Functional roles of HMGA2 in human embryonic stem cells and human tumor cells

Li, Ou

2009

Li, O. (2009). Functional roles of HMGA2 in human embryonic stem cells and human tumor cells. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/19319

https://doi.org/10.32657/10356/19319

FUNCTIONAL ROLES OF HMGA2 IN HUMAN EMBRYONIC STEM CELLS AND HUMAN TUMOR CELLS

LI OU

School of Biological Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2009

ACKNOWLEGEMENT

First I would like to thank my supervisor, A/P Peter Dröge for his valuable guidance, support and encouragement; and for his willingness to teach and share his knowledge and experience. His tremendous drive, intense results-oriented focus and excellent scientific stewardship have been a great source of inspiration throughout the three and half years, and will be far into the future.

I would like to also thank Dr. Curt A Davey and Dr. Li Jinming for their excellent collaboration and advice in the project.

Many thanks also go to my fellow lab colleagues, Bao Qiuye, Tan Shen Mynn, Klaus Neef, Renji Reghunathan, Heike Summer, Li Heng, Zhou Ruijie, Feng Shu and Sabrina Peter for all of their aid and advice.

Last but not least, I would like to thank my family who has supported me all these years, and especially my wife, Bai Jing, who has always been the inspiration to me!

TABLE OF CONTENTS

LIST OF TABLE

LIST OF FIGURE

ABSTRACT

The state of chromatin in human embryonic stem cells (hESCs) is a key factor determining stem cell identity. The non-histone chromatin-associated factor high-mobility group protein A2 (HMGA2) has been studied mostly in the mouse where its function seems critical for embryonic cell growth and adipocytic cell differentiation, leading to a pygmy phenotype with greatly reduced fat tissue in homozygous knock out mice.

We showed that among the major HMG proteins, HMGA2 was highly expressed in two hESC lines. Interestingly, expression was further upregulated during early embryoid body formation before it quickly droped to or below the level found in undifferentiated cells. We also showed that HMGA2 was stably associated with inter- and meta-phase hESC chromatin, and that up to 12 HMGA2 protomers stably associated *in vitro* with a single nucleosome core particle of known atomic structure. These suggested that HMGA2 might interact with nucleosomes in a way that imposed a global effect on the state of hESC chromatin, which might contribute to the establishment of both hESCs identity and the initiation of specific differentiation programs. We then established HMGA2 as a regulator of human genes linked to mesodermal cell differentiation, adipogenesis, and hESCs growth through siRNA technology in combination with quantitative reverse transcriptase polymerase chain reaction, stem cell-specific microarray analyses, and cell proliferation assays.

HMGA proteins are also present in high copy numbers in most neoplasias, and correlations between the degree of malignancy, patient prognostic index, and HMGA levels have been firmly established. Intriguingly, HMGA2 is also found in rare tumor-

inducing cells, which are resistant to chemotherapy. Here, we demonstrated that HMGA1a/b and HMGA2 possessed intrinsic dRP and AP lyase activities, which most likely resided in their AT-hook DNA binding domains. We also showed that HMGA2 could be covalently trapped at genomic abasic sites in cancer cells. The associated lyase activities promoted cellular resistance against DNA damage, which was targeted by base excision repair (BER) pathways, and this protective effect directly correlated with the level of HMGA2 expression. Furthermore, we demonstrated an interaction between AP endonuclease 1 and HMGA2 in cancer cells, which strongly supported our conclusion that HMGA2 could be incorporated into the cellular BER machinery. This study thus identified an unexpected crucial role for HMGA2 in DNA repair in cancer cells. Taking together the information that HMGA2 seems to be a key signature of Oct3/4 positive cancer stem cell of at least some types of cancer, these findings have important implications for cancer therapy as well as stem cell biology, including genome maintanence and stability, developmental regulation, etc.

In all, we have identified functional roles of HMGA2 in both hESCs and human tumor cells, and provided better understandings on both stem cell and cancer research.

I INTRODUCTION

I.1 Human Embryonic Stem Cells

I.1.1 Overview

Human embryonic stem cells, or hESCs, are derived from the inner cell mass of the blastocyst-stage preimplantation embryo. The term 'Embryonic Stem Cell' was given by Gail R. Martin who first derived these cells from mouse embryos in 1981, although Martin Evans and Matthew Kaufman did the same work independently at the same time. Only until November 1998, a breakthrough in human embryonic stem cell research came in when a group led by James Thomson at the University of Wisconsin-Madison first developed a technique to isolate and grow the cells derived from human blastocysts (Figure I.1). These cells are immortal and retain the potential to differentiate into cells of all three germ layers, and are thus pluripotent in culture over many cell generations (Thomson et al., 1998; Reubinoff et al., 2000). This property allows research into many aspects of human development, for instance, molecular mechanisms of differentiation, organogenesis and even fetal maturation. Eventually, it is hoped that hESCs can live up to clinical promises of cellular transplantation therapy for many degenerative disorders and diseases.

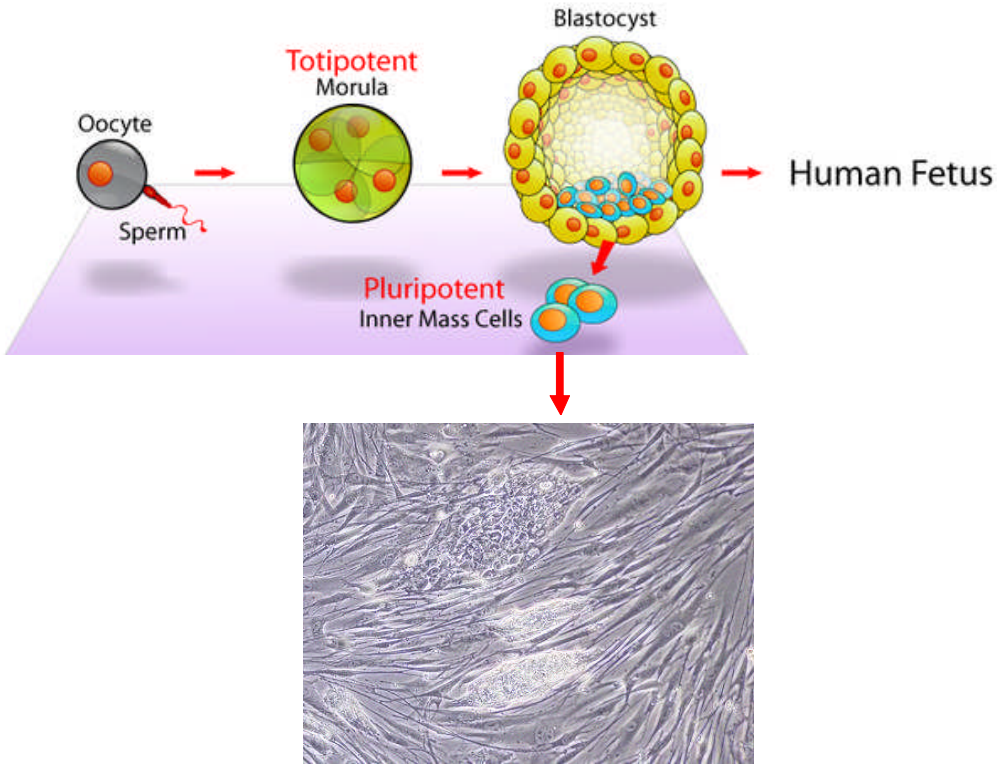


Figure I.1 Derivation of human embryonic ES cell. (Adapted from Wikipedia)

ES cells were derived from the Inner Mass Cells from blastocyst which was developed from either in vitro fertilization or doner oocyte. Briefly, embryos were first cultured to the blastocyst stage in specialized medium and the zona pelliucida was digested with pronase. To remove the trophoblastic cells, the difference in the permeability of inner cell masses (ICMs) and trophoblastic cells to certain antibodies was exploited: these cells are killed when the blastocysts are exposed to antiserum alone and then transferred to guinea pig complement (Solter and Knowles, 1975). Cells were then cultured on the Mouse embryonic fibroblast which functions as feeders.

I.1.2 Properties

The properties and general features of hESCs are essentially similar to those of mouse ESCs (mESCs). It is vital that they possess all of the following characteristics for being considered embryonic stem cells (Pera et al., 2000):

- Originate from a pluripotent cell population
- Maintain a normal karyotype and be stably diploid in vitro
- Immortal and can be cultured indefinitely in the primitive embryonic state
- Ability to differentiate spontaneously into cell types of all three germ layers either:
    - In teratomas after grafting in immune-deficient mice
    - In suspended embryoid body cultures in vitro
    - Under appropriate culture conditions in vitro
- Ability to contribute to any cell type when allowed to colonize a host blastocyst

These criteria can be easily and routinely tested during hESC characterization, except the last one, due to ethical opposition against human chimerical experiments.

I.1.3 Differentiation

Although hESCs retain the potential to differentiate into all cell types of the endoderm, mesoderm and ectoderm, to successfully differentiate them into a specific cell type of choice is not that easy. Thus far, hESCs have been successfully coaxed to differentiate into cardiomyocytes, hematopoietic progenitor cells, leucocytes, endothelial cells, neurons, glia, insulin-producing cells, and even germ cells (Hoffman and Carpenter, 2005). However, the efficiency of various protocols varies along with the utility of the differentiated cells. Essentially, success in this field depends on the knowledge of mechanisms controlling lineage decisions during embryogenesis, and subsequent recreation of these conditions in vitro (Spagnoli and Hemmati-Brivanlou, 2006).

I.2 Pluripotency of Human Embryonic Stem Cells

Pluripotency, or the potential for differentiation, is one of the defining features of hESCs.

I.2.1 Molecular mechanisms involved in pluripotency and self-renewal

Oct4, Sox2, Nanog network

The maintenance of pluripotency within hESCs requires extrinsic factors, either added to the growth medium or provided by growing on a feeder layer of differentiated cells (Smith et al., 1988; Williams et al., 1988), thereby creating an appropriate external signalling environment. It has become apparent that these factors are likely to be ultimately responsible for the maintenance of a network of key transcription factors that controls pluripotency within the cell. This network includes the homeodomain transcription factor Oct4 (Niwa et al., 2000), the variant homeodomain transcription factor Nanog (Chambers et al., 2003; Mitsui et al., 2003; Kuroda et al., 2005; Hough et al., 2006) and the high mobility group (HMG)—box transcription factor Sox2 (Avilion et al., 2003).

