

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Identification of target genes for wild type and truncated HMGA2 in mesenchymal stem-like cells

BMC Cancer 2010, **10**:329 doi:10.1186/1471-2407-10-329

Jorn Henriksen (jorn.henriksen@imbv.uio.no)
Marianne Stabell (marianne.stabell@imbv.uio.no)
Leonardo A Meza-Zepeda (Leonardo.Meza-Zepeda@rr-research.no)
Silje AU Lauvrak (Silje.Anett.Ugland.Lauvrak@rr-research.no)
Moustapha Kassem (mkassem@health.sdu.dk)
Ola Myklebost (ola.myklebost@imbv.uio.no)

ISSN 1471-2407

Article type Research article

Submission date 15 January 2010

Acceptance date 25 June 2010

Publication date 25 June 2010

Article URL <http://www.biomedcentral.com/1471-2407/10/329>

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

Identification of target genes for wild type and truncated HMGA2 in mesenchymal stem-like cells

Jørn Henriksen^{*}, Marianne Stabell^{*}, Leonardo A. Meza-Zepeda^{*#}, Silje AU Lauvrak^{*}, Moustapha Kassem[§] and Ola Myklebost^{*#}

^{*} CAST, Cancer Stem Cell Innovation Centre, Department of Tumor Biology, The Norwegian Radium Hospital, Oslo University Hospital, N-0027 Oslo, Norway, [#] Norwegian Microarray Consortium, Department of Molecular Bioscience, University of Oslo, Oslo, Norway, [§] Department of Endocrinology, University Hospital of Odense, Denmark & Stem Cell Unit, College of Medicine, King Saud University, Riyadh, KSA

Corresponding author: OM, Tel: +47-2278-1779, Fax: +47-2278-1795,

OM: ola.myklebost@imbv.uio.no

JH: jorn.henriksen@imv.uio.no

MS: marianne.stabell@rr-research.no

LAMZ: leonardo.meza-zepeda@rr-research.no

SAUL: silje.u.lauvrak@rr-research.no

KM: mkassem@health.sdu.dk

Abstract

Background

The HMGA2 gene, coding for an architectural transcription factor involved in mesenchymal embryogenesis, is frequently deranged by translocation and/or amplification in mesenchymal tumours, generally leading to over-expression of shortened transcripts and a truncated protein.

Methods

To identify pathways that are affected by sarcoma-associated variants of HMGA2, we have over-expressed wild type and truncated HMGA2 protein in an immortalized mesenchymal stem-like cell (MSC) line, and investigated the localisation of these proteins and their effects on differentiation and gene expression patterns.

Results

Over-expression of both transgenes blocked adipogenic differentiation of these cells, and microarray analysis revealed clear changes in gene expression patterns, more pronounced for the truncated protein. Most of the genes that showed altered expression in the HMGA2-overexpressing cells fell into the group of NF- κ B-target genes, suggesting a central role for HMGA2 in this pathway. Of particular interest was the pronounced up-regulation of SSX1, already implicated in mesenchymal oncogenesis and stem cell functions, only in cells expressing the truncated protein. Furthermore, over-expression of both HMGA2 forms was associated with a strong repression of the epithelial marker CD24, consistent with the reported low level of CD24 in cancer stem cells.

Conclusions

We conclude that the c-terminal part of HMGA2 has important functions at least in mesenchymal cells, and the changes in gene expression resulting from overexpressing a protein lacking this domain may add to the malignant potential of sarcomas.

Background

HMGA2 belongs to a family of nuclear proteins that contain an N-terminal part that recognizes and binds to AT-rich regions in the DNA [1] and an acidic C-terminal tail that probably modulates their interactions with DNA [2], and proteins [3]. HMGA proteins are not transactivators on their own, but modulate the assembly of transcriptional complexes at various levels [4-8], and may also regulate gene transcription through protein-protein interactions without direct contact with DNA [9-11] or by altering chromatin structure [12-16]. Interestingly, a recent study suggested that both HMGA1 and HMGA2 are taking active parts in the formation of senescence-associated heterochromatin foci and maintenance of the growth-arrested state [17, 18].

HMGA2 is expressed during embryonic development while it is almost undetected in normal adult tissues [19, 20]. HMGA2 plays a critical role early in adipogenesis, probably by regulating the proliferation of preadipocytes during differentiation [21], and mice lacking functional HMGA2 protein exhibit a *pygmy* phenotype with drastic reduction of adipose tissue [22]. On the other hand, over-expression of a truncated HMGA2 protein lacking the acidic C-terminal tail results in a giant phenotype with multiple lipomas [23, 24].

Aberrations in the chromosomal region 12q14-15 that affect *HMGA2* are frequent in a variety of tumours. In benign tumours of mesenchymal origin, *HMGA2* is often rearranged by translocation, and the resulting chimeric transcripts are formed by fusion of the DNA-binding domains, coded by exons 1-3, to ectopic sequences [25-27]. In sarcomas, *HMGA2* is frequently and selectively amplified and rearranged [28]. We have cloned and sequenced a number of these cancer-associated ectopic

sequences from 12q as well as other chromosomes [29]. The only common factor that was found was the loss of the 3'-untranslated region (UTR) as well as the last two exons, resulting in fusion proteins containing as little as one extra amino acid replacing the C-terminal part [29].

Recently it was shown that the HMGA2 3'-UTR contains target sites for the *let-7* miRNA, and thus the above mentioned rearrangements lead to increased levels of HMGA2 protein due to loss of miRNA-mediated repression [30, 31]. Thus the attention has shifted from possible oncogenic effects of loss of the acidic domain to effects of increased expression, even of the wild type protein. Furthermore, it turns out that the balance between *let-7* and HMGA2 governs the exit of cells from the undifferentiated and self-renewing state, and HMGA2 is now thought to be central in cancer in general [32-35]. Because HMGA2 most frequently is rearranged in well-differentiated liposarcomas, border-line tumours resembling adipose tissue, most sarcoma cell lines, representing highly malignant cancers with a different tissue type, would not be appropriate to detect a phenotype when the gene is over-expressed. We therefore chose an immature, stem-like mesenchymal cell line with adipogenic potential to investigate whether the wild type and truncated forms of the protein activate the same pathways, by performing gene expression analysis.

Methods

DNA constructs

The coding sequence and the part of the HMGA2 gene encoding the DNA-binding domain (exon 1-3) were PCR-amplified from pBluescript-HMGI-C [27] and cloned into pEGFP-C3 (Clontech, Palo Alto, CA, USA) generating pEGFP-HMGA2_{WT} and pEGFP-HMGA2_{TRUNC}, respectively.

Cell culture and stable transfectants

The human mesenchymal stem cell line (hMSC) transduced with human telomerase reverse transcriptase (hTERT), hMSC-TERT20 cells [36, 37] were transfected with the HMGA2 constructs described above, using Fugene (Roche Diagnostics, Basel, Switzerland) according to the manufacturer's protocol. Stable clones were selected with 400 µg/ml G418 (Gibco BRL, Gaithersburg, MD, USA) and clones with strong GFP-fluorescence were established using fluorescence activated cell sorting. Cells were maintained in RPMI1640 (Gibco) supplemented with 10% fetal bovine serum and 100 µg/ml G418.

Fluorescence microscopy

Transfected hMSC-TERT20 cells grown on coverslips were fixed with 10% formalin solution (Sigma-Aldrich, St. Louis, MO, USA), permeabilized with 0.1% Triton X-100 in PBS and blocked with 5% FCS. Staining for HMGA2 and B23 (nucleophosmin - a marker for the nuclear matrix)[38] was performed using rabbit-anti HMGA2 (Strategic Diagnostics Inc., Newark, DE, USA), and mouse-anti B23 followed by DyLight549-conjugated donkey anti-rabbit and DyLight649-conjugated donkey anti-mouse antibodies (Jackson ImmunoResearch Laboratories, West Grove, PA, USA), respectively. The samples were mounted in Prolong Gold antifade reagent with DAPI

(Invitrogen, Carlsbad, CA, USA). The fluorescence was visualized using a Zeiss LSM 510 confocal microscope and pictures were taken of thin single plane sections.

Differentiation assay

The hMSC-TERT20 cells as described above, were seeded at 10 000 cells/cm² in medium supplemented with a cocktail of 0.5 mM methylisobutylxanthine (Sigma-Aldrich, St. Louis, MO, USA), 1 µM dexamethasone (Decadron™, MSD, Haarlem, Nederland), 2 µM insulin (Sigma-Aldrich) and 20 µM Rosiglitazone (GlaxoSmithKline, Harlow, UK) for 14 days. To monitor the differentiation process, cultures were stained with Oil Red O (Serva, Heidelberg, Germany). The stained neutral lipids were extracted and the absorbance at 490 nm measured. The number of cells was estimated by using SRB protein assay on the same cultures. Briefly, proteins were precipitated with 10% trichloroacetic acid (Merck, Darmstadt, Germany) and assayed by staining with sulforhodamine B (Sigma-Aldrich), extraction and measuring the dye absorbance at 540 nm [39].

Northern blot analysis and quantitative RT-PCR

Total RNA was isolated from sub-confluent growing cells using Trizol reagent (Invitrogen) according to the manufacturer's instructions. For Northern blot analysis, the samples of RNA (10 µg) were run on a 1% formaldehyde agarose gel, blotted onto a positively charged nylon membrane and hybridizations were performed as described [40]. Probes were sequentially hybridized to the membrane and signals from an 18S rRNA oligonucleotide were used to correct for unequal sample loading.

