at the transcriptional level. Conversely, pathway stimulation associated with channel or pore class transporter activity indicates a possible interference with the decrease in the dehydration status of cells, interfering with the rheological properties of the cell. The study of genes associated with these pathways and the description of new genes that are up- or downregulated by HC may also represent a potential tool to identify new targets for SCA patient therapy.

Acknowledgements

We would like to thanks Dr Nicola Conran for English review. This study was supported by CAPES/ FAPESP (Proc. no. 02/13801-7), Brazil.

References

Adragna, N.C., Fonseca, P. & Lauf, P.K. (1994) Hydroxyurea affects cell morphology, cation transport, and red blood cell adhesion in cultured vascular endothelial cells. *Blood*, **83,** 553–560.

Audic, S. & Claverie, J.M. (1997) The significance of digital gene expression profiles. *Genome Research*, **7,** 986–995.

Bonafoux, B., Lejeune, M., Piquemal, D., Quere, R., Baudet, A., Assaf, L., Marti, J., Aguilar-Martinez, P. & Commes, T. (2004) Analysis of remnant reticulocyte mRNA reveals new genes and antisense transcripts expressed in the human erythroid lineage. *Haematologica*, **89,** 1434–1438.

Bunn, H.F. (1997) Pathogenesis and treatment of sickle cell disease. *New England Journal of Medicine*, **337,** 762–769.

Chabardes-Garonne, D., Mejean, A., Aude, J.C., Cheval, L., Di Stefano, A., Gaillard, M.C., Imbert-Teboul, M., Wittner, M., Balian, C., Anthouard, V., Robert, C., Segurens, B., Wincker, P., Weissenbach, J., Doucet, A. & Elalouf, J.M. (2003) A panoramic view of gene expression in the human kidney. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 13710–13715.

de Chaldee, M., Gaillard, M.C., Bizat, N., Buhler, J.M., Manzoni, O., Bockaert, J., Hantraye, P., Brouillet, E. & Elalouf, J.M. (2003) Quantitative assessment of transcriptome differences between brain territories. *Genome Research*, **13,** 1646–1653.

Charache, S. (1997) Mechanism of action of hydroxyurea in the management of sickle cell anemia in adults. *Seminars in Hematology*, **34**(3 Suppl. 3), 15–21.

Cokic, V.P., Smith, R.D., Beleslin-Cokic, B.B., Njoroge, J.M., Miller, J.L., Gladwin, M.T. & Schechter, A.N. (2003) Hydroxyurea induces fetal hemoglobin by the nitric oxide-dependent activation of soluble guanylyl cyclase. *The Journal of Clinical Investigation*, **111,** 231–239.

Demura, M., Takeda, Y., Yoneda, T., Furukawa, K., Usukura, M., Itoh, Y. & Mabuchi, H. (2002) Two novel types of contiguous gene deletion of the AVPR2 and ARHGAP4 genes in unrelated Japanese kindreds with nephrogenic diabetes insipidus. *Human Mutation*, **19,** 23–29.

Ewing, B. & Green, P. (1998) Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Research*, **8,** 186–194.

Ferry, A.E., Baliga, S.B., Monteiro, C. & Pace, B.S. (1997) Globin gene silencing in primary erythroid cultures. An inhibitory role for interleukin-6. *The Journal of Biological Chemistry*, **272,** 20030–20037.

Foley, H.A., Ofori-Acquah, S.F., Yoshimura, A., Critz, S., Baliga, B.S. & Pace, B.S. (2002) Stat3 beta inhibits gamma-globin gene expression

in erythroid cells. *The Journal of Biological Chemistry*, **277,** 16211–16219.

Fu, X. & Kamps, M.P. (1997) E2a-Pbx1 induces aberrant expression of tissue-specific and developmentally regulated genes when expressed in NIH 3T3 fibroblasts. *Molecular and Cellular Biology*, **17,** 1503–1512.

Gashler, A. & Sukhatme, V.P. (1995) Early growth response protein 1 (Egr-1): prototype of a zinc-finger family of transcription factors. *Progress in Nucleic Acid Research and Molecular Biology*, **50,** 191–224.

Goossens, M. & Kan, Y.Y. (1981) DNA analysis in the diagnosis of hemoglobin disorders. *Methods in Enzymology*, **76,** 805–817.

Halsey, C. & Roberts, I.A. (2003) The role of hydroxyurea in sickle cell disease. *British Journal of Haematology*, **120,** 177–186.

Hashimoto, S., Nagai, S., Sese, J., Suzuki, T., Obata, A., Sato, T., Toyoda, N., Dong, H.Y., Kurachi, M., Nagahata, T., Shizuno, K., Morishita, S. & Matsushima, K. (2003) Gene expression profile in human leukocytes. *Blood*, **101,** 3509–3513.

Ikuta, T., Ausenda, S. & Cappellini, M.D. (2001) Mechanism for fetal globin gene expression: role of the soluble guanylate cyclase-cGMP-dependent protein kinase pathway. *Proceedings of the National Academy of Sciences of the United States of America*, **98,** 1847–1852.