The identification of the many common targets and regulatory networks established by these transcription factors provides a current focus for understanding and controlling pluripotency. Each of these factors is required for pluripotency both *in vivo* and *in vitro*. For example, loss of Oct4 or Nanog results in the loss of pluripotency and the spontaneous differentiation to trophectoderm and primitive endoderm respectively (Nichols et al., 1998; Niwa et al., 2000; Chambers et al., 2003; Mitsui et al., 2003; Hough et al., 2006). In fact, the 3 transcription factors often work together and many

target genes are bounded by the combinations of Oct4, Nanog and Sox2 in hESCs, as reported recently on research using chromatin immuno-precipitation (ChIP) study (Boyer et al., 2005).

The optimal expression levels of these key transcription regulators are also important in guarding ESCs pluripotency. In particular, Oct4 and Sox2 have been found to bind to adjacent domains in regulatory regions of OCT4, NANOG (Okumura-Nakanishi et al., 2005; Kuroda et al., 2005; Rodda et al., 2005), as well as many other pluripotency-related genes such as UTF1, REX1 and TDGF1 (Rao, 2004; Boyer et al., 2005; Babaie et al., 2006).

"Reprogramming"

Pluripotency can also be re-established in a differentiated or non-pluripotent cell by exposure of nucleus to external factors. The introduction of the nucleus of a somatic cell into an enucleated oocyte, termed Somatic Cell Nuclear Transfer (SCNT), results in a cell gaining the ability to give rise to all cells of the mature organism, as demonstrated by the creation of Dolly the sheep and more than a dozen species of mammalian clones subsequently (Wilmut et al., 1997; reviewed in Gurdon and Byrne, 2004). More recently, the use of four specific transcription factors, c-myc, Sox2, Oct4 and Klf4, has been shown to induce pluripotency in both human and mouse somatic cells (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Lowry et al., 2008; Park et al., 2008a).

I.2.2 Chromatin structures and epigenetic features in ESCs

The state of chromatin in undifferentiated ESCs is regarded as a key factor in determining stem cell identity. Although the large-scale organization of genomes at the level of chromosomes is not significantly different between ES and differentiated cells, ES cell nuclei appear to have a distinct chromatin structure at key loci, which is involved in the maintenance of pluripotency or in the process of differentiation. For example, NANOG is localized significantly more centrally in the nuclei of hESCs compared with a more peripheral location in B cells, and OCT4 loops out from the chromosome territory in hESCs (Wiblin et al., 2005).

The ESC genome is transcriptionally globally hyperactive and undergoes large-scale silencing as cells differentiate. Whole-genome tiling arrays demonstrate widespread transcription in coding and noncoding regions in ESCs, whereas the transcriptional landscape becomes more discrete as differentiation proceeds. The transcriptional hyperactivity in ESCs is accompanied by disproportionate expression of the chromatin-remodeling genes and the general transcription machinery. Global transcription is proposed as a hallmark of pluripotent ESCs by contributing to their plasticities, and the lineage specification is driven by reduction of the transcribed portion of the genome (Cheng at al., 2008).

With recent analysis of mESCs, a model of its global chromatin dynamics has been raised. It is believed in pluripotent ESCs, DNA architectural proteins are loosely associated with the chromatin, and this allows a rapid recognition of the chromatin structure during differentiation (Meshorer and Misteli, 2006). Compared with that in non-pluripotent cells, this state is characterized by the rapid turnover of a subset of

proteins, associated with chromatin, including the H2B and H3 core histones (Meshorer et al., 2006).

The identification and characterization of stem cell chromatin regulators is therefore important for our understanding of stem cell pluripotency and embryonic development (Lee et al., 2006; Boyer et al., 2006; Meshorer et al., 2006; Buszczak and Spradling, 2006). As High Mobility Group (HMG) proteins are the most abundant non-histone DNA architechtual factor, we hypotesize that they could also be involved in ESCs chromotin remodeling.

I.2.3 Determination of pluripotency in hESC cells

There are three different aspects of pluripotency that can be easily assessed: cell morphology, transcript status and protein marker expression.

When cultured on MEFs, hESCs proliferate in tight and rounded colonies, and are morphologically quite distinct from other differentiated cell types, such as fibroblasts. In particular, hESCs should have a high nuclear to cytoplasm ratio, prominent nucleoli and be compact in size (Amit *et al.*, 2000).

The pluripotency of hESCs in culture can also be routinely assessed by mRNA transcript expression levels via either microarray, semi-quantitative reverse transcription PCR (RTPCR) or quantitative real-time PCR (qRT-PCR) techniques. Other techniques such as serial analysis of gene expression (SAGE) (Richards et al., 2006) and massively parallel signature sequencing (MPSS) (Wei et al., 2005) are also used to confirm pluripotency related transcripts as well as to elucidate new “stemness” genes.

The expression of certain cell-surface or intracellular proteins or markers can also be analyzed to gauge hESC pluripotency. Certain surface markers like the globoseries glycolipid stage specific embryonic antigens (SSEA) and the keratin sulphate-associated tumor recognition antigens (TRA) have been well characterized; pluripotent hESCs stain positively for TRA-1-60, TRA-1-81, and SSEA-4, but negatively for SSEA-1 (Henderson et al., 2002; Carpenter et al., 2004). This profile is similar to that of human embryonic carcinoma and human ICM cells. In contrast, mESCs stain positively for SSEA-1, but negatively for the other 4 markers. Despite being associated closely with hESCs pluripotency, the function of these surface antigens presents a conundrum, with various roles in cell-cell adhesion or embryonic pathology (Henderson et al., 2002).

In addition to the surface marker analyses, expression levels of transcription factor Oct4 and Sox2 were also examined via immunostaining in our lab for regular pluripotency determination. Another marker, alkaline phosphatase, a hydrolase enzyme with high expression in hESCs and responsible for catalyzing dephosphorylation, is also commonly used to determine the pluripotency of hESCs cultures based on its enzymatic actvities (Laslett et al., 2003).

I.3 Human Embryogenesis

Human embryogenesis is defined as the process of cell division and cellular differentiation of the human embryo during early prenatal development. In general, it is recognized as the first 8 weeks of gestational age after fertilization, the embryo is then usually called a fetus (Medical Encyclopedia, NIH, http://www.nlm.nih.gov/medlineplus/ency/article/002398.htm). In vivo, the inner cell mass forms a two layered embryo, composed of the epiblast and the hypoblast, which is made of columnar cells and cuboidal cells, respectively. The epiblast, now called primitive ectoderm will perform gastrulation, approximately at day 16 after fertilization. In this process, it gives rise to all three germ layers of the embryo: ectoderm, mesoderm, and endoderm. The hypoblast, or primitive endoderm, will give rise to extraembryonic structures only, such as the lining of the primary yolk sac. As human embryonic stem cells are derived from the ICM of the blastocyst, we refer our *in vitro* early embryogenesis in this work as the hESCs differentiation into all three germ layers (Figure I.2).

Ectoderm: emerges first and forms from the outermost of the germ layers. In vertebrates, the ectoderm consists of external ectoderm (also known as surface ectoderm) and neuroectoderm which consists of neural crest and neural tube. The ectoderm later forms central nervous system, lens of eyes, pigment cells, head connective tissues, epidermis, hair, and mammary glands.

Endoderm: is formed by cells migrating inward along the archenteron from the inner layer of the gastrula. The endoderm consists at first of flattened cells, which become columnar later. It forms the epithelial lining of the whole of the digestive tube, except part of the mouth, pharynx and the terminal part of rectum (which are lined by involutions of the ectoderm); the lining cells of all the glands which open into the digestive tube, including those of liver and pancreas; the epithelium of auditory tube and tympanic cavity, of trachea, bronchi, and alveoli of the lungs, of urinary bladder and part of urethra; and the cells which line the follicles of thyroid gland and thymus.

Mesoderm: forms during gastrulation when some of the cells migrate inward to form the endoderm, producing an additional layer that lies between the endoderm and the ectoderm. Mesoderm is found in the embryos of almost all large, complex animals that are more complex than cnidarians. The apperance of this layer allows the formation of a coelom (body cavity) that creats more room for independent growth of the body organs and the protective coelomic fluid.

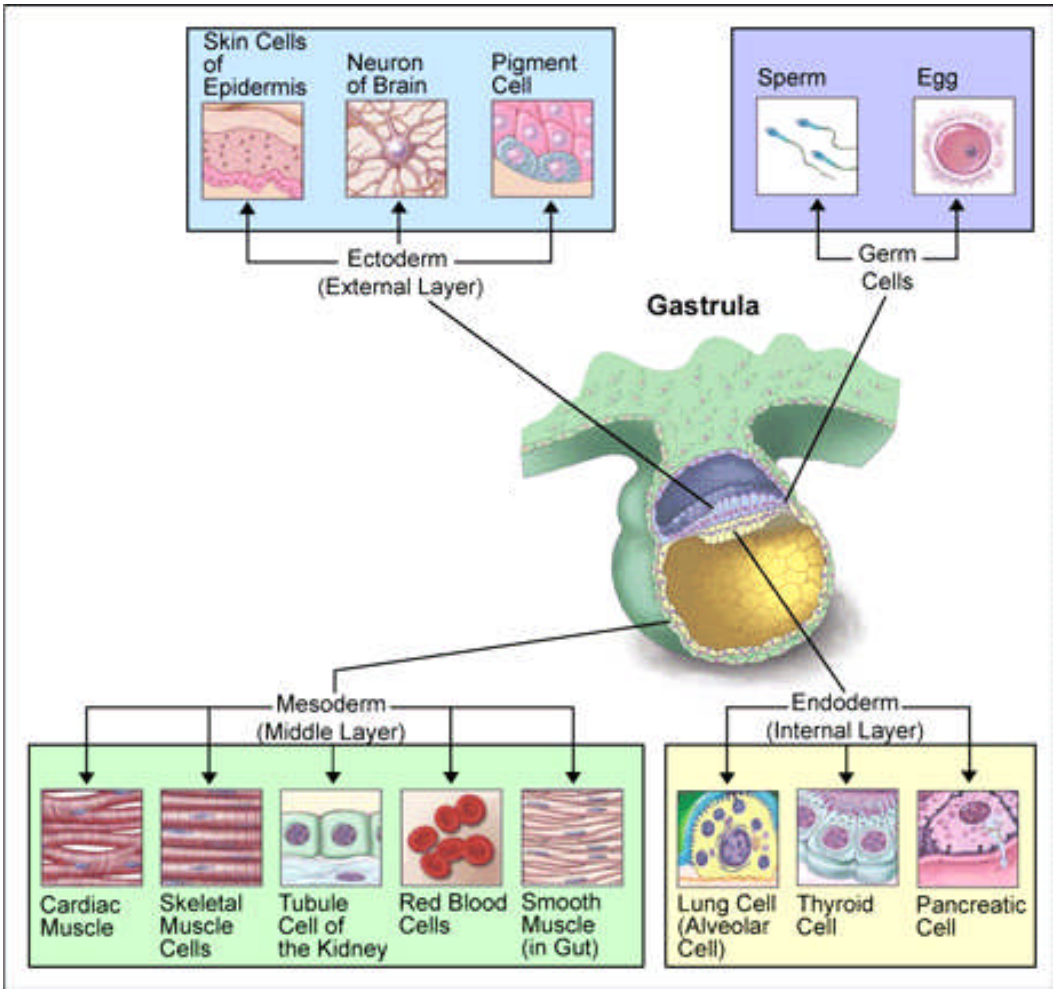


Figure I.2 Embryonic cells differentiate into a variety of different cell types.