For quantitative RT-PCR, 0.5 µg of RNA from each sample was reverse transcribed using the High-Capacity cDNA Archive Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. The resulting cDNAs were used for real-time PCR using TaqMan Gene Expression Assays (Applied BioSystem;

Additional file 1 Table S1). Threshold cycle (C_t) values were measured and the comparative C_t method [41] was used to calculate the \log_2 fold difference in transcript level of the over-expressing cell lines relative to the parental. See Additional file 1 Table S1 for probes.

Western blot analysis

Cells were harvested and protein lysates were prepared by sonicating cells resuspended in a nuclear lysis buffer of 50 mM Tris-HCl, pH 7.5, 0.1 % SDS, 0.5 M NaCl, 1 % NP-40, 1 % DOC, 2 mM EDTA supplemented with protease inhibitors (all from Sigma-Aldrich). 5 μ g protein were subjected to SDS-PAGE, transferred onto an Immobilon-P membrane (Millipore, Billerica, MA, USA) and probed with anti-HMGA2 (Santa Cruz Technologies, Santa Cruz, CA, USA) and anti-tubulin (Santa Cruz). Secondary antibodies conjugated to horse radish peroxidase (Dako, Glostrup, Denmark) were used according to the manufacturer's instructions. Detection of antibody signals was performed with chemiluminescence using SuperSignal West Dura Extended Duration Substrate (Pierce, Rockford, IL, USA).

Microarray analysis

Total RNA was isolated from sub-confluent growing cells using Trizol reagent according to the manufacturer's protocol and global expression analysis was performed on HG-U133 Plus 2.0 oligonucleotide arrays (Affymetrix, Santa Clara, CA, USA). Labelling, hybridisation and scanning were performed at the Rikshospitalet University of Oslo Microarray Core Facility (<http://microarray.rikshospitalet.no>) using standard conditions. As we aimed to detect changes above the variation expected from different sub-lines, i.e. above 3-fold, one array per sample was sufficient (the technical variation between arrays being much smaller).

Flow cytometry analysis of surface markers

To verify differential expression of CD24 and HLA-DR on the cell surface, subconfluent growing cells were harvested, incubated with phycoerythrin-conjugated antibodies and rinsed several times in PBS. IFN γ -treated cells were cultured with 100 U/ml IFN γ (Sigma) for 40 hours prior to antibody staining. Antibody staining was detected with a triple-focus DIVA flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). Antibodies used; anti-HLA-DR (Becton Dickinson); anti-CD24 (Beckman-Coulter, Fullerton, CA, USA). Data were analyzed with FlowJo software (Tree star, Ashland, Oregon, USA).

Analysis of promoter sequences

Analysis of putative transcription factor binding sites was performed using MatInspector (www.genomatix.de/cgi-bin/matinspector/matinspector.pl) and TESS (<http://www.cbil.upenn.edu/tess/>). The genomic DNA sequences were retrieved from the Ensembl genome browser (<http://www.ensembl.org/index.html>) covering 1500 base pairs of the 5' flanking regions (upstream from the transcriptional start sites) and 500 base pairs downstream. A core similarity of at least 0.80 from MatInspector or TESS analysis was used as cut-off for potential match to known transcription factor recognition sequences.

Microarray data analysis

For GeneChip arrays, the raw intensity of individual samples was normalized using Bioconductor and the GC-RMA algorithm (<http://www.bioconductor.org>) [42]. Normalized gene expression levels for each transfected cell line were compared with the parental cell line and probe sets that exhibited less than 3-fold intensity change were removed from further analysis. The remaining probe sets were initially annotated according to entries in the NetAffx database (<http://www.affymetrix.com>).

Using the Database for Annotation, Visualization and Integrated Discovery (DAVID located at <http://apps1.niaid.nih.gov/david>) [43], lists of differentially expressed genes, common or specific for the samples, were analysed to identify Gene Ontology (GO) terms that were significantly overrepresented using a modified Fisher Exact test (EASE-score<0.05). The redundancy in the annotations was reduced by applying the clustering algorithm available in DAVID using intermediate stringency and default settings for the Kappa statistics.

Results

Exogenous HMGA2 is localized to the nucleus

To be able to monitor localization of the proteins, an enhanced green fluorescence protein (eGFP) tag was fused to the amino terminal end of both HMGA2 proteins. Cells expressing eGFP-HMGA2_{WT} or eGFP-HMGA2_{TRUNC} accumulate the eGFP-tagged proteins in the nuclei, especially in DAPI-dense and nucleolar regions (Fig. 1a), whereas eGFP alone showed general fluorescence throughout the cell, with lower level in nucleoli than in the rest of the nuclei (not shown). Noteworthy, both endogenous and fusion proteins co-localized with DAPI-dense heterochromatin but they were also present in numerous discrete foci that did not stain well by DAPI. Moreover, both fusion proteins appeared to be enriched in the nucleoli, detected by immunostaining against the nuclear matrix marker nucleophosmin/B23 [38]. Immunostaining for HMGA2, on the other hand, gave a more even fluorescence also covering the nucleoli for both endogenous and ectopic HMGA2, and it is unclear whether this indicates that the apparent nucleolar enrichment of eGFP fusion proteins was a result of overexpression, or reduced nucleolar immunostaining could be due to incomplete penetrance of antibodies into the nucleoli compared to the efficient fluorescence from the fluorescent eGFP tags. Interestingly, endogenous HMGA2 seemed to be distributed throughout the granular region in the nucleolus whereas the granular nucleophosmin/B23 was restricted to the periphery. In addition, the fusion proteins stayed bound to the condensed chromosomes during mitosis (data not shown).

The endogenous *HMGA2* transcript and protein was readily detected in hMSC-TERT20 cells (Fig. 1b and 1c), and the ectopic transcript showed a lower steady-state

level than the endogenous transcript, as less than two-fold increase of HMGA2 transcripts was measured by quantitative PCR in the over-expressing cell lines. However, the levels of ectopically expressed protein was increased several fold relative to endogenous HMGA2.

HMGA2 abolishes differentiation into the adipocytic lineage

Compared to the parental cells, forced expression of full-length or truncated HMGA2 abolished the ability of the cells to differentiate towards adipocytes and accumulate fat (Fig. 1d). Noteworthy, the growth of the HMGA2-overexpressing cells was unaffected by addition of differentiation medium, whereas the proliferation of the parental cells was inhibited when differentiation was induced. A control cell line expressing eGFP only showed growth and differentiation patterns similar to the parental line (i.e. within the variation expected when different passages of these cells was used, Fig 1d). The proliferation of the different transfectants was indistinguishable from that of the parental cells in the absence of differentiation (not shown).

Identification of transcriptional targets

To find out how HMGA2 affects transcription, and possible effects of truncation, we analyzed the gene expression profiles of proliferating cultures of the over-expressing and the parental cell lines using the HG-U133 Plus 2.0 Affymetrix GeneChips. We identified 408 transcripts that were up-regulated more than 3-fold relative to the recipient cells in either of the two HMGA2-expressing lines, while 477 transcripts were down-regulated more than 3-fold. The microarray data were deposited in the Array Express public database (<http://www.ebi.ac.uk/microarray-as/ae/>, accession number E-MEXP-2647), and a list of the most differentially expressed genes is shown

in Table 1 (the complete list of transcripts changed more than 3-fold is given in Additional file 2 Table S2). Although the microarray technique can be used to detect much smaller changes, we chose 3-fold as a level where we would avoid smaller differences due to the prolonged separate culturing of each transfectant line.

Interestingly, although both protein variants similarly affected many genes, a large number of genes were differently affected. Out of 408 genes that showed up-regulation, 66 genes were up regulated by both full-length and truncated HMGA2. In addition, 84 and 258 were up-regulated in the cells expressing full-length and truncated HMGA2 respectively. On the other hand, 139 genes were down-regulated by both proteins while 99 and 239 were down-regulated by full-length and truncated HMGA2 respectively. As the cells were analysed in their undifferentiated state, we could not monitor genes directly involved in differentiation. Although we detected no change of the known HMGA2 target *Imp2 (IGF2BP2)* [44] in this system, the about 2-fold down regulation by both protein variants of both PPAR α and PPAR γ , central regulators of adipogenesis, could contribute to the differentiation block. In near all cases quantitative real-time PCR (qPCR) analyses confirmed the direction of changes in transcript levels observed by microarrays (Table 1). An exception was members of the SSX gene family, which were poorly distinguished by the microarray probes, and of which only SSX1 was confirmed by real-time PCR. Discrepant results were also found for *LCPI*, *FGF1* and *HDAC9*, which most likely were caused by alternative splicing according to entries in the Alternative Splicing and Transcript Diversity database (<http://www.ebi.ac.uk/astd/main.html>).

The over all correlation between microarray and real-time ratios between HMGA2 transfectants and parental cells was 0,96. The cells transfected with eGFP only were

included in the real-time validation as indicated in Table 1, and although there were some discrepancies, the overall correlation between real-time ratios relative to the eGFP-transfectants and those for the parental cells was 0.94, indicating their high degree of similarity.