Jison, M.L., Munson, P.J., Barb, J.J., Suffredini, A.F., Talwar, S., Logun, C., Raghavachari, N., Beigel, J.H., Shelhamer, J.H., Danner, R.L. & Gladwin, M.T. (2004) Blood mononuclear cell gene expression profiles characterize the oxidant, hemolytic, and inflammatory stress of sickle cell disease. *Blood*, **104,** 270–280.

Katoh, M. (2003) FNBP2 gene on human chromosome 1q32.1 encodes ARHGAP family protein with FCH, FBH, RhoGAP and SH3 domains. *International Journal of Molecular Medicine*, **11,** 791–797.

Katoh, Y. & Katoh, M. (2004) Identification and characterization of ARHGAP27 gene in silico. *International Journal of Molecular Medicine*, **14,** 943–947.

Krishnaraju, K., Hoffman, B. & Liebermann, D.A. (2001) Early growth response gene 1 stimulates development of hematopoietic progenitor cells along the macrophage lineage at the expense of the granulocyte and erythroid lineages. *Blood*, **97,** 1298–1305.

Lal, A., Lash, A.E., Altschul, S.F., Velculescu, V., Zhang, L., McLendon, R.E., Marra, M.A., Prange, C., Morin, P.J., Polyak, K., Papadopoulos, N., Vogelstein, B., Kinzler, K.W., Strausberg, R.L. & Riggins, G.J. (1999) A public database for gene expression in human cancers. *Cancer Research*, **59,** 5403–5407.

Lubarski, I., Pihakaski-Maunsbach, K., Karlish, S.J., Maunsbach, A.B. & Garty, H. (2005) Interaction with the Na, K-ATPase and tissue distribution of FXYD5 (related to ion channel). *The Journal of Biological Chemistry*, **280,** 37717–37724.

Malago, Jr, W., Sommer, C.A., Del Cistia Andrade, C., Soares-Costa, A., Abrao Possik, P., Cassago, A., Santejo Silveira, H.C. & Henrique-Silva, F. (2005) Gene expression profile of human Down syndrome leukocytes. *Croatian Medical Journal*, **46,** 647–656.

Margulies, E.H. & Innis, J.W. (2000) eSAGE: managing and analysing data generated with serial analysis of gene expression (SAGE). *Bioinformatics*, **16,** 650–651.

May, P., Schniertshauer, U., Gerhartz, C., Horn, F. & Heinrich, P.C. (2003) Signal transducer and activator of transcription STAT3 plays a major role in gp130-mediated acute phase protein gene activation. *Acta Biochimica Polonica*, **50,** 595–601.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, **29,** e45.

Pollock, J.D. (2002) Gene expression profiling: methodological challenges, results, and prospects for addiction research. *Chemistry and Physics of Lipids*, **121,** 241–256.

Schaefer, A., Kosa, F., Bittorf, T., Magocsi, M., Rosche, A., Ramirez-Chavez, Y., Marotzki, S. & Marquardt, H. (2004) Opposite effects of inhibitors of mitogen-activated protein kinase pathways on the egr-1 and beta-globin expression in erythropoietin-responsive murine erythroleukemia cells. *Cellular Signalling*, **16,** 223–234.

Schwering, I., Brauninger, A., Distler, V., Jesdinsky, J., Diehl, V., Hansmann, M.L., Rajewsky, K. & Kuppers, R. (2003) Profiling of Hodgkin's lymphoma cell line L1236 and germinal center B cells: identification of Hodgkin's lymphoma-specific genes. *Molecular Medicine*, **9,** 85–95.

Steinberg, M.H., Barton, F., Castro, O., Pegelow, C.H., Ballas, S.K., Kutlar, A., Orringer, E., Bellevue, R., Olivieri, N., Eckman, J., Varma, M., Ramirez, G., Adler, B., Smith, W., Carlos, T., Ataga, K., DeCastro, L., Bigelow, C., Saunthararajah, Y., Telfer, M., Vichinsky, E., Claster, S., Shurin, S., Bridges, K., Waclawiw, M., Bonds, D. & Terrin, M. (2003) Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to 9 years of treatment. *JAMA: the Journal of the American Medical Association*, **289,** 1645–1651.

Taylor, S.C., Shacks, S.J., Mitchell, R.A. & Banks, A. (1995) Serum interleukin-6 levels in the steady state of sickle cell disease. *Journal of Interferon & Cytokine Research*, **15,** 1061–1064.

Tribioli, C., Droetto, S., Bione, S., Cesareni, G., Torrisi, M.R., Lotti, L.V., Lanfrancone, L., Toniolo, D. & Pelicci, P. (1996) An X chromosome-linked gene encoding a protein with characteristics of a rhoGAP predominantly expressed in hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America*, **93,** 695–699.

Varela, C., Cardenas, J., Melo, F. & Agosin, E. (2005) Quantitative analysis of wine yeast gene expression profiles under winemaking conditions. *Yeast*, **22,** 369–383.

Velculescu, V.E., Zhang, L., Vogelstein, B. & Kinzler, K.W. (1995) Serial analysis of gene expression. *Science*, **270,** 484–487.

Wang, M., Tang, D.C., Liu, W., Chin, K., Zhu, J.G., Fibach, E. & Rodgers, G.P. (2002) Hydroxyurea exerts bi-modal dose-dependent effects on erythropoiesis in human cultured erythroid cells via distinct pathways. *British Journal of Haematology*, **119,** 1098–1105.