The inner cell mass (light blue), which is derived to be embryonic cell is a cluster of cells inside the outer layer of blastocyst (yellow). The inner cell mass can give rise to the germ cells—eggs and sperm—as well as cells derived from all three germ layers (ectoderm, *light blue*; mesoderm, *light green*; and endoderm, *light yellow*), including nerve cells, muscle cells, skin cells, blood cells, bone cells, and cartilage.
(http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html)

I.4 High Mobility Group Proteins

High mobility group (HMG) proteins are the most abundant non-histone chromatin components (Bustin and Reeves, 1996). They were discovered more than 35 years ago as the acid-extractable protein components of chromatin with fast migration in the low pH electrophoretic gels. All HMG members are nuclear proteins with small size, i.e. less than 30 kDa, and most of them will undergo extensive post-translational modification. They are classified into three families with revised names of HMGA, HMGB, and HMGN with systematic reference to the respective domains contained (Bustin, 2001) (Table I.1).

HMGA: Featured with AT-hooks comprised of 8 or 9 amino acids each. These domains are unstructured in solution, and bind AT-rich DNA regions in the minor groove of the double-helix structure with low sequence specificity (Reeves, 2001).

HMGB: Featured with HMG Box comprised of about 80 amino acids. These domains also bind DNA in the minor grooves with limited or even no sequence specificity.

HMGN: Named by its property of binding between the Histone octamer and the DNA spires inside Nucleosome.

Members of these families usually bind DNA with relaxed sequence specificity and thereby modulate the functional status of individual chromatin domains, forming a Yin-Yang relationship with Histone H1 protein (Bianchi and Agresti, 2005). For example, upon DNA binding, HMGA and HMGB proteins can alter the path of DNA segments through bending and thereby govern the formation of specialized nucleoprotein structures involved in genomic transactions, such as transcription,

recombination, and retroviral gene expression (Grosschedl et al., 1994).

Generally, HMGs are considered transcription-friendly, and they are also known in modulating specific genes. HMGNs do so by direct binding to nucleosomes, directly competing with histone H1. It was reported that H1 mobility on the nucleosomes is increased with increasing HMGN proteins amount (Catez et al., 2002). On the other hand, HMGAs regulate specific gene transcription by organizing multiprotein complexes of transcription factors and cofactors, eg. *IFN-γ*. HMGBs seem to have both properties of that in HMGAs and HMGNs (Agresti et al., 2003).

Table I.1 Members of the HMG protein family.

Protein, New Name	Protein, Old (Alternate) Name	Gene, New Symbol	Accession #
HMGB Proteins			
HMGB1	HMG1; HMG-1;HMG 1; (amphoterin)	HMGB1 Hmgb1	H: U51677 M: U00431
HMGB2	HMG2; HMG-2; HMG 2	HMGB2 Hmgb2	H: M83665 M: X67668
HMGB3	HMG2a; HMG-2a; HMG 2a (HMG-4)	HMGB3 Hmgb3	H: NM_005342 M: Y10044
HMGN Proteins			
HMGN1	HMG-14; HMG14; HMG 14	HMGN1 Hmgn1	H: M21339 M: NM_008251
HMGN2	HMG-17; HMG17; HMG 17	HMGN2 Hmgn2	H: X13546 M: NM_016957
HMGN3	Trip7	HMGN3 Hmgn3	H: L40357
HMGN3a		HMGN3 Hmgn3	
HMGN4		HMGN4	H: NM_006353
HMGA Proteins			
HMGA1a	HMG-I; HMGI; HMG I HMG-I/Y; a-protein	HMGA1 Hmga1	H: L17131(1) M: J04179(2)
HMGA1b	HMG-Y; HMGY; HMG Y	HMGA1 Hmga1	H: M23618(2) M: J04179(2)
HMGA2	HMGI-C	HMGA2 Hmga2	H: L46353 M: L41617

I.4.1 HMGA sub family

Mammalian HMGA proteins are encoded by two genes, HMGA1 and HMGA2. Alternative mRNA splicing gives rise to different protein isoforms. These chromatin-associated proteins have no intrinsic transcriptional activity but bind to AT-rich DNA sequences and potentiate the effects of transcription factors by altering local chromatin structure (Fusco and Fedele, 2007). HMGA proteins are characterized by their small sizes (11 to 12 kDa), an acidic C-terminal tail, and the presence of three individual AT-hooks (Figure I.3).

Figure I.3 Aligned amino acid sequences of the three major human HMGA molecules.

The three conserved AT-hook DNA-binding domains (AT-hook) are in bold. The C-terminal carboxy acidic tail is frequently lost in human neoplasia (Cerignoli et al., 2004).

The 3 AT-hooks are responsible for the binding of the protein to the DNA minor grooves with AT-stretches. The motif of the AT-hook is BBXRGRXB of which B is a basic amino acid (Lysine or Arginine), and X is either Glycine or Proline. When bound to the DNA minor groove, an AT-hook adopts a crescent-shaped configuration (Huth et al., 1997), and the mostly conserved RGR sequence was reported as the key

for this interaction through analyzing oligonuleotide-protein complex with NMR technology (Geierstanger et al., 1994).

The C-terminal of the HMGA proteins is generally regarded as the acidic tail as it contains many hydroxyl amino acids, mainly Glutamic acid. The function of the acidic tail is still not fully known, although many amino acids could be phosphorylated by casein kinase II (Heath, 1995). It is thought to have an influence on HMGA's protein–DNA or protein–protein binding, and thus modify the "architectural transcription" activity of HMGA proteins for certain specific target genes (Yie et al., 1997; Borrmann et al., 2003).

Apart from the detailed studies on topology, very little is known about how HMGA proteins containing three separate DNA binding domains interact stably with nucleosomes, and how this interaction affects nucleosome structure and ultimately the higher-order organization of chromatin fibers (Reeves and Nissen, 1993; Reeves and Wolfe, 1996; Sgarra et al., 2004, Reeves, 2004).

The study of HMGA proteins has shown that they are involved in many cellular processes, including gene regulation, cell cycle, differentiation, and viral integration (Reeves, 2001). Therefore, HMGA mutations are often related with many common diseases, such as benign and malignant tumors (Sgarra et al., 2004), diabetes (Foti et al., 2005), obesity (Anand and Chada, 2000), and atherosclerosis (Schlueter et al., 2005). In fact, the HMGA genes are the only known proto-oncogenes coding for DNA architectural chromosomal proteins consistently over-expressed in almost all types of naturally occurring cancers. Their expression level correlates directly with the degree of malignancy and metastatic potential (Tallini and Dal, 1999; Fusco and Fedele, 2007;

Cleynen and Van de Ven, 2008; Di Cello et al, 2008). In line with these observations, recent micro-array analyses showed that both HMGA genes belong to a small class of genes whose expression are critical to the cancer phenotype in cells that carry two key oncogenic mutations in p53 and Ras (McMurray et al., 2008).

I.4.2 HMGA2

I.4.2.1 Overview

HMGA2 is encoded by a gene located on Chromosome 10 in mouse and Chromosome 12q13-15 in human. In human, the protein is encoded in 5 exons from its gene. The first exon encodes the first 37 amino acids including the first AT-hook. The second and third AT-hooks are encoded in the next two exons. The fourth exon encodes an 11 amino-acid segment that separates the third AT-hook from the acidic C-terminal tail which is encoded by the last exon. Recently, it was reported that HMGA2, as a DNA architectural protein, has its nuclear localization signal overlaped with the second AT-hook, which directly interacts with nuclear import receptor Importin-α2 (Cattaruzzi et al., 2007).

It was rarely mentioned in the literatures that HMGA2 has a splice variant, or isoform b, consisting of the same first 3 exons and an alternate exon 4, which comes from intron 3 and results in an early stop codon. The resulting protein has a distinct C-terminus and is shorter than isoform a (Figure I.4).

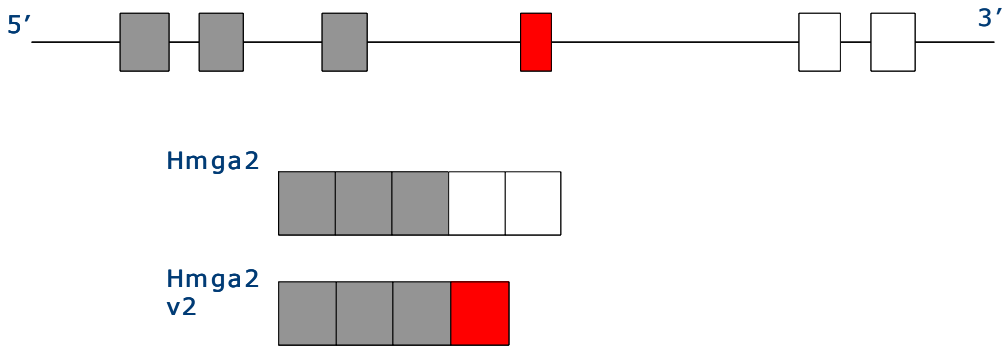


Figure I.4 Alternate splicing of HMGA2 gene.

The first 3 exons (in grey) were the same between isoform a and isoform b. However, isoform be consists an alternate exon 4 which comes from intron 3 (shown in red), results in an early stop codon, thus a shorter gene product than isoform a which contains normal exon 4 and 5 (in white).