Transcriptional effects common to wild type and truncated HMGA2

Sixty-six genes were up- and 139 genes were down-regulated by both proteins. A large fraction of the affected genes are known direct targets for NF- κ B, and potential intrinsic binding sites for HMGA proteins could be found within many of the NF- κ B sites (Additional file 3 Table S3). Most of these genes code for proteins involved in cytokine signalling. All of the genes in the *CXC* chemokine cluster in chromosome band 4q13.3 that were expressed in the parental cell line were down-regulated by both protein variants. Other cytokines that were initially expressed in the parental cells and down-regulated upon HMGA2 expression were *IL1B*, *IL1RN*, *IL6*, *IL11*, *IL15* and *LIF1*, all containing validated or predicted NF- κ B sites and only *IL11* and *IL15* lack intrinsic binding sites for HMGA proteins. Several other possible NF- κ B targets were down regulated by the over-expression of either form of HMGA2, e.g. *THBS2*, *S100A4*, *IFI44L*, *SOD2*, *PSMB9*, *IGFBP1*, *SERPINB1*, *NFKBIA*, *BDKRB1* and *PTGES*, all associated with increased NF- κ B activity (<http://people.bu.edu/gilmore/nf-kb/target/index.html> and references therein). We also observed a four-fold decrease of the primary microRNA miR-146A (MIRN146A), which is known to be regulated by NF- κ B [45]. Noteworthy, a less stringent criterion (>2-fold) showed down-regulation of several other potential NF- κ B target genes, such as *BCL2A1*, *JUNB*, *LGALS3*, *HAS1*, *TWIST1*, *PTGS2*, *CFLAR* and *BIC* (encoding

miR-155). However, *EDN1*, *MDK* and *TGM2*, three genes thought to be positively regulated by NF- κ B [46-48], were strongly up-regulated in both cell lines.

In addition, several putative interferon-responsive genes were down-regulated, such as *IFI27*, *IFITM1*, *ISGF3G*, *AIM2*, *KYNU*, *C1S* and *OASL*. Since the regulatory region of both *IFI27* and *IFITM1* contain IFN- α stimulated response element (ISRE; Additional file 4 Table S4) that are recognized by the ISGF3 complex [49] they work downstream of *ISGF3G* and are therefore strongly dependent on the expression of the *ISGF3G* gene. The *ISGF3G* gene itself is known to be regulated by IL-6 [50] which indicates that these genes are secondary targets to HMGA2.

Another large group of genes that showed expression change was the Krüppel-type zinc finger proteins located in several gene clusters mainly on the long arm of chromosome 19 [51, 52]. Some of these zinc finger proteins contain a potent repressor domain, the Krüppel-associated box (KRAB), and they are likely involved in gene silencing [53]. Interestingly, there was a broader up-regulation of several other genes mapped to the chromosome band 19q13 in the cell line over-expressing the truncated HMGA2, including 35 Krüppel-type zinc finger genes of which 27 contain a KRAB domain. This might indicate a common regulatory mechanism for these genes that could involve modulation of chromatin since there was also an increased expression of several unrelated genes near these clusters. Cells expressing HMGA2_{WT} had an increased expression of 12 Krüppel-type zinc finger genes, nine of which mapped to the long arm of chromosome 19 and were up-regulated also in HMGA2_{TRUNC} transfectants. Six of the common up-regulated genes that mapped to 19q contain a KRAB repressor domain, while the three genes mapped to other chromosome regions,

showing elevated expression only in the cells expressing HMGA2_{WT}, all contain KRAB domains.

A gene ontology (GO) analysis revealed a highly significant decrease in expression of genes encoding proteins involved in signal transduction ($p < 0.001$) and response to wounding ($p < 0.001$), mainly affecting genes involved in cytokine-receptor signalling ($p < 0.001$), chemotaxis ($p < 0.001$) and inflammatory response ($p < 0.001$). In addition, down-regulation of genes involved in regulation of cellular proliferation ($p = 0.0033$) and epithelial cell differentiation ($p = 0.026$) was observed. A few genes involved in vascular development ($p = 0.033$) were up-regulated. However, six out of 66 up-regulated genes encode Krüppel-type zinc finger proteins containing a KRAB repressor domain ($p = 0.0051$), indicating a common regulatory mechanism for these genes (Table 2).

Reciprocal effects of different forms of HMGA2

There were several genes up-regulated by HMGA2_{WT} and down-regulated in cells expressing the truncated form, such as *FGF13*, *EHF*, *HCLS1*, *MEST*, *G0S2* and *PTPRN2*. Two of the gene products, G0S2 and MEST, are both implicated in adipogenesis [54-56] while *FGF13*, *EHF*, *HCLS1* and *PTPRN2* are potential regulators of stem cell activity [57], senescence [58], adhesion [59] and insulin secretion [60], respectively.

Transcriptional effects specific to full-length HMGA2

Eighty-four genes were up-regulated and 99 genes were down-regulated more than three-fold by wild-type but not truncated HMGA2. Most genes were classified into groups of more general GO terms providing limited information about the underlying processes affected by full-length HMGA2. However, a significant number of the up-

regulated genes are cell-type-specific genes involved in embryonal development and likely reflect a transition in the differentiation state of the cells. Some of the down-regulated genes affect gene transcription. However, most of the affected genes were poorly annotated and gave little information about their cellular function.

Transcriptional effects specific to truncated HMGA2

Two hundred and fifty-eight genes were up-regulated and 239 genes were down-regulated in the cells expressing the truncated protein. HMGA2_{TRUNC} cells showed a clear increase in expression of *CXCL12* whereas the large MHC class II gene cluster was down-regulated (Table 1b). The down-regulation of the MHC class II gene HLA-DRA was verified by qPCR (Table 1b) and using flow cytometry, we could show that HLA-DRA was expressed at a basal level on the surface of both the parental cells and the cells over-expressing the wild type HMGA2, but was strongly down-regulated in HMGA2_{TRUNC} cells in according with the expression data (Fig. 1f). Adding interferon γ increased the expression of HLA-DRA, resulting in a 40-fold increase in the level on the HMGA2_{TRUNC} cells, to give a level of surface protein similar to that on the parental line (not shown).

Based on the microarray data, the SSX gene family appeared to be co-induced in cells expressing HMGA2_{TRUNC}, but real-time PCR was only able to verify the induction of SSX1 transcripts. Interesting, the expression increased by 1000-fold indicating a specific activation of this gene.

The GO analysis revealed a significant down-regulation of genes involved in antigen presentation and processing ($p < 0.001$), and organism development ($p = 0.0021$). A large proportion of the up-regulated genes encode proteins with roles in regulation of transcription ($p < 0.001$) and most of them were KRAB-zinc finger proteins. In

addition, genes encoding proteins involved in intracellular protein transport across membranes ($p = 0.0032$), organization of cytoskeletal structures ($p = 0.042$) and Ras signalling ($p = 0.031$), were significant up-regulated.

Discussion

The deregulation of expression caused by the absence of the 3'-UTR in our transgenes, and thus loss of *let-7*-mediated inhibition, resulted in abundant expression of HMGA2 protein (Fig. 1c). The eGFP fusion proteins were shown to localise in the nucleus just like the endogenous protein. Their distribution to the nucleoli (Fig. 1a, arrowheads) was unexpected though HMGA2 has previously been found to interact with the mainly nucleolar nucleophosmin/B23 protein [61].

Furthermore, the localization of endogenous HMGA2 in the nucleoli was clearly restricted to numerous discrete foci (Fig. 1a, arrow). The distribution of these sites was occasionally scattered throughout the nucleolus or more often restricted to the periphery of the nucleolar body where they seemingly co-localized with B23. In this outer part of nucleolus the late processing of rRNAs takes place before they assemble into ribosomal subunits [62], suggesting that HMGA2 might be involved in these processes. However, this could also be a result of B23 acting as a molecular chaperone [63], sequestering HMGA2 to the nucleolus, as described for IRF-1 [64], pRb [65] and TRF2 [66].

Both wild type and truncated HMGA2 abrogated growth inhibition and adipogenesis (Fig. 1d), probably by a common mechanism. It appears that both HMGA2 proteins prevented the growth arrest necessary for adipogenic differentiation, as cellular protein accumulated at the same rate as in un-induced cells. Interestingly, a recent study showed that the level of HMGA2 is highly regulated during adipogenesis by *let-*

7 and knockdown of HMGA2 by this miRNA inhibited both clonal expansion and the transition to terminal differentiation [67]. HMGA2 seems to act in a narrow window early in the adipogenesis where it is involved in expanding the population of preadipocytes and keeps the cells in an undifferentiated state. This would be consistent with our *HMGA2* over-expressing keeping the mesenchymal stem-like cells in a proliferative state, preventing differentiation into adipocytes. Transgenic mice overexpressing truncated HMGA2 gene, on the other hand, develop well differentiated lipomas and abundant adipose tissue [24]. This discrepancy is not surprising, as the conditions during culture *in vitro* and development *in vivo* are very different, among other things due to the microenvironment. Furthermore, the absence in hMSC- TERT20 cells of p16, encoded by the *INK4A/ARF* locus, and possibly other oncogenic alterations in our model system [37] may be expected to contribute to the results, although this blockage of differentiation was confirmed in preliminary experiments using another, non-tumourigenic telomerase-immortalised MSC line (Stabell et al, unpublished).

It has recently been reported that HMGA2 represses transcription of *INK4A/ARF* [68] and suggested that this might be mediated through JUNB, a transactivator of *INK4A/ARF* [69]. Consistent with previous findings [3, 68], both HMGA2 forms down-regulated the expression of NF- κ B regulated genes in the cells, and in addition to binding sites for NF- κ B, many of the down-regulated genes identified here, including *JUNB*, also contain potential intrinsic binding sites for HMGA2 (Additional file 3 Table S3). These findings suggest that HMGA2 may repress JUNB by interfering with the NF- κ B mediated transactivation of this gene. However, since the *INK4A/ARF* locus is deleted in our cell model, we were not able to confirm an effect on this gene by over-expressing HMGA2. One might speculate that the higher number

of genes affected by the truncated protein could be due to a regulatory or moderating function of the c-terminal domain, but this remains to be investigated further.