The expression of HMGA2 is usually high during early embryogenesis and undetectable in most human adult tissues (Reeves, 2001). Recent reports identify HMGA2 as a target of the let-7 family of microRNAs (miRNAs), of which the binding sites are located on the 3' untranslated region (UTR).

miRNAs encode a class of small, regulatory noncoding RNAs, which have come to prominence as a key mechanism of post-transcriptional regulation within the last decade (Ambros, 2004; Bartel, 2004; Zamore and Haley, 2005). Let-7 is one of the first miRNA genes to be discovered in Caenorhabditis elegans, a heterochronic gene regulating the developmental timing of the larvae. The let-7 family's expression is also developmentally regulated in mice and humans (Schulman et al., 2005; Pasquinelli et al., 2000). Expression of HMGA2 is high in embryonic cells and very low in differentiated cells, a pattern exactly inverse to that of let-7 family members, further suggesting a regulatory relationship between them (Droge and Davey, 2007).

I.4.2.2 HMGA2 and Development

HMGA2 has been studied in detail in mouse where it is expressed in mESCs and during embryogenesis. Recent reports suggested that its expression is upregulated in mESCs undergoing differentiation towards cardiomyocytes lineage (Caron et al., 2005). HMGA2 expression *in vivo* was observed throughout embryos, before 15.5 days of post coital age, with exception of the brain, where the expression was only observed in a small portion in the forebrain (Zhou et al., 1995).

Although HMGA2 expression is mainly embryonic (Zhou et al., 1995), its expression has been detected in testis (Chieffi et al., 2002), adipose tissues (Anand and Chada, 2000), CD34+ hematopoietic stem cells, uterine myoblasts, testicular cells, and skeletal muscle cells (Hirning-Folz et al., 1998; Caron et al., 2005) of adult mice, suggesting functional links between these tissues and HMGA2.

HMGA2 function seems critical for cell growth and adipocytic cell differentiation because mice lacking functional HMGA2 exhibit a pygmy phenotype with greatly reduced fat tissue (Zhou et al., 1995). The involvement of HMGA2 in adipogenesis was confirmed by recent studies that showed mice lacking HMGA2 are resistant to obesity (Anand et al., 2000). In detail, HMGA2–/– and HMGA2+/– mice fed a high-fat diet for 26 weeks fail to develop obesity whereas wild-type mice on the same diet do become obese. Similarly, HMGA2–/– mice when bred to obese/obese mice fail to acquire an obese phenotype and are weighed approximately only 1/4 of obese/obese mice. In addition, over-expression of a truncated HMGA2 protein lacking C-terminal tail results in gigantism and lipomatosis (Battista et al., 1999). However, very little is known about the underlying molecular mechanisms which lead to the development of

these phenotypes.

HMGA2 may also play an age-associated role in the self-renewal of mouse neural stem cells (NSCs) possibly by regulating the Ink4a/Arf expression through repressing the expression of their transcription activator JunB. Although HMGA2 does not appear to be required for the generation of NSCs during fetal development, NSCs from HMGA2-deficient mice have defects in proliferation and self-renewal. With the observation that differentiating neural progenitors from these mice do not exhibit proliferative defects, it was suggested that HMGA2 loss does not lead to a global decrease in cellular replication but specifically affects stem cell self-renewal (Nishino et al., 2008).

In humans, HMGA2 expression has been detected in most fetal cells with gestational age less than 22 weeks, with higher expression in fetal kidney, liver and uterus. However, it is only detected in lung and kidney samples in human adult. On the other hand, HMGA2 is observed to be upregulated in many cell cultures *in vitro*. The reason is speculated to be the involvement of some stimulatory component in the serum used (Gattas et al., 1999).

Interestingly, a germ-line truncated HMGA2 gene was recently found to be associated with a case of extreme somatic overgrowth and multiple lipomas, thus also linking HMGA2 to general cell growth and adipogenesis during human development (Ligon et al., 2005). Furthermore, recent data hinted at novel roles for HMGA2 in heterochromatin formation in cultured human cells and in epithelial–mesenchymal transition during embryogenesis and metastasis (Thuault et al., 2006; Narita et al., 2006).

These observations and the finding that cyclin A is a direct target of HMGA2 (Tessari et al., 2003), point at important roles for HMGA2 in the regulation of ESC growth, cell growth during embryogenesis, and cell transformation. While it is reasonable to assume that HMGA2 acts locally and participates in the regulation of individual genes, such as cyclin A, it remains possible that HMGA2 also has a more global effect on the state of ESC chromatin and may in this way contribute to the establishment of ESC identity. In this respect, it is interesting to note that a high copy number of HMGA2-encoding mRNAs is already present during oogenesis in Xenopus leavis (Hock et al., 2006).

I.4.2.3 HMGA2 and Cancer

High HMGA2 protein levels are often associated with malignant tumors such as liposarcomas and osteosarcomas, acute lymphoblastic leukemia, neuroblastomas, as well as many lung carcinomas (Berner et al., 1997; Xu et al., 2004; Cerignoli et al., 2004; Sarhadi et al., 2006). HMGA2 gene rearrangements by reciprocal translocations are frequently found in benign tumors of mesenchymal origin (Fedele et al., 2001), such as lipomas, lung hamartomas, uterine leiomyomas, endothelial polyps, fibroadenomas, and adenolipomas of the breast (Ashar et al., 1995; Kazmierczak et al., 1995, 1996; Schoenmakers et al., 1995). In most HMGA2 translocations, the breakpoint occurs within the ~140 kb third intron, giving rise to chimeric transcripts that are a fusion of HMGA2's three AT-hooks with various other genes (Goodwin, 1998). It was therefore presumed that the fusion proteins are the cause of tumorigenicity.

Surprisingly, however, a truncated form of HMGA2 itself, containing only the three

AT-hook domains, is sufficient for transformation *in vitro*. Therefore the loss of the acidic tail, rather than the chimeric proteins generated from the translocation, is likely to be the critical step in causing HMGA2 oncogenicity (Fedele et al., 1998). It is proposed that the loss of HMGA2's 3' UTR via chromosomal translocations within the mRNA results in the loss of let-7 binding and negative regulation, and in consequence contributes to tumorigenesis (Park et al., 2007).

Interestingly, the let-7/HMGA2 linkage was recently identified as a candidate signature of cancer stem cells derived from primary human cancer tissues. These cells are resistant to an *in vivo* therapy which introduces chemical modifications of DNA bases, such as alkylation (Yu et al., 2007; Park et al., 2007; Dröge and Dave, 2008). There is evidence that HMGA2 is involved in cancer stem cell differentiation and proliferation control (Yu et al., 2007), but due to the pleiotropic functions of HMGA2 as a DNA architectural chromatin factor, the molecular mechanism(s) which connect the let-7/HMGA2 linkage with chemotherapy-resistance of cancer stem cells remain elusive. It is noteworthy here that the presence of HMGA1 is also functionally linked to chemo resistance of certain types of human carcinomas (Liau and Whang, 2008).

I.5 HMGA2 and Base Excision Repair (BER)

I.5.1 DNA damages repaired by BER

Alkylation damage is one common type of DNA damage. Cells have several pathways to repair small alkyl adducts including BER, whereas the large alkyl adducts are processed by nucleotide excision repair (NER) (Sedgwick et al., 2007).

Methyl Methanesulfonate (MMS) is the most important environmental alkylating agent, which can create covalent modifications at ring nitrogen residues of DNA bases, in particular 7-methylguanine (7-meG) and 3-methyladenine (3-meA) (Strauss et al., 1975). Whereas 7-meG appears to be relatively innocuous, 3-meA has a strong cytotoxic effect by blocking both replication and transcription through the aberrant protrusion of the methyl group into the DNA minor groove (Boiteux and Laval, 1982; Larson et al., 1985).

Another type of DNA damage reagent, reactive oxygen species (ROS), such as hydrogen peroxide, superoxide, and hydroxyl radicals, are byproducts of the normal aerobic metabolism, but can also be produced after ionizing radiation, for instance (Gajewski et al., 1990). DNA bases are very susceptible to ROS-mediated oxidation, resulting in oxidized bases, formation of apurinic/apyrimidinic (AP) sites and strand breaks (Neeley and Essigmann, 2006).

One of the most abundant lesions provoked after oxidative treatments is 8-oxoguanine (8-oxoG). 8-oxoG is best studied as strongly mutagenic because it can result in G-to-T transversion mutations through misincorporation of adenine, which is additionally cytotoxic (Wiederholt and Greenberg, 2002; Ober et al., 2003). Replicative DNA

polymerases can bypass this lesion very efficiently, in contrast to many other types of DNA damage (Shibutani et al., 1991). In human cells, it is repaired by BER with 8-oxoG DNA glycosylase 1 (OGG1) cleavage as the first step.

Apurine/Apyrimidine (AP) site itself is a kind of DNA damage. In general, AP site formation can either stem from the enzymatic action of *N*-glycosylases after processing a base lesion or from spontaneous hydrolysis of the *N*-glycosyl bond, by low pH (Siede et al., 2006). Low pH, high temperature and chemical modification, such as oxidation as well as alkylation, of both base and sugar can greatly speed up the formation of AP site (Lindahl et al., 1993; Loeb and Preston, 1986). In fact, AP site formation is such a frequent event that human genome undergoes an estimated 10,000 times base loss per cell per day. The rate of depurination is 100 to 500 times greater than depyrimidination. The consequence of leaving an un-repaired AP site is, in some extreme cases, fatal, since it could stall replication folk and induce double strand break. AP sites are also a major cause of mutagenesis and carcinogenesis, due to either the translation synthesis of AP site by PolV or the low fidelity of Polβ (Loeb and Preston, 1986).

I.5.2 BER mechinary

The repair of AP site and many kinds of base modification are mostly accomplished by base excision repair (BER) pathway. BER is a highly conserved DNA repair mechanism through evolution. The BER pathway fixes lesions in bases that are similar in size and shape to the normal bases. These base lesions include deaminated cytosine, 5-methylcytosine, and adenine, but also the oxidation products of all four bases and some types of base alkylation. Typically, only a small region (1 to 13 nucleotides)

around the damaged base is removed and replaced during BER, in contrast to some other excision repair mechanisms, such as NER.