In spite of abnormal dynamics of growth and differentiation of our model system, we expect the transcription response to HMGA2 overexpression to reflect authentic regulatory functions in an immature mesenchymal context, as these cells show differentiation properties similar to primary mesenchymal stem-like cell cultures.

HMGA2 is known to participate in the formation of heterochromatin in senescent cells [18] and we found that both full-length and truncated HMGA2 proteins co-localized extensively with DAPI-dense regions, probably enriched in heterochromatin (Fig. 1a). The modulation of chromatin by HMGA2 could also be responsible for the coordinated induction of the chromosome 19 KRAB-ZNF genes. Several members of this family modulate cell growth and survival, and they are implicated in malignant disorders [70], although the functions of most of them are largely unknown.

Noteworthy, a recently study showed that Apak, encoded by the *ZNF420* gene, repress p53-mediated apoptosis [71] and it was suggested that various KRAB-type zinc-finger proteins may modulate p53 activity. Although the level of *ZNF420* transcript increased more than two-fold in cells over-expressing HMGA2, no effects on pro-apoptotic genes were observed (data not shown).

Interestingly, our most up-regulated gene, *SSX1*, encodes a protein that also is involved in chromatin modulation, and is already involved in sarcoma oncogenesis [72]. Members of this gene family are seen expressed in germ cells and mesenchymal stem cells, decrease upon differentiation into adipocytes and osteocytes [73, 74], and are expressed in various tumor types [75]. SSX proteins co-localize with Polycomb group (PcG) proteins in the nucleus and likely act through these chromatin regulators

to maintain repression of target genes [76, 77]. This could suggest that they play a key role in maintaining mesenchymal cell stemness [78], and that increased SSX1 expression could reflect a cancer stem cell phenotype. Although the SSX gene family is highly conserved, the variable activation of SSX genes reported in malignant cells [74] suggests gene-specific transcription control. Their restricted expression pattern is mainly governed by epigenetic silencing [79, 80], which could be affected by HMGA2_{TRUNC}. An AT-rich region 1 kb upstream from transcription start of SSX1 contains a likely binding site for HMGA2 (⁻¹⁰⁷⁶ TATTAATAT), although it remains to be proven that it binds specifically to HMGA2_{TRUNC}. Recent findings suggest that both HMGA2 [81] and SSX proteins [73, 82] are taking part in the epithelial mesenchymal transition, where epithelial cells convert to motile mesenchymal cells [83]. However, the significance of these processes in already mesenchymal cells is unclear. Interestingly, we observed a change in both cell lines to a more restricted but still mesenchymal phenotype by the repression of some epithelial markers, such as keratin 14 (*KRT14*) and the P selectin ligand CD24, and up-regulation of mesenchymal-specific genes such as *CD44*, probably shifting from a multi-lineage potential [84] to a more restricted but proliferative state. The down-regulation of the NF-κB pathway might signify an undifferentiated mesenchymal phenotype committed to adipogenic lineage depending on further signalling cues.

The induction of *CXCL12* gene expression by HMGA2_{TRUNC} is probably the cause of down-regulation of HLA-DRA and other major histocompatibility complex class II (MHC-II) proteins on the cell surface, due to this cytokine's ability to repress the class II transactivator [85]. The fact that interferon induced HLA-DRA protein levels to a similar level in parental and HMGA2_{TRUNC} cells supports this interpretation. Loss of MHC-II expression in lymphomas correlates with poor survival [86], and absence

of HLA-DR expression is seen in other cancers [87]. Their lost ability to present tumor antigens on their cell surfaces may contribute to their escape from immunosurveillance and therefore this might also provide a selective advantage to sarcoma cells expressing truncated HMGA2.

Since HMGA2 truncation and amplification is observed initially in the development of well-differentiated liposarcomas, borderline tumours of adipocytic differentiation, and always in combination with amplification of the p53-blocking MDM2, we do not expect a very “malignant” phenotype when HMGA2 is overexpressed on its own. On the other hand, a phenotype where adipogenesis is inhibited would be consistent with a role in adipogenic tumours. When truncated HMGA2 is overexpressed in transgenic mice, they develop adipose hypertrophy, which indicates that adipocytic differentiation is not completely blocked [24]. Like our constructs, theirs lack the 3’ UTR, thus preventing down regulation of HMGA2 by *let-7* as differentiation progresses, and the most likely explanation for the discrepancy is that the balance of proliferation and differentiation over weeks or months *in vivo* is sufficiently different for our short-term *in vitro* conditions, or there are differences in how human and mouse cells are regulated at this level. Although we get the same result in another non-transformed hMSC line, this model is also based on ectopic activation of telomerase, which could affect the result[88, 89].

To investigate these questions, we will determine the activity of truncated HMGA2 in sarcoma cells by siRNA-based knock-down in cells with the natural malignant genetic background.

Conclusions

Our results show that over-expression of truncated HMGA2 induces a more mesenchymal (stem-like) phenotype, characterized by resistance toward differentiation, over-expression of SSX1, lost expression of certain epithelial markers and strengthened expression of mesenchymal markers. We suggest that amplification and truncation of HMGA2 in sarcoma provide the tumors with a more stem-like phenotype.

Competing interests

The authors have no competing interests

Authors' contributions

JH participated in the planning of the experiments, performed all the transfections and experiments for array analysis, performed most of the array data analysis, and writing of the manuscript. MS performed the differentiation analysis and participated in the writing of the manuscript. LAMZ conceived the study, supervised the practical work, and participated in writing of the manuscript. SAUL performed the confocal microscopy analyses and participated in the writing of the manuscript. MK constructed and provided the mesenchymal model system used, and critically read the manuscript. OM conceived the study, supervised the practical work, and participated in the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Marwa Jalal for having performed the FACS, Russell Castro for the analysis of surface markers, and Hanne S. Sorte and Ingrid Østensen for their assistance with the microarray analyses. Microarray analyses were supported by the FUGE Platform for Microarray Technology (Norwegian Microarray Consortium,

NMC, microarray.no), funded by the Norwegian Research Council, and this study was supported by The Norwegian Cancer Society.

References

1. Reeves R, Nissen M: **The AT-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure.** *J Biol Chem* 1990, **265**(15):8573-8582.
2. Nissen MS, Reeves R: **Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y.** *J Biol Chem* 1995, **270**(9):4355-4360.
3. Noro B, Licheri B, Sgarra R, Rustighi A, Tessari MA, Chau KY, Ono SJ, Giancotti V, Manfioletti G: **Molecular dissection of the architectural transcription factor HMGA2.** *Biochemistry* 2003, **42**(15):4569-4577.
4. Chase MB, Haga SB, Hankins WD, Williams DM, Bi Z, Strovel JW, Obriecht C, Berg PE: **Binding of HMG-I(Y) elicits structural changes in a silencer of the human beta-globin gene.** *Am J Hematol* 1999, **60**(1):27-35.
5. Edelstein LC, Lagos L, Simmons M, Tirumalai H, Gelinas C: **NF-kappa B-dependent assembly of an enhanceosome-like complex on the promoter region of apoptosis inhibitor Bfl-1/A1.** *Mol Cell Biol* 2003, **23**(8):2749-2761.
6. Klein-Hessling S, Schneider G, Heinfling A, Chuvpilo S, Serfling E: **HMG I(Y) interferes with the DNA binding of NF-AT factors and the induction of the interleukin 4 promoter in T cells.** *Proc Natl Acad Sci U S A* 1996, **93**(26):15311-15316.
7. Luan J, Shattuck-Brandt R, Haghnegahdar H, Owen JD, Strieter R, Burdick M, Nirodi C, Beauchamp D, Johnson KN, Richmond A: **Mechanism and biological significance of constitutive expression of MGSA/GRO chemokines in malignant melanoma tumor progression.** *J Leukoc Biol* 1997, **62**(5):588-597.
8. Thanos D, Maniatis T: **The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene.** *Cell* 1992, **71**(5):777-789.
9. Chin MT, Pellacani A, Wang H, Lin SS, Jain MK, Perrella MA, Lee ME: **Enhancement of serum-response factor-dependent transcription and DNA binding by the architectural transcription factor HMG-I(Y).** *J Biol Chem* 1998, **273**(16):9755-9760.
10. Fedele M, Pierantoni GM, Visone R, Fusco A: **E2F1 activation is responsible for pituitary adenomas induced by HMGA2 gene overexpression.** *Cell Div* 2006, **1**:17.
11. Tessari MA, Gostissa M, Altamura S, Sgarra R, Rustighi A, Salvagno C, Caretti G, Imbriano C, Mantovani R, Del Sal G *et al*: **Transcriptional activation of the cyclin A gene by the architectural transcription factor HMGA2.** *Mol Cell Biol* 2003, **23**(24):9104-9116.

12. Bonnefoy E, Bandu MT, Doly J: **Specific binding of high-mobility-group I (HMGI) protein and histone H1 to the upstream AT-rich region of the murine beta interferon promoter: HMGI protein acts as a potential antirepressor of the promoter.** *Mol Cell Biol* 1999, **19**(4):2803-2816.
13. Li O, Vasudevan D, Davey CA, Droge P: **High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells.** *Genesis* 2006, **44**(11):523-529.
14. Reeves R, Leonard WJ, Nissen MS: **Binding of HMG-I(Y) imparts architectural specificity to a positioned nucleosome on the promoter of the human interleukin-2 receptor alpha gene.** *Mol Cell Biol* 2000, **20**(13):4666-4679.
15. Reeves R, Nissen MS: **Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles.** *J Biol Chem* 1993, **268**(28):21137-21146.
16. Zhao K, Kas E, Gonzalez E, Laemmli UK: **SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin.** *Embo J* 1993, **12**(8):3237-3247.
17. Funayama R, Saito M, Tanobe H, Ishikawa F: **Loss of linker histone H1 in cellular senescence.** *J Cell Biol* 2006, **175**(6):869-880.
18. Narita M, Narita M, Krizhanovsky V, Nunez S, Chicas A, Hearn SA, Myers MP, Lowe SW: **A novel role for high-mobility group proteins in cellular senescence and heterochromatin formation.** *Cell* 2006, **126**(3):503-514.
19. Chiappetta G AV, Visconti R, Fedele M, Battista S, Trapasso F, Merciai BM, Fidanza V, Giancotti V, Santoro M, Simeone A, Fusco A.: **High level expression of the HMGI (Y) gene during embryonic development.** *Oncogene* 1996, **13**(11):2439-2446.
20. Rogalla P, Drechsler K, Frey G, Hennig Y, Helmke B, Bonk U, Bullerdiek J: **HMGI-C expression patterns in human tissues. Implications for the genesis of frequent mesenchymal tumors.** *Am J Pathol* 1996, **149**(3):775-779.
21. Anand A, Chada K: **In vivo modulation of Hmgic reduces obesity.** *Nat Genet* 2000, **24**(4):377-380.
22. Zhou X, Benson KF, Ashar HR, Chada K: **Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C.** *Nature* 1995, **376**(6543):771-774.
23. Arlotta P, Tai AK, Manfioletti G, Clifford C, Jay G, Ono SJ: **Transgenic mice expressing a truncated form of the high mobility group I-C protein develop adiposity and an abnormally high prevalence of lipomas.** *J Biol Chem* 2000, **275**(19):14394-14400.
24. Battista S, Fidanza V, Fedele M, Klein-Szanto AJ, Outwater E, Brunner H, Santoro M, Croce CM, Fusco A: **The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis.** *Cancer Res* 1999, **59**(19):4793-4797.
25. Ashar HR, Fejzo MS, Tkachenko A, Zhou X, Fletcher JA, Weremowicz S, Morton CC, Chada K: **Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains.** *Cell* 1995, **82**(1):57-65.
26. Kazmierczak B, Wanschura S, Rosigkeit J, Meyer-Bolte K, Uschinsky K, Haupt R, Schoenmakers EF, Bartnitzke S, Van de Ven WJ, Bullerdiek J:

- Molecular characterization of 12q14-15 rearrangements in three pulmonary chondroid hamartomas.** *Cancer Res* 1995, **55**(12):2497-2499.
27. Schoenmakers EF, Wanschura S, Mols R, Bullerdiek J, Van den Berghe H, Van de Ven WJ: **Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours.** *Nat Genet* 1995, **10**(4):436-444.
 28. Berner J, Meza-Zepeda L, Kools, Forus A, Schoenmakers E, Ven Wvd, Fodstad O, Myklebost O: **HMGIC, the gene for an architectural transcription factor, is amplified and rearranged in a subset of human sarcomas.** *Oncogene* 1997, **14**(24):2935-2941.
 29. Meza-Zepeda LA, Berner JM, Henriksen J, South AP, Pedeutour F, Dahlberg AB, Godager LH, Nizetic D, Forus A, Myklebost O: **Ectopic sequences from truncated HMGIC in liposarcomas are derived from various amplified chromosomal regions.** *Genes Chromosomes Cancer* 2001, **31**(3):264-273.
 30. Lee YS, Dutta A: **The tumor suppressor microRNA let-7 represses the HMGA2 oncogene.** *Genes Dev* 2007, **21**(9):1025-1030.
 31. Mayr C, Hemann MT, Bartel DP: **Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation.** *Science* 2007, **315**(5818):1576-1579.
 32. Langelotz C, Schmid P, Jakob C, Heider U, Wernecke KD, Possinger K, Sezer O: **Expression of high-mobility-group-protein HMGI-C mRNA in the peripheral blood is an independent poor prognostic indicator for survival in metastatic breast cancer.** *Br J Cancer* 2003, **88**(9):1406-1410.
 33. Motoyama K, Inoue H, Nakamura Y, Uetake H, Sugihara K, Mori M: **Clinical significance of high mobility group A2 in human gastric cancer and its relationship to let-7 microRNA family.** *Clin Cancer Res* 2008, **14**(8):2334-2340.
 34. Shell S, Park SM, Radjabi AR, Schickel R, Kistner EO, Jewell DA, Feig C, Lengyel E, Peter ME: **Let-7 expression defines two differentiation stages of cancer.** *Proc Natl Acad Sci U S A* 2007, **104**(27):11400-11405.
 35. Yu F, Yao H, Zhu P, Zhang X, Pan Q, Gong C, Huang Y, Hu X, Su F, Lieberman J *et al*: **let-7 regulates self renewal and tumorigenicity of breast cancer cells.** *Cell* 2007, **131**(6):1109-1123.
 36. Simonsen J, Rosada C, Serakinci N, Justesen J, Stenderup K, Rattan S, Jensen T, Kassem M: **Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells.** *Nature Biotechnology* 2002, **20**(6):592-596.
 37. Serakinci N, Guldborg P, Burns JS, Abdallah B, Schrodder H, Jensen T, Kassem M: **Adult human mesenchymal stem cell as a target for neoplastic transformation.** *Oncogene* 2004, **23**(29):5095-5098.
 38. De Angelis P, Stokke T, Clausen O: **NO38 expression and nucleolar counts are correlated with cellular DNA content but not with proliferation parameters in colorectal carcinomas.** *Mol Pathol* 1997, **50**(4):201-208.
 39. Skehan P, Storeng R, Scudiero D, Monks A, McMahon J, Vistica D, Warren JT, Bokesch H, Kenney S, Boyd MR: **New colorimetric cytotoxicity assay for anticancer-drug screening.** *J Natl Cancer Inst* 1990, **82**(13):1107-1112.
 40. Forus A, Flørenes V, Maelandsmo G, Meltzer P, Fodstad O, Myklebost O: **Mapping of amplification units in the q13-14 region of chromosome 12 in human sarcomas: some amplica do not include MDM2.** *Cell Growth Differ* 1993, **4**(12):1065-1070.

41. Schmittgen T, Livak K: **Analyzing real-time PCR data by the comparative C(T) method.** *Nat Protoc* 2008, **3**(6):1101-1108.
42. Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S (eds.): **Bioinformatics and Computational Biology Solutions Using R and Bioconductor:** Springer; 2005.
43. Huang dW, Sherman B, Lempicki R: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nature Protocols* 2009, **4**(1):44-57.
44. Cleynen I, Brants JR, Peeters K, Deckers R, Debiec-Rychter M, Sciort R, Van de Ven WJ, Petit MM: **HMG2A2 regulates transcription of the Imp2 gene via an intronic regulatory element in cooperation with nuclear factor-kappaB.** *Mol Cancer Res* 2007, **5**(4):363-372.
45. Taganov KD, Boldin MP, Chang KJ, Baltimore D: **NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses.** *Proc Natl Acad Sci U S A* 2006, **103**(33):12481-12486.
46. Kuncio GS, Tsyganskaya M, Zhu J, Liu SL, Nagy L, Thomazy V, Davies PJ, Zern MA: **TNF-alpha modulates expression of the tissue transglutaminase gene in liver cells.** *Am J Physiol* 1998, **274**(2 Pt 1):G240-245.
47. Woods M, Wood EG, Bardswell SC, Bishop-Bailey D, Barker S, Wort SJ, Mitchell JA, Warner TD: **Role for nuclear factor-kappaB and signal transducer and activator of transcription 1/interferon regulatory factor-1 in cytokine-induced endothelin-1 release in human vascular smooth muscle cells.** *Mol Pharmacol* 2003, **64**(4):923-931.
48. You X, Wu Y, Zeng Y, Deng Z, Qiu H, Yu M: **Mycoplasma genitalium-derived lipid-associated membrane proteins induce activation of MAPKs, NF-kappaB and AP-1 in THP-1 cells.** *FEMS Immunol Med Microbiol* 2008, **52**(2):228-236.
49. Qureshi S, Salditt-Georgieff M, Darnell JJ: **Tyrosine-phosphorylated Stat1 and Stat2 plus a 48-kDa protein all contact DNA in forming interferon-stimulated-gene factor 3.** *Proc Natl Acad Sci U S A* 1995, **92**(9):3829-3833.
50. Weihua X, Hu J, Roy SK, Mannino SB, Kalvakolanu DV: **Interleukin-6 modulates interferon-regulated gene expression by inducing the ISGF3 gamma gene using CCAAT/enhancer binding protein-beta(C/EBP-beta).** *Biochim Biophys Acta* 2000, **1492**(1):163-171.
51. Tadepally HD, Burger G, Aubry M: **Evolution of C2H2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains.** *BMC Evol Biol* 2008, **8**:176.
52. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA *et al*: **The sequence of the human genome.** *Science* 2001, **291**(5507):1304-1351.
53. Margolin JF, Friedman JR, Meyer WK, Vissing H, Thiesen HJ, Rauscher FJ, 3rd: **Kruppel-associated boxes are potent transcriptional repression domains.** *Proc Natl Acad Sci U S A* 1994, **91**(10):4509-4513.
54. Takahashi M, Kamei Y, Ezaki O: **Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size.** *Am J Physiol Endocrinol Metab* 2005, **288**(1):E117-124.
55. Zandbergen F, Mandard S, Escher P, Tan N, Patsouris D, Jatkoa T, Rojas-Caro S, Madore S, Wahli W, Tafuri S *et al*: **The G0/G1 switch gene 2 is a**