BER is initiated by a wide variety of glycosylases, which catalyze the breakage of *N*-glycosyl bond and consequently produces an AP site (Krokan et al., 1997). The AP site is then further cleaved by an AP endonuclease via a hydrolytic mechanism (Siede et al., 2006), which produces 3'-hydroxyl and 5'-deoxyribose phosphate (dRP) group. Alternatively, the AP site can also be cleaved first by an AP lyase through β-elimination (McCullough et al., 1999; Dodson et al., 1994), which produces 3'-α, β-unsaturated aldehyde and 5'-phosphate. In short patch BER, either 5'-dRP end or 3'-α, β-unsaturated aldehyde end will be further processed by dRP lyase or 3' diesterase activity associated with AP endonuclease (Sander and Wilson, 2005), respectively, for the release of abasic sugar ring and an intact 3'-hydroxyl terminus. Finally, DNA polymerases utilize the newly generated 3'-hydroxyl terminus to fill in the gap, followed with the ligation by DNA ligases. In long patch BER, DNA polymerase such as Poly (ADP-ribose) polymerase (PARP) is responsible for extension from the 3'-OH group left by AP endonuclease, replacing the strand containing the 5'-dRP. The flap structure is removed by structure specific flap endonuclease (FEN1). The resulting strand break is ligated by DNA ligase (Figure I.4).

One important point in BER is that if any of these steps are interrupted in the middle of the repair process, a much more serious DNA damage will be produced. Therefore, BER components must work in a tightly regulated synergistic fashion. Evidence has been found regarding the intimate partnership between ligase Ⅲ and XRCC1 (Caldecott et al., 1994), and the close interaction between bi-functional glycosylase, which contains also the AP lyase, and human AP endonuclease Ape1 activity

(Sidorenko et al., 2007).

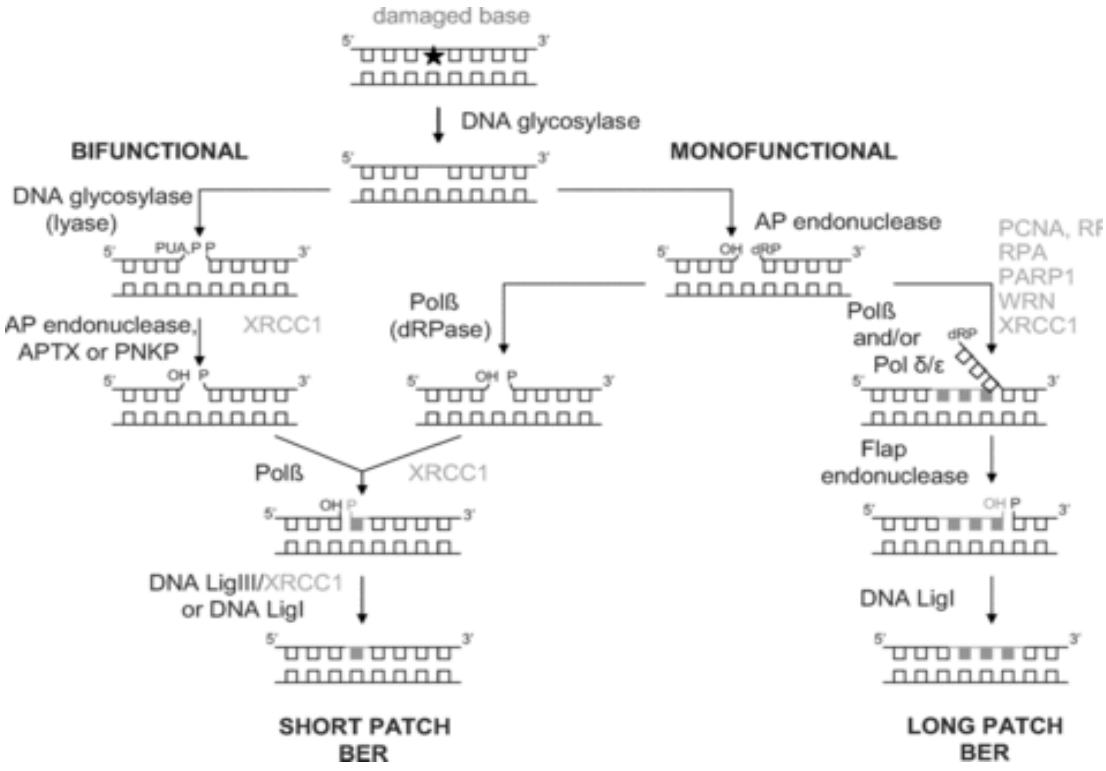


Figure I.5 Schematic BER pathway and different sub-pathways in mammals.

BER starts with the recognition and removal of a lesion (star) by a DNA glycosylase. Only bifunctional DNA glycosylases are able to cleave the sugar-phosphate backbone and create a 5' phosphate (P) and a 3' phosphate or 3' polyunsaturated aldehyde (PUA), depending on the DNA glycosylase. After removal of the damaged base by monofunctional DNA glycosylases, strand scission is exerted by AP endonuclease, creating 3' hydroxyl (OH) and 5' deoxyribose-phosphate (dRP). These unconventional termini have to be restored to 3' OH and 5' P to allow further repair through deoxyribose-phosphatase diesterase (dRPase) activity of Pol β (5' dRP), diesterase activity of AP endonuclease (3' PUA), phoshatase activity of polynucleotide kinase phosphatase (PNKP) (3' P), or phosphatase activity of aprataxin (APTX) (3' P). Repair then proceeds via short-patch or long-patch repair. Pol β incorporates one nucleotide, followed by nick ligation by the XRCC1/LigIIIα complex (predominantly) or LigI. If the 5' lesion is refractory to cleavage by Pol β, the long-patch branch of BER is taken. Pol β and/or Pol δ/ε accomplish strand displacement by incorporating multiple nucleotides, followed by removal of the DNA flap containing the 5' refractory moiety by Flap endonuclease and ligation of the resulting nick by LigI. Supportive BER proteins are indicated in gray (Figure from Baute and Depicker, 2008).

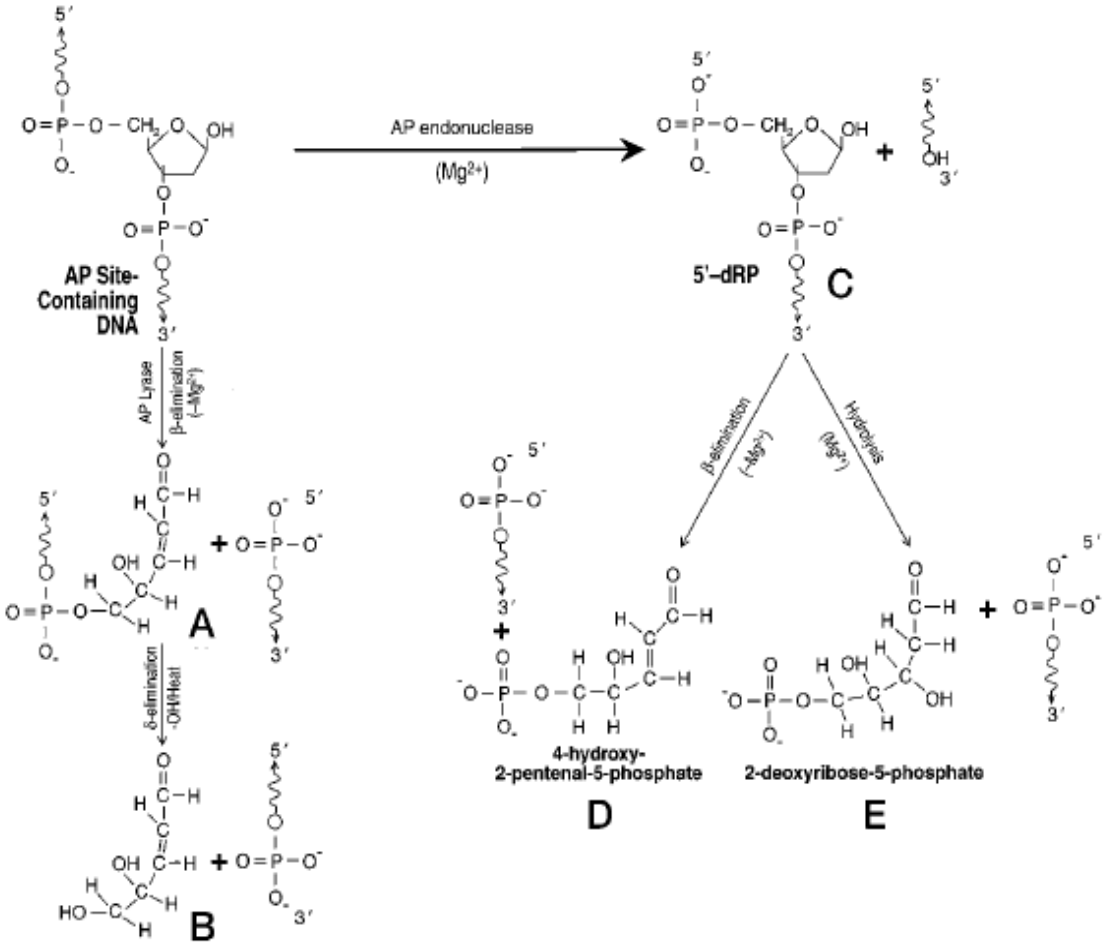


Figure I.6 AP site chemistry.

Polβ can cleave AP site DNA by a β-elimination mechanism leaving 5'-phosphate and 3'-a,b-unsaturated aldehyde termini (A), which may undergo d-elimination in the presence of alkali/heat (B). The 3'- and the 5'-termini generated by this reaction are phosphate groups. The β, δ-elimination reactions do not require Mg2+. Alternatively, the AP site can be hydrolyzed by AP endonuclease producing DNA containing 5' 2-deoxyribose-5-phosphate (5'-dRP) (C). This reaction requires Mg2+. The 5'-dRP is cleaved enzymatically, producing either 4-hydroxy-2-pentenal-5-phosphate by β-elimination (D) or 2-deoxyribose-5-phosphate by hydrolysis (E) (Figure adapted from Sobol et al., 2000).

I.5.3 AP/dRP lyase and HMGA2

The AP lyase was discovered first as an associated enzymatic activity from bifunctional *N*-glycosylases, such as human 8-oxoG DNA glycosylase (hOGG1), *E. coli* endonuclease Ⅲ (Endo Ⅲ), uracil DNA glycosylase (UDG), etc. The glycosylase/AP lyases recognize a damaged base or AP site as their substrate. An activated Schiff base nucleophile from the active center attacks the C1 at the sugar ring of the substrate and forms a protonated Schiff base intermediate (Figure I.5). The protonation imparts positive charge on the Schiff base intermediate, making it a good "electron sink" (Gerlt et al., 1993), thereby lowering the pKa of the sugar 2'-hydrogen, and making it a good leaving group for subsequent β-elimination chemistry. After the concomitant bond breakage between C3 and the oxygen, the enzyme is released from its substrate through hydrolysis (McCullough et al., 1999).

During our studies of DNA lyase activities associated with proteins belonging to the *E. coli* DNABII family of DNA binding proteins (unpublished result), we discovered that recombinant human HMGA proteins efficiently cleave DNA containing AP sites *in vitro*.

In this study, we present here a biochemical analysis of these associated lyase activities and demonstrate with cell-based experiments that HMGA2 is directly involved in base excision repair (BER) and protects cancer cells from DNA-damage occurred, for example, during chemotherapy.

II MATERIALS and METHODS

II.1 Routine Cell Line Maintenance

Cell lines:

Name	Details	Source
HeLa	Human cervical carcinoma cells	Lab Stock ATCC No. CCL-2
A549	Human lung alveolar basal epithelial carcinoma cells	Lab Stock ATCC No. CCL-185
MCF7	Human breast carcinoma cells	Lab Stock ATCC No. HTB-22
F3	Primary human foreskin fibroblast cells	Technion-Israel Institute of Technology
PMEF-N	Neomycin-resistant Mouse Embryonic Fibroblast cells	Speciality Media

Culture media:

Dulbecco's modified eagle medium (DMEM)
10% (v/v) fetal bovine serum (FBS)
2 mM L-glutamine
100 unit/ml penicillin/streptomycin

Freezing media:

90% (v/v) FBS
10% (v/v) Dimethyl sulfoxide (DMSO)

Cell lines were cultured on tissue-culture grade dishes and flasks (Nunc or TPP) at 37°C with 5% CO2, and passaged when confluency was reached more than 90% in a proper split ratio. Phosphate buffered saline, PBS (Gibco), was first used to wash the cells before the addition of 0.25% trypsin-ethylenediaminetetraacetic acid (EDTA)

(Gibco). Cells were then incubated at 37°C for 5 min before dissociation by pipetting. Culture media were added to the cells to neutralize trypsin, and cells were collected by centrifugation at 600 g for 3 min followed by being plated out with new culture media on new dishes/flasks.

These cell lines were frozen via the standard slow cooling protocol. Briefly, cultures were centrifuged at 600 g for 3 min after trypsinization and neutralization. This pellet was then resuspended in freezing media and split into cryovials (Nalgene; 500 µl per vial). The vials were placed into a Mr. Frosty freezing aid (Nalgene) and incubated at -80°C overnight. The next day, they were transferred to either a -152°C freezer or liquid nitrogen tanks. For thawing, frozen cryovials were incubated in a 37°C water bath until the frozen contents were almost fully thawed and then added drop-wise to 10 ml of culture media. Cells were collected by centrifugation at 600 g for 3 min, re-suspended in culture media, and plated out.

II.2 Human Embryonic Stem Cell Culture

II.2.1 Gelatin- and matrigel-coated plate preparation

1% gelatin (w/v, diluted with distilled water) (Sigma) stock solution was autoclaved and stored at -20°C. Plates were then coated by incubating 0.1% gelatin solution on them for at least one hour in a 5% CO2 incubator before MEF was plated. Growth factor-reduced matrigel matrix (BD Biosciences) was diluted 1:20 in cold DMEM (Gibco), dispensed onto dishes and incubated at 4°C overnight to form a thin matrigel layer on the culture dish.

II.2.2 Mouse embryonic fibroblast feeder plate preparation

MEFs were cultured until passage 7 at the time when they were mitotically inactivated. A confluent T175 flask was incubated with 10 µg/ml mitomycin C (Sigma) in culture media for three hours in a 5% CO2 incubator. These inactivated MEFs were washed with culture media and collected via trypsinization, and quantified using a haemocytometer. Cells were frozen down at density of 1.2×10^6 cells per vial (enough to plate 6 Falcon centre well organ culture dishes or three 3 cm dishes). MEFs were either thawed or seeded directly from fresh, collected cultures at density of 1.5×10^5 cells onto each gelatin-coated organ culture dish (twice this number for a 3 cm dish). These were incubated overnight to allow cell attachment before they were ready for hESC plating. Plates between 24 hrs and 48 hrs after MEF plating were ideal for hESC plating. MEF plates older than 96 hrs were not usable.

II.2.3 hES cell lines

Name	Karyotype	Race	Time Derived	Source
hES2	46XX	Chinese	1997	ES Cell International Pte Ltd
HUES7	46XY	*in vitro* Fertilized	2003	Melton hES Cell Facility, Harvard University
HUES8	46XY	*in vitro* Fertilized	2003	

II.2.4 hES2 cell culture

Culture Media:

Dulbecco's modified eagle medium (DMEM)
10% (v/v) defined fetal bovine serum (dFBS)
2 mM L-glutamine
1X Insulin-Transferrin-Selenium (ITS)
Non-Essential Amino Acids (NEAA)
0.1% (v/v) f3-Mercaptoethanol
100 unit/ml penicillin/streptomycin

hES2 cells were cultured on K4 MEF-plated organ culture dishes (Falcon). hES2 colonies were mechanically cut and dissociated using a 27G1/2 needle as a scalpel (6 pieces from each large colony), and plated onto MEF dishes at density of 6-8 colonies per organ culture dish (colonies or colony pieces that looked differentiated were not passaged). Media were changed everyday; hES2 colonies were passaged every 7 days. Colony manipulation was performed under an Olympus microscope equipped with a heated stage at 37°C at 12.5X magnification.

II.2.5 hES2 vitrification

ES-HEPES Media:

DMEM with 20% (v/v) defined FBS and 20 mM HEPES

10% Vitrification solution:

ES-HEPES media with 10% (v/v) Ethylene Glycol and 10% (v/v) DMSO

20% Vitrification solution:

ES-HEPES media with 20% (v/v) Ethylene Glycol and 20% (v/v) DMSO

hES2 cells were frozen via the vitrification process. Briefly, hES2 colonies were dissociated into large pieces (4 pieces from each large colony), and placed in ES-HEPES media for 5 min. The pieces were transferred to the 10% vitrification solution for 1min, and then to the 20% vitrification solution for 25 sec. A maximum of 8 colony pieces were transferred together via micropipette into a small 20 µl droplet. All the colony pieces were then transferred to a smaller 3 µl droplet. Subsequently, the colony pieces were drawn into a vitrification straw via capillary action, flash frozen in liquid nitrogen, and placed in a 5 ml perforated cryovial (Nalgene) for storage in liquid nitrogen.

II.2.6 hES2 thawing

hES2 colony pieces were thawed by placing the vitrification straw directly from liquid nitrogen into 0.2 M sucrose/ES-HEPES solution at 37°C. All the colony pieces, which were expelled from the straw by the thawing vitrification solution, were then transferred to another well containing 0.2 M sucrose/ES-HEPES solution. Thereafter, they were transferred into ES-HEPES media for 5 min, and transferred to another well with ES-HEPES media for an additional 5 min, before finally being plated onto a

MEF-layered dish. Media were changed everyday thereafter, and hES2 colonies were ready to be passaged after 10 to 14 days.

II.2.7 HUES cell culture

Culture Media: Knockout DMEM

10% (v/v) Knockout Serum Replacement
10% (v/v) Plasmanate
2 mM Gluta-MAX
Non-Essential Amino Acids (NEAA)
10 ng/mL basic fibroblast growth factor (bFGF)
0.1% (v/v) f3-Mercaptoethanol
100 unit/ml penicillin/streptomycin

HUES hESCs were cultured on PMEF or PMEF-N MEF-layered 3 cm dishes and passaged via trypsinization. Confluent cultures were washed with PBS and dissociated via incubation with 0.6 ml of 0.05% trypsin-EDTA per 3 cm dish for 5 min and resuspended. 0.6 ml of HUES culture media was then added to the cell suspension to neutralize the trypsin. 0.24 ml of this culture was seeded into each MEF layered 3 cm dish (equivalent to a 1:5 split), and topped up to 2 ml with HUES culture media. Media were changed everyday and cultures were usually confluent and ready for the next passage in 5-7 days.

HUES cells were frozen and thawed via the procedure as described in II.1.

II.3 Regular Pluripotency Examination of hESCs

II.3.1 Alkaline phosphatase staining

Alkaline phosphatase (AP) is a hydrolase enzyme responsible for catalyzing dephosphorylation and is believed to be highly expressed in embryonic stem cells. AP staining was performed on hESCs using the ES cell characterization kit (Chemicon). Briefly, the cells were grown up to 50% in confluency before fixed with 90% Methanol/10% Formalin (Merck) for 1min, washed with PBS, and incubated with freshly prepared mixture of Napthol/Fast Red Violet/water at a 2:1:1 ratio for 15 min in the dark. The cells were then washed with PBS again and viewed under normal reverse-phase light microscope. Photomicrographs were taken using the IX71 and DP70 digital imaging system (Olympus).

II.3.2 Pluripotency marker immunostaining

Human ESCs to be stained were cultured on MEF-layered 96-well black/clear bottom plates (Greiner) until they reach 70% in confluency. They were first fixed in 4% Paraformaldehyde (Sigma)/PBS for 15min, washed twice with rinse buffer, and then permeabilized with 0.1% Triton X-100 (Sigma)/PBS for 10min. Subsequently, the cells were washed twice with rinse buffer and blocked with 4% goat serum (Santa Cruz)/PBS for 30 min. The hESCs were incubated with the primary antibodies (1:20 dilution with blocking solution) for 60 min and washed thrice with rinse buffer. Secondary antibodies were then added (1:50 dilution with PBS) to the hESCs, and cells were incubated for 60 min and washed thrice with rinse buffer. Photomicrographs were taken using the IX71 and DP70 digital imaging system (Olympus).