- novel PPAR target gene.** *Biochem J* 2005 Dec 1;392(Pt 2):313-24 2005, **392**(Pt 2):313-324.
56. Kamei Y ST, Kohda T, Ishino F, Yasuda K, Miura S, Ezaki O, Ogawa Y.: **Peg1/Mest in obese adipose tissue is expressed from the paternal allele in an isoform-specific manner.** *FEBS Lett* 2007, **581**(1):91-96.
 57. Kawano M, Suzuki S, Suzuki M, Oki J, Imamura T: **Bulge- and basal layer-specific expression of fibroblast growth factor-13 (FHF-2) in mouse skin.** *J Invest Dermatol* 2004, **122**(5):1084-1080.
 58. Fujikawa M, Katagiri T, Tugores A, Nakamura Y, Ishikawa F: **ESE-3, an Ets family transcription factor, is up-regulated in cellular senescence.** *Cancer Sci* 2007, **98**(9):1468-1475.
 59. Gomez T, McCarney S, Carrizosa E, Labno CM CE, Nolz JC, Zhu P, Freedman BD, Clark MR, Rawlings DJ, Billadeau DD, Burkhardt JK.: **HS1 functions as an essential actin-regulatory adaptor protein at the immune synapse.** *Immunity* 2006, **24**(6):741-752.
 60. Doi A, Shono T, Nishi M, Furuta H, Sasaki H, Nanjo K: **IA-2beta, but not IA-2, is induced by ghrelin and inhibits glucose-stimulated insulin secretion.** *Proc Natl Acad Sci U S A* 2006, **103**(4):885-890.
 61. Sgarra R, Tessari M, Di Bernardo J, Rustighi A, Zago P, Liberatori S, Armini A, Bini L, Giancotti V, Manfioletti G: **Discovering high mobility group A molecular partners in tumour cells.** *Proteomics* 2005, **5**(6):1494-1506.
 62. Spector D: **Macromolecular domains within the cell nucleus.** In: *Annu Rev Cell Biol.* vol. 9; 1993: 265-315.
 63. Szebeni A, Olson MO: **Nucleolar protein B23 has molecular chaperone activities.** *Protein Sci* 8, 905-912 1999, **8**(4):905-912.
 64. Kondo T, Minamino N, Nagamura-Inoue T, Matsumoto M, Taniguchi T, Tanaka N: **Identification and characterization of nucleophosmin/B23/numatrin which binds the anti-oncogenic transcription factor IRF-1 and manifests oncogenic activity.** *Oncogene* 1997, **15**(11):1275-1281.
 65. Takemura M, Ohoka F, Perpelescu M, Ogawa M, H. M, Takaba T, Akiyama T, Umekawa H, Furuichi Y, Cook P *et al*: **Phosphorylation-dependent migration of retinoblastoma protein into the nucleolus triggered by binding to nucleophosmin/B23.** *Exp Cell Res* 2002, **276**(2):233-241.
 66. Zhang S, Hemmerich P, Grosse F: **Nucleolar localization of the human telomeric repeat binding factor 2 (TRF2).** *J Cell Sci* 2004, **117**(Pt17):3935-3945.
 67. Sun T, Fu M, Bookout AL, Kliwer SA, Mangelsdorf DJ: **MicroRNA let-7 regulates 3T3-L1 adipogenesis.** *Molecular endocrinology (Baltimore, Md)* 2009, **23**(6):925-931.
 68. Nishino J, Kim I, Chada K, Morrison SJ: **Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression.** *Cell* 2008, **135**(2):227-239.
 69. Passegué E, Wagner E: **JunB suppresses cell proliferation by transcriptional activation of p16(INK4a) expression.** *EMBO J* 2000, **19**(12):2969-2979.
 70. Tommerup N, Vissing H: **Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identify putative candidate genes for developmental and malignant disorders.** *Genomics* 1995, **27**(2):259-264.

71. Tian C, Xing G, Xie P, Lu K, Nie J, Wang J, Li L, Gao M, Zhang L, He F: **KRAB-type zinc-finger protein Apak specifically regulates p53-dependent apoptosis.** *Nat Cell Biol* 2009, **11**(5):580-591.
72. Nagai M, Tanaka S, Tsuda M, Endo S, Kato H, Sonobe H, Minami A, Hiraga H, Nishihara H, Sawa H *et al*: **Analysis of transforming activity of human synovial sarcoma-associated chimeric protein SYT-SSX1 bound to chromatin remodeling factor hBRM/hSNF2 alpha.** *Proc Natl Acad Sci U S A* 2001, **98**(7):3843-3848.
73. Cronwright G, Le Blanc K, Gotherstrom C, Darcy P, Ehnman M, Brodin B: **Cancer/testis antigen expression in human mesenchymal stem cells: down-regulation of SSX impairs cell migration and matrix metalloproteinase 2 expression.** *Cancer Res* 2005, **65**(6):2207-2215.
74. Gure AO, Tureci O, Sahin U, Tsang S, Scanlan MJ, Jager E, Knuth A, Pfreundschuh M, Old LJ, Chen YT: **SSX: a multigene family with several members transcribed in normal testis and human cancer.** *International journal of cancer* 1997, **72**(6):965-971.
75. Tureci O, Chen YT, Sahin U, Gure AO, Zwick C, Villena C, Tsang S, Seitz G, Old LJ, Pfreundschuh M: **Expression of SSX genes in human tumors.** *Int J Cancer* 1998, **77**(1):19-23.
76. De Bruijn DR, van Dijk AH, Willemsse MP, van Kessel AG: **The C terminus of the synovial sarcoma-associated SSX proteins interacts with the LIM homeobox protein LHX4.** *Oncogene* 2008, **27**(5):653-662.
77. Soulez M, Saurin AJ, Freemont PS, Knight JC: **SSX and the synovial-sarcoma-specific chimaeric protein SYT-SSX co-localize with the human Polycomb group complex.** *Oncogene* 1999, **18**(17):2739-2746.
78. Bracken AP DN, Pasini D, Hansen KH, Helin K.: **Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions.** *Genes Dev* 2006, **20**(9):1123-1136.
79. dos Santos N, de Bruijn D, van Kessel A: **Molecular mechanisms underlying human synovial sarcoma development.** *Genes Chromosomes Cancer* 2001, **30**(1):1-14.
80. de Bruijn D, dos Santos N, Kater-Baats E, Thijssen J, van den Berk L, Stap J, Balemans M, Schepens M, Merckx G, van Kessel A: **The cancer-related protein SSX2 interacts with the human homologue of a Ras-like GTPase interactor, RAB3IP, and a novel nuclear protein, SSX2IP.** *Genes Chromosomes Cancer* 2002, **34**(3):285-298.
81. Thuault S, Valcourt U, Petersen M, Manfioletti G, Heldin C, A M: **Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition.** *J Cell Biol* 2006, **174**(2):175-183.
82. Saito T, Nagai M, Ladanyi M: **SYT-SSX1 and SYT-SSX2 interfere with repression of E-cadherin by snail and slug: a potential mechanism for aberrant mesenchymal to epithelial transition in human synovial sarcoma.** *Cancer Res* 2006, **66**(14):6919-6927.
83. Hay E: **The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it.** *Dev Dyn* 2005, **233**(3):706-720.
84. Abdallah B, Haack-Sørensen M, Burns J, Elsnab B, Jakob F, Hokland P, Kassem M: **Maintenance of differentiation potential of human bone marrow mesenchymal stem cells immortalized by human telomerase reverse transcriptase gene despite [corrected] extensive proliferation.** *Biochem Biophys Res Commun* 2005, **326**(3):527-538.

85. Sheridan C, Sadaria M, Bhat-Nakshatri P, Goulet R, Jr., Edenberg HJ, McCarthy BP, Chang CH, Srouf EF, Nakshatri H: **Negative regulation of MHC class II gene expression by CXCR4.** *Exp Hematol* 2006, **34**(8):1085-1092.
86. Rimsza LM, Roberts RA, Miller TP, Unger JM, LeBlanc M, Brazier RM, Weisenberger DD, Chan WC, Muller-Hermelink HK, Jaffe ES *et al*: **Loss of MHC class II gene and protein expression in diffuse large B-cell lymphoma is related to decreased tumor immunosurveillance and poor patient survival regardless of other prognostic factors: a follow-up study from the Leukemia and Lymphoma Molecular Profiling Project.** *Blood* 2004, **103**(11):4251-4258.
87. Satoh A, Toyota M, Ikeda H, Morimoto Y, Akino K, Mita H, Suzuki H, Sasaki Y, Kanaseki T, Takamura Y *et al*: **Epigenetic inactivation of class II transactivator (CIITA) is associated with the absence of interferon-gamma-induced HLA-DR expression in colorectal and gastric cancer cells.** *Oncogene* 2004, **23**(55):8876-8886.
88. Armstrong L, Saretzki G, Peters H, Wappler I, Evans J, Hole N, von Zglinicki T, Lako M: **Overexpression of telomerase confers growth advantage, stress resistance, and enhanced differentiation of ESCs toward the hematopoietic lineage.** *Stem Cells* 2005, **23**(4):516-529.
89. Yang C, Przyborski S, Cooke MJ, Zhang X, Stewart R, Anyfantis G, Atkinson SP, Saretzki G, Armstrong L, Lako M: **A key role for telomerase reverse transcriptase unit in modulating human embryonic stem cell proliferation, cell cycle dynamics, and in vitro differentiation.** *Stem Cells* 2008, **26**(4):850-863.

Table 1. Identification of genes most differentially expressed between the parental cell line and the over-expressing cell lines.

Changes in HMGA2 _{WT} cells							Changes in HMGA2 _{trunc} cells				
Rank	Gene	Description	MA	qPCR	Rel. to EGFP*	Gene	Description	MA	qPCR	Rel. to EGFP*	
a)	1	<i>CXCL6</i>	Chemokine (C-X-C motif) ligand 6	-8.4	-6.7	-2.2	<i>CXCL6</i>	Chemokine (C-X-C motif) ligand 6	-9.0	-6.3	-2.2
	2	<i>IL8</i>	Interleukin 8	-7.5	-7.1	-6.4	<i>HLA-DRA</i>	Major histocompatibility complex class II antigen DRA	-8.7	-8.1	-7.2
	3	<i>CD24</i>	CD24 molecule	-7.1			<i>HLA-DPA1</i>	Major histocompatibility complex class II antigen DPA1	-8.0	-6.8	-5.5
	4	<i>C13orf33</i>	Chromosome 13 open reading frame 33	-6.5			<i>ChGn</i>	Chondroitin beta-1,4-N-acetylgalactosaminyltransferase	-6.5		
	5	<i>PTGES</i>	Prostaglandin E synthase	-6.3			<i>PKIB</i>	cAMP-dependent protein kinase β 2	-6.4	-4.9	-5.0
	6	<i>ChGn</i>	Chondroitin beta-1,4-N-acetylgalactosaminyltransferase	-5.6			<i>IFI44L</i>	Interferon-induced protein 44-like	-6.2		
	7	<i>IL6</i>	Interleukin 6	-5.5			<i>LAMA4</i>	Laminin alpha chain	-6.2		
	8	<i>PDE4D</i>	cAMP-specific phosphodiesterase 4D	-5.2			<i>SLC38A5</i>	Solute carrier family 38, member 5	-6.0		
	9	<i>IL1RN</i>	Interleukin 1 receptor antagonist	-5.1	-4.4	-3.0	<i>PTPRN2</i>	Protein tyrosine phosphatase, receptor type, N2	-5.5		
	10	<i>CES1</i>	Carboxylesterase 1	-5.1			<i>C13orf33</i>	Chromosome 13 open reading frame 33	-5.4		
	11	<i>KIAA1644</i>	KIAA1644 protein	-5.0	-3.4	-1.7	n/a	Similar to Wnt-1 inducible signalling pathway protein 1	-5.4		
	12	<i>SAMSN1</i>	SH3-SAM adaptor protein	-4.8			<i>IL1B</i>	Interleukin1 beta	-5.3	-4.3	-2.2
	13	<i>SERPIND1</i>	Serpin peptidase inhibitor, clade D, member 1	-4.7			<i>F2RL1</i>	Coagulation factor II receptor-like 1	-5.2		
	14	<i>FAM5C</i>	Family with sequence similarity 5, member C	-4.4			<i>CD74</i>	CD74 antigen-associated invariant chain	-5.2	-5.4	-4.3
	15	n/a	Transcribed sequence FLJ 26764	-4.3			<i>IL1RN</i>	Interleukin 1 receptor antagonist	-5.2	-5.2	-3.8
b)	1	<i>ZBED2</i>	Zinc finger BED domain-containing protein 2	5.3	4.5		n/a	transcribed sequence FLJ33010	6.7		
	2	<i>SHROOM2</i>	Shroom family member 2	4.6			<i>CXCL12</i>	Chemokine (C-X-C motif) ligand 12	6.4	2.7	2.0
	3	<i>FGF13</i>	Fibroblast growth factor 13	4.5	4.6	5.6	<i>PRSS7</i>	Serine protease 7	6.0		
	4	<i>BAI3</i>	Brain-specific angiogenesis inhibitor 3	4.5	4.4	5.4	<i>SULT1B1</i>	Sulfotransferase 1B1	5.8		
	5	<i>PLXNA4</i>	Plexin A4	4.4			<i>SSX1</i>	Synovial sarcoma, X breakpoint 1	4.9	9.6	10.2
	6	<i>LCP1</i>	Lymphocyte cytosolic protein 1	4.1	3.9	2.1	<i>MMP3</i>	Matrix metalloproteinase-3	4.7		
	7	<i>C20orf197</i>	Chromosome 20 open reading frame 197	4.0			<i>CYB5R2</i>	Cytochrome b5 reductase 2	4.5		
	8	<i>G0S2</i>	G0/S switch 2	3.7	5.6	5.4	<i>KISS1</i>	KISS-1 metastasis-suppressor	4.5		
	9	<i>HDAC9</i>	Histone deacetylase 9	3.7	1.5		<i>EDN1</i>	Endothelin 1	4.5		
	10	<i>SIGLEC15</i>	Sialic acid binding Ig-like lectin 15	3.7			<i>SYNE1</i>	Synaptic nuclear envelope protein 1	4.4		
	11	<i>HCLS1</i>	Hematopoietic cell-specific Lyn substrate 1	3.6	2.5	3.6	<i>HAPLN1</i>	Hyaluronan and proteoglycan link protein 1	4.4		
	12	<i>LOC646340</i>	similar to Ankyrin repeat domain-containing protein 18A	3.6			n/a	transcribed sequence FLJ35091	4.4		
	13	<i>GATA3</i>	GATA binding protein 3	3.5	3.6	>10	n/a	Similar to TATA box-binding protein-associated factor 1B	4.3		
	14	<i>SH3BGRL2</i>	SH3 domain binding glutamic acid-rich protein like 2	3.5			<i>BAI3</i>	Brain-specific angiogenesis inhibitor 3	4.2	4.6	5.6
	15	<i>PDE4B</i>	cAMP-specific phosphodiesterase 4B	3.5			n/a	Similar to cytochrome c oxidase I	4.1		

The 15 most down- (a) and up- (b) regulated genes were ranked according to the expression fold change between transfectants and parental cells based on the microarray data. Changes in gene expression were validated by real time PCR using TaqMan assays, values in brackets are relative to those for the

eGFP controls, for comparison. The expression level of each target gene was normalized for TBP expression and represented as fold induction over parental cells. Fold change ratios from both microarray (MA) and real-time assay (qPCR), including qPCR values (marked by an asterisk) calculated relative to those from eGFP only control cells, are shown.

Table 2. Functional annotation analysis of differentially expressed genes.

Biological processes affected by over-expression of wild type and truncated HMGA2	
Vascular development (4 of 66 up-regulated genes)	p=0.039
Response to wounding (17 of 139 down-regulated genes)	p<0.001
Signal transduction (40/139)	p<0.001
Chemotaxis (11/139)	p<0.001
Cytosolic calcium ion homeostasis (5/139)	p<0.001
Phosphate transport (5/139)	p=0.0017
Regulation of cell proliferation (11/139)	p=0.0033
Cell migration (8/139)	p=0.0057
Cell recognition (3/139)	p=0.027
Epithelial cell differentiation (3/139)	p=0.026
Wound healing (4/139)	p=0.042
Biological processes affected by over-expression of wild type HMGA2	
Multi-cellular organismal development (20 of 84 up-regulated genes)	p=0.0011
Response to hormone stimulus (4/84)	p=0.0034
Regulation of signal transduction (7/84)	p=0.022
Pyridine nucleotide biosynthetic process (3 of 99 down-regulated genes)	p=0.0017
Steroid hormone receptor activity (3/99)	p=0.0032
Negative regulation of transcription factor activity (3/99)	p=0.0034
Positive regulation of transcription (7/99)	p=0.0065
Response to external stimulus (8/99)	p=0.042
Biological processes affected by over-expression of truncated HMGA2	
Regulation of transcription (61 of 258 up-regulated genes)	p<0.001
Intracellular protein transport across a membrane (6/258)	p=0.0032
Membrane organization and biogenesis (11/258)	p=0.011
Ras protein signaling transduction (9/258)	p=0.031
Regulation of cytoskeleton organization and biogenesis (4/258)	p=0.042
Antigen processing and presentation (10 of 239 down-regulated genes)	p<0.001
Multi-cellular organismal development (41/239)	p=0.0021
Cell adhesion (16/239)	p=0.019
Eye morphogenesis (3/239)	p=0.035

The genes were classified into biological processes using gene ontology terms. Each biological process was determined if it was significantly over- or underrepresented ($p < 0.05$) by comparing the observed frequency of genes to the expected frequency of genes on the Affymetrix GeneChip that cover 50 547 known genes. Significantly over- or under-represented processes were grouped and ranked according to the geometric mean of their p-values. A representative annotation term for each cluster, the number of genes associated with the term and its p-value was listed.