II.3.3 Embryoid body (EB) formation

Conditional media were used for the culture of EB. To prepare them, ES media were first used to culture MEF cells for 24 hrs. They were then drawn out and filtered through syringe driven PVDF filter with 0.22 μM in pore size before added into plates containing EBs.

hES2 cell EBs were generated by manually dissociating cells with a 26.5 gauge needle (BD Biosciences) into pieces of approximately 5 x 10^3 cells each. 30 to 35 pieces were then transferred into low-binding tissue culture dishes (NUNC) and cultured in conditional media for designated periods before harvesting. For EB generation from HUES cells, the cells were grown up to 90% in confluency, and dislodged from the plate by tripsinizition before being cultured in the low-binding tissue culture dishes (Figure II.1).

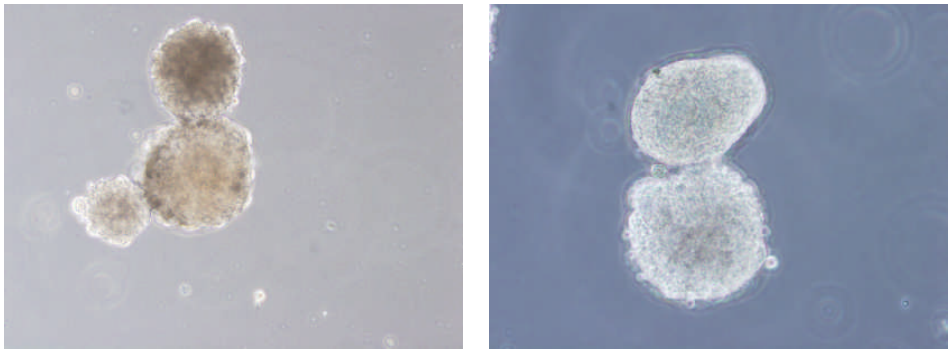


Figure II.1 Images of Embryoid Bodies after 7 days of culture.

The Embryoid body has a sphere-like 3-D structure and is non-adherent to the culture surface. hES2 cell EBs were generated by manually dissociating cells with a 26.5 gauge needle into pieces and HUES cell EBs were generated by dislodging cells from the plate by tripsinizition, both was then cultured in low cell-binding plates with condition media.

II.4 Expression Study of HMGA2 in hESCs

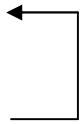
II.4.1 Reverse Transcriptase-PCR

Total RNA from the cells was extracted using RNeasy Mini Kit (QIAGEN) according to the manufacturer's protocol, and concentrations were determined spectrophotometrically at λ260. 100 ng of RNA was used as the template for RT-PCR using a One-step RT-PCR Kit (QIAGEN). The following oligonucleotides designed for β-actin and each of the HMG genes were used as primers:

β-actin	F	5'-CAC CTT CTA CAA TGA GCT GCG-3'
	R	5'-TGC TTG CTG ATC CAC ATC TGC-3'
HMGA1a/b	F	5'-ATG AGT GAG TCG AGC TC-3';
	R	5'-TCA CTG CTC CTC CTC CG-3'
HMGA2	F	5'-ATG AGC GCA CGC GGT GAG-3'
	R	5'-GGC TTC TTC TGA ACA ACT TG-3'h
HMGA2b	F	5'- ATG AGC GCA CGC GGT GAG-3'
	R	5'-GGT ACT GTT TCC FCAG AGA TG-3'
HMGB1	F	5'-GGA GAT CCT AAG AAG CCG-3'
	R	5'-TTC AGC CTT GAC AAC TCC C-3'
HMGB2	F	5'-ATG TCC TCG TAC GCC TTC-3'
	R	5'-CAT CTT CTG GTT CGT TC-3'
HMGB3	F	5'-GAC CCC AAG AAA CCA AAG-3'
	R	5'- CAG GAC CCT TTG CAC CAT C-3'
HMGN1	F	5'-GAT GCC CAA GAG GAA GG-3'
	R	5'-GCT TCA TCA GAG GCT GG-3'
HMGN2	F	5'-ATG CCC AAG AGA AAG G-3'
	R	5'-GGC ATC TCC AGC ACC TTC-3'
HMGN3a/b	F	5'-GTC TCC AGA GAA TAC AGA G-3'
	R	5'-GAT TCA GTT TTC TGT GCC-3'
HMGN4	F	5'-GAC GGC CAC GAG ACT C-3'
	R	5'-CTG GAG TGT AGA GGC ATC-3'

Table II.1 Primer sequences for RT-PCR analysis on HMGs.

The reaction conditions were as follows:

Reverse transcription	50ºC	30 min	
Initial denaturation	95ºC	5 min	
Denaturation	94ºC	30 sec	
Annealing	50ºC	30 sec	35 cycles
Extension	72ºC	1 min	
Final extension	72ºC	10 min	

RT-PCR products were loaded on 2% agarose gel for electrophoresis and stained with ethidium bromide.

II.4.2 Copy-DNA synthesis and quantitative PCR

c-DNA was synthesized from 500 ng of total RNA from each biological sample using the SuperScript First Strand Synthesis Kit for RT-PCR (Invitrogen), with oligodTs as the primer in a 20 µl reaction system according to the manufacturer's protocol.

1 µl of the final 21 µl reaction was then used as template for quantitative PCR using QuantiTech SYBR Green PCR Kit (QIAGEN) with the following primers:

hGAPDH	F	5' - CCC ACT CCT CCA CCT TTG AC - 3'
	R	5' - CTC CCC TCT TCA AGG GGT CT - 3'
HMGA2	F	5'-ATG AGC GCA CGC GGT GAG-3'
	R	5'- GGC TTC TTC TGA ACA ACT TG -3'

Reactions were run in triplicate on an ABI 7500 real-time PCR System (Applied Biosystem). PCR conditions were:

Initial denaturation	95ºC	10 min	
Denaturation	95ºC	15 sec	
Annealing	50ºC	30 sec	45 cycles
Extension	72ºC	1 min	

Plus the programmed dissociation step to check the specificity of the PCR products.

Results were then analyzed with the sequence detection software version 1.2.3 (Applied Biosystem), with the threshold set at 0.09, and the Ct values of HMGA2 amplification curve were normalized to corresponding GAPDH Ct values. The identity of PCR products was always analyzed by 2% agarose gel electrophoresis.

II.4.3 Immunoblotting

SDS-PAGE

15% SDS-PAGE gel was usually used in this study. The gel was casted as follows:

For resolving gel:

Component	Final concentration
30% Acrylamine:Bis (29:1)	15%
1.5 M Tris-Cl pH 8.8	0.375 M
10% SDS	1%
10% Ammonium Persulphate	1%
TEMED	0.05% (v/v)
H_2O	23% (v/v)

For stacking gel:

Component	Final concentration
30% Acrylamine:Bis (29:1)	5%
1.0 M Tris-Cl pH 6.8	0.125 M
10% SDS	1%
10% Ammonium Persulphate	1%
TEMED	0.05% (v/v)
H_2O	70% (v/v)

Harvested cell pellets were lysed in a NP40 buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl and 1% NP-40) with addition of 1% SDS. Lysates were sonicated on ice-cold water bath with output power of 5 W for 10 x 3 sec. Protein concentrations were determined with BCA protein Assay Kit (Pierce) according to the manufacturer's instruction with BSA as standard. Lysates equivalent to 15 μg of total protein or 1 x 10^5 cells were analyzed through 15% SDS-PAGE and subsequently transferred to PVDF membrane with 0.2 μm pore size (Bio-rad).

The blot was blocked with 5% nonfat milk in TBST for 1 hr shaking at room temperature and incubated with goat polyclonal antibody against HMGA2 (S15, Santa Cruz) at a dilution of 1:500. The blot was then washed by TBST for 3 x 10 min followed by the incubation with a secondary anti-goat HRPconjugated antibody (Santa Cruz) and revealed with Lumi-light Western blotting substrate (Roche). β-actin was probed as an internal standard using monoclonal antibodies (Sigma).

II.4.4 Immunostaining

hES2 cells were seeded on sterilized glass cover slips with MEFs before staining. Cells were treated to be arrested in metaphase with 10 ng/ml Karyo Max Colcemid (Invitrogen) for 19 hr. Cells were fixed with 4% paraformaldehyde in PBS, treated with 0.1% Triton X in PBS, and blocked with 5% nonfat milk, followed by incubation with anti-HMGA2 antibody (S15 Santa Cruz) at a dilution of 1:100. The glass slides were washed with PBS for 3 times before the bound antibodies were detected by a secondary anti-goat FITC-conjugated antibody (Santa Cruz) at a dilution of 1:50. At the same time of the secondary antibody incubation, cells were incubated with 4',6-diamidino-2-phenylindole dihydrochloride (Sigma, St. Louis) at 10 μg/ml. Cover slips

with cells were mounted onto glass slides with mounting media (Sigma) and analyzed through a LSM 510 META confocal microscope (Carl Zeiss).

II.4.5 Cloning of HMGA2-His *E.coli* expression vector

Restriction Digests and Ligation

Restriction digests were performed with restriction enzymes (REs) from New England Biolabs (NEB). In general, 1 µg of DNA was incubated with 1 µl of RE (10 u/µl) in 20 µl of buffered solution for 4 hrs, unless indicated otherwise. DNA products were then visualized via agarose gel electrophoresis (Biorad) and extracted via Qiagen's gel extraction kit, if needed. DNA ligation reactions were performed at a 3:1 stoichiometric ratio (insert to backbone) with 1 µl of T4 DNA Ligase (10 u/µl) in 10 µl of buffered solution (NEB) at 16 ºC overnight.

Agarose gel electrophoresis and DNA purification

DNA products and plasmids were separated via agarose gel electrophoresis (Biorad) and stained with ethidium bromide (Biorad) for 10min before visualization with Biorad's gel documentation system. Desired DNA fragments were extracted and purified via Qiagen's gel extraction kit. Qiagen's PCR purification kit was also used to extract solitary DNA products from PCR reaction mixtures, which were subsequently used for ligation.

Polymerase chain reaction (PCR)

PCR reactions were performed using a thermal cycler (ABI), and *Pfu* or *Taq* DNA polymerases with relevant buffers from Promega and Qiagen respectively, according to manufacturer's protocol. Primers used are described in the following individual sections.

PCR conditions were:

Initial denaturation	95ºC	10 min	
Denaturation	95ºC	30 sec	
Annealing	50ºC	30 sec	32 cycles
Extension	72ºC	1 min	
Final Extension	72ºC	10 min	

Bacteria work

Media:

Luria Broth (LB) containing 10 g of tryptone, 5 g of NaCl and 5 g of yeast extract in 1 liter sterile water was used for all bacteria culture in this project. LB agar plate was used for initial bacteria growth. 200 μg/ml Ampicillin was used for selection of bacteria carrying transformed vectors whenever necessary.