Figure legend

Fig. 1. Characterization of hMSC-TERT20 cells stably over-expressing eGFP-HMGA2_{WT} and eGFP-HMGA2_{TRUNC}.

a) Nuclear localization of eGFP-tagged HMGA2 proteins was observed in hMSC-TERT20 over-expressing eGFP-HMGA2_{WT} or eGFP-HMGA2_{TRUNC}. Endogenous HMGA2 was visualized by immunofluorescent staining with an anti-HMGA2 antibody (The intensity in parental cells was adjusted to make observation of nucleolar fluorescence possible). The nucleolus was detected with an anti-B23 antibody, while heterochromatin was stained by DAPI, showing co-localization of HMGA2 with both nucleoli and heterochromatin. Arrowheads show concentration of endogenous or eGFP-tagged HMGA2 at sites that correspond with nucleoli, while an arrow indicates discrete foci of HMGA2 within nucleoli. b) Expression levels of endogenous and exogenous *HMGA2* transcripts; expected size of the endogenous transcript is 4468 bp, while the exogenous is only 327bp due to the lack of the 3' UTR. Top panel, northern blot of *HMGA2* mRNAs. Lower panel, quantitation of total *HMGA2* mRNA levels by real-time PCR, normalized for TBP expression and represented as fold induction over parental cells. c) Expression levels of endogenous and exogenous HMGA2 proteins, as detected on western blot. d) Accumulation of fat in cell cultures grown in basal medium supplemented with MDI and Rosiglitazone for 14 days. e) Relative accumulation of fat-bound Oil Red O after 14 days of differentiation of hMSC-TERT20 cells. Data from two independent experiments are shown. Par, Parental cells; eGFP, eGFP-transduced; WT, transduced with eGFP-HMGA2_{WT}; Trunc, transduced with

eGFP-HMGA2_{TRUNC} f) Expression of CD24 and HLA-DRA on each hMSC TERT20-derived cell lines as measured by flow cytometric mean fluorescence intensity (MFI).

Additional files

Additional Table S1 TaqMan primers for real-time PCR

TableSup11.doc

Additional Table S2 The complete list of transcripts changed more than 3-fold

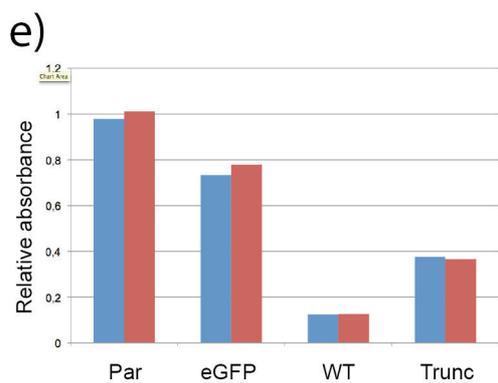
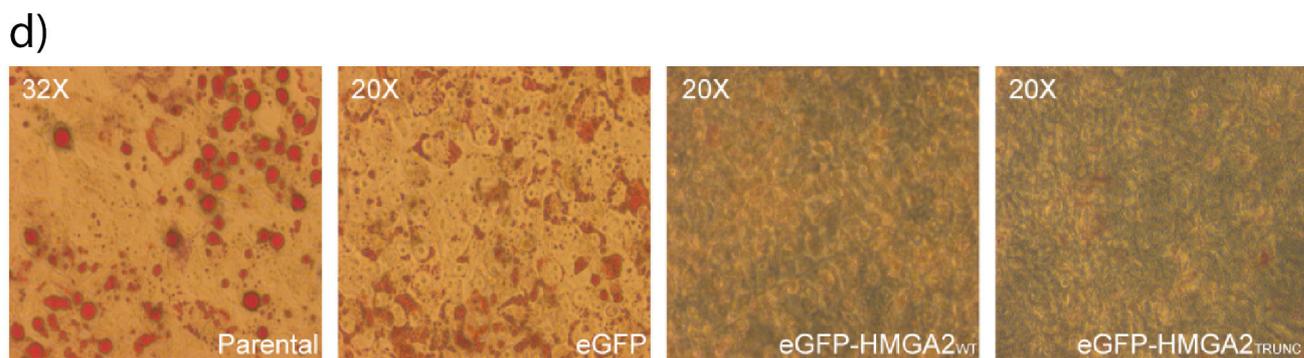
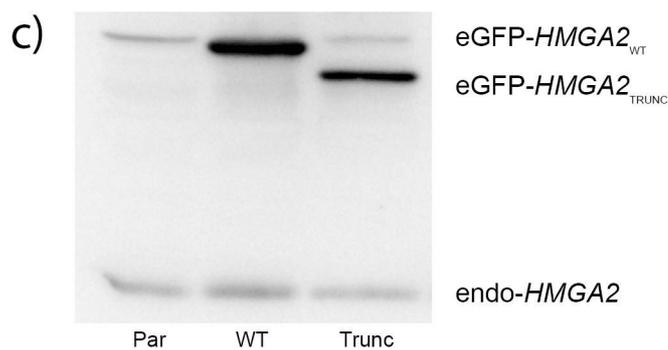
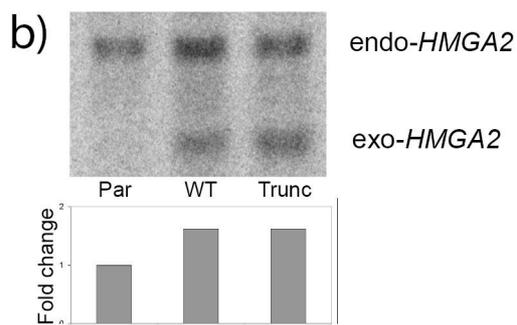
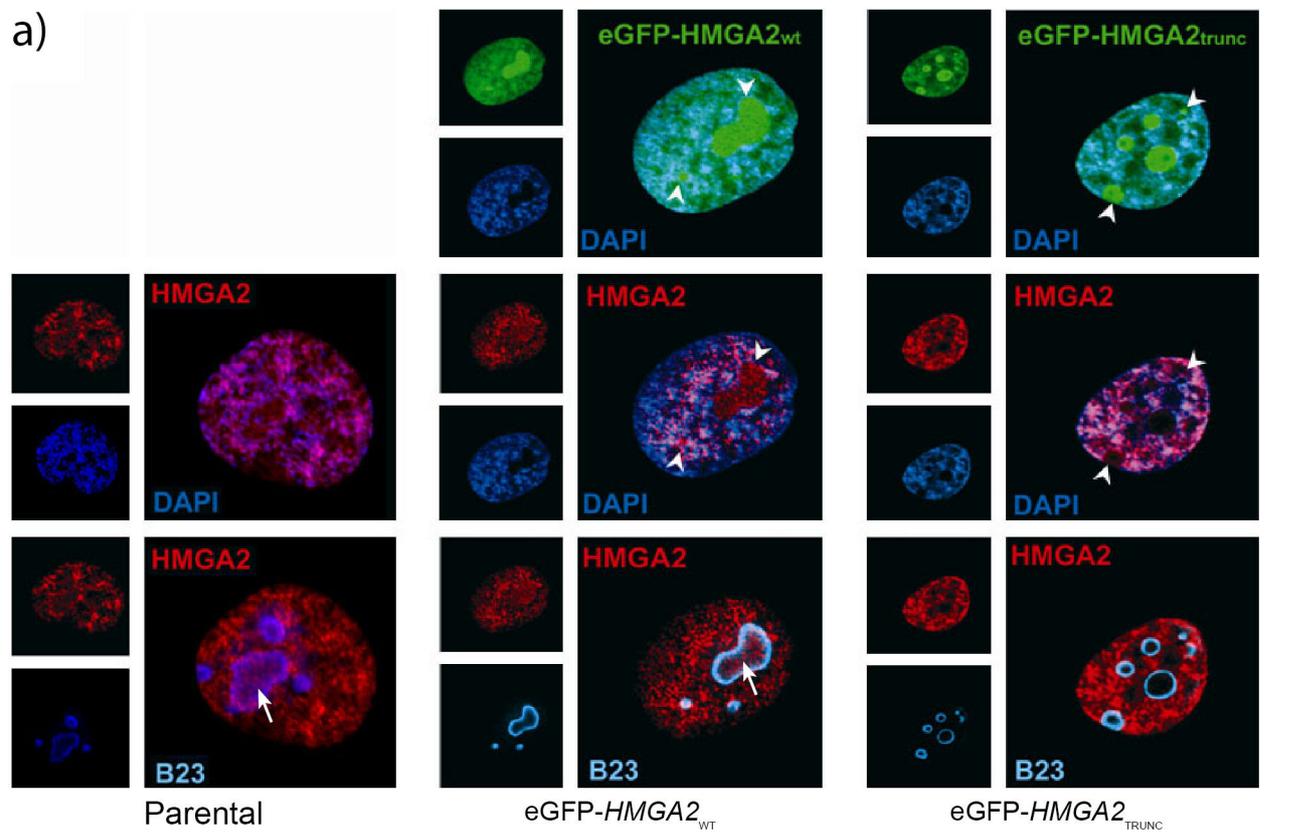
TableSup12.doc

Additional Table S3 Predicted NF- κ B recognition sequences in affected genes

TableSup13.doc

Additional Table S4 Predicted STAT2 recognitions sequences in affected genes

TableSup14.doc



f)

Protein	Parental	WT	Trunc
CD24	21	5.3	5.6
HLA-DRA	54	74	5.6

Additional files provided with this submission:

Additional file 1: TableSupl1.doc, 29K

<http://www.biomedcentral.com/imedia/1393581236345905/supp1.doc>

Additional file 2: TableSupl2.pdf, 589K

<http://www.biomedcentral.com/imedia/1652699770376230/supp2.pdf>

Additional file 3: TableSupl3.doc, 41K

<http://www.biomedcentral.com/imedia/1090339562376231/supp3.doc>

Additional file 4: TableSupl4.doc, 28K

<http://www.biomedcentral.com/imedia/1050013341376232/supp4.doc>