Competent cell preparation

DH5α *supE*44 Δ*lacU*169 (φ80 *lacZ* ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1 *E. coli* cells (Hanahan, 1983) were streaked onto LB agar plates and cultured overnight at 37°C. A single colony was then picked and inoculated into 50 ml LB media and

incubated at 300 rpm, 37°C until the OD_{600} value reached between 0.4 to 0.6. The cultures were harvested via centrifugation at 4000 g for 5 min at 4°C. The resultant pellet was resuspended in 30 ml of cold 0.1 M $MgCl_2$ (Sigma) and centrifuged again at 4000 g for 5 min at 4 °C. With the supernatant discarded, the pellet was re-suspended in 20 ml of 0.1M $CaCl_2$ (Sigma) and kept on ice for 30 min. The cells were collected via centrifugation at 4000 g for 5 min at 4 °C, re-suspended in 2 ml of cold 0.1 M $CaCl_2$/15% glycerol (Sigma) and aliquoted to 100 µl/tube to be stored at -80 °C. These bacteria were then ready for transformation.

Transformation of *E. coli* and plasmid DNA preparation

20 µl of DNA ligation mix was added to 100 µl of competent DH5α bacteria and incubated on ice for 10min. Heat shock was applied at 42°C for 2 min, and the bacteria were placed on ice again for 5 min. Recovery was performed by incubating at 230 rpm, 37°C for 50 min, with the addition of 900 µl of fresh LB media. The cell pellet was collected via centrifugation, and re-suspended in 100 µl of LB media and plated out onto LB amp+ plates. The plates were then incubated overnight at 37°C and colonies were picked, inoculated and expanded in LB media with 200 µg/ml ampicillin. Plasmid DNA was prepared from transformed *E. coli* cell pellets via the Qiagen miniprep kit for general molecular biology work, or the Qiagen endo-free maxi kit for cell culture work, performed according to manufacturer's protocols.

II.4.6 Production and purification of His-Tagged hHMGA2

hHMGA2 coding region was cloned from mRNA isolated from hES2 cells, and tagged with a sequence encoding a 6- histidine at the C-terminal by RT-PCR using primers:

HMGA2Nde1F, 5'-GGA ATT CCA TAT GAG CGC ACG CGG TGA G-3';

HMGA2HisEcoR1R, 5'-GGA ATT CTT ACT AGT GGT GGT GGT GGT GGT GAG CGC TGT CCT CTT CGG CAG ACT C-3'.

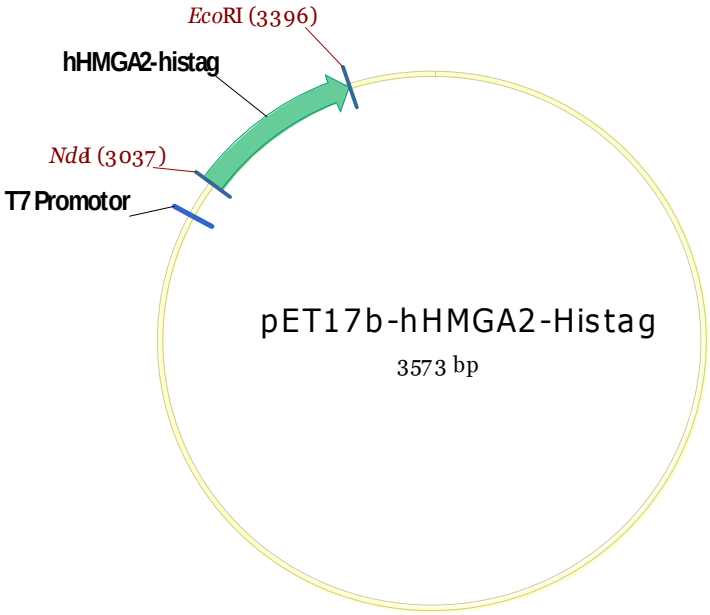


Figure II.2 Cloning vector map of pET-hHMGA2-His

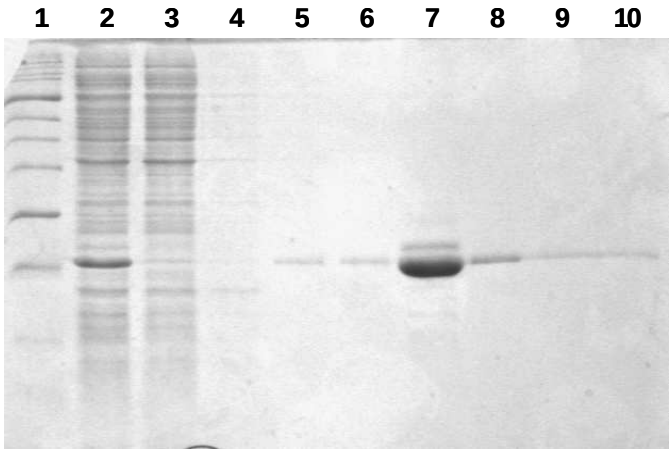
hHMGA2 coding region was cloned from mRNA isolated from hES2 cells, and tagged with a sequence encoding a 6- histidine at the C-terminal by RT-PCR, the fragment was then cloned into pET17b using NdeI and EcoRI, the sequence of which was included in the PCR primers. The cloning procedure was designed with Vector NTI.

pET17b-hHMGA2-His was constructed by inserting the hHMGA2 coding sequence, including the His tag, into pET17b using NdeI and EcoRI (NEB). All DNA constructs were sequenced (Figure II.2).

100 ng of the mini-prepared plasmid was transformed into bacteria line BL21-DE3 as described above without recovery step. Antibiotics (Ampicillin 200 µg/ml, tetracycline 15 µg/ml, chloramphenicole 34 µg/ml) were applied overnight to select positive bacteria cells bearing pET17b-hHMGA2-His plasmid. Culture was then expended in 1 litre LB medium containing the 3 antibiotics until OD_{600} reached 0.6–0.8. Isopropyl-β- D-thiogalactopyranoside (IPTG) was added at the final concentration of 1 mM to induce expression of hHMGA2-His.

After incubation for 4 hr, bacteria were harvested by centrifugation at 10,000 g for 15 min at 4ºC, and resuspended in lysis buffer consists of 50 mM NaH_2PO_4, 300 mM NaCl and 10 mM Imidazole, pH 8.0. Harsh Sonication, with output power of 10 watts and short breaks every 5 sec, was carried out to the sample in ice cold water baths for 10 min. The extract was then cleared by centrifugation at 16,100 g for 15 min at 4 °C.

The supernatant was mixed with 5 ml of Ni-NTA–agarose (Qiagen) and incubated overnight at 4 °C. The resin was packed into a small column, washed firstly with 2 x 5 ml of normal lysis buffer, secondly with 5 x 5 ml lysis buffer containing 50 mM imidazole. hHMGA2-His was then eluted with 6 x 5 ml of the lysis buffer containing 250 mM Imidazole. Each fraction of the eluted material was run on 15% SDS-PAGE gel and the gel was stained with coomassie-blue to determine the purity and amount of the desired protein in each fraction (Figure II.3).



Lane	
1	MW marker
2	lysate
3	flow through
4	wash with lysis buffer
5	wash with wash buffer
6	elution fraction 1
7	elution fraction 2
8	elution fraction 3
9	elution fraction 4
10	elution fraction 5

Figure II.3 Comassie-blue staining of SDS-PAGE gel with samples from each step of HMGA2 affinity purification through Ni^{+} column.

The protein band of HMGA2-His present in lysate was not present in the flow through, indicating the beeds were in excess for a complete binding of the protein. The majority of the protein was eluted in elution fraction 2, with an extra band above it, indicating the affinity binding was not specific enough. However, this band disappears in the reanalysis after the fraction undergone dialysis (Figure II.4).

The fraction contained the highest amount and purity of the protein was then subjected to dialysis overnight to eliminate imidazole. After dialysis against 50 mM NaH_2PO_4 (pH 8.0) and 300 mM NaCl, hHMGA2-His was dialysed again against 50 mM Tris-Cl pH 8.0 for desalting. Sample after dialysis was again run on 15% SDS-PAGE gel for purity determination before stored at $4^{\circ}C$ (Figure II.4). Protein concentrations were determined using the BCA assay kit (Pierce Biotechnology) with BSA as standard.

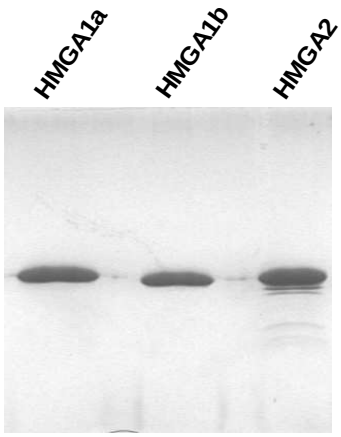


Figure II.4 Comassie-blue staining on SDS-PAGE gel with samples from finally purified HMGA proteins after dialysis.

The affinity-purified proteins were undergone 2-step dialysis to remove imidazole and later all other salt in the elution buffer. In most cases, no obvious extra band could be detected after this step. In the above picture, the extra protein bands were most likely from the degradation of HMGA2-His.

II.4.7 Nucleosome core particle (NCP)-HMGA2 binding assays

NCP was assembled from recombinant Xenopus laevis histones and a 147 bp palindromic DNA fragment as described in Davey et al., 2002. Complexes were prepared by mixing NCP or 147 bp DNA with the appropriate stoichiometry of purified HMGA2 protein, in a buffer containing 20 mM Tris (pH 7.5) and 1 mM EDTA, followed by incubation at 20°C for 1 hr prior to band shift analysis. Complex formation was analyzed via EMSA using non-denaturing gels containing 5% acrylamide/0.083% bis-acrylamide in a running buffer of 20 mM Tris, 20 mM boric acid, and 0.5 mM EDTA (pH 8.3). Gels were subjected to pre-running at 4 °C under a constant voltage of 100 V. Samples containing a constant amount of 150 ng of DNA and 10% (w/v) sucrose were loaded on the gel. Gel was run at 100 V with continuous

buffer recirculation and was stained with ethidium bromide followed by Coomassie Blue for visualization of DNA and protein, respectively (Experiment done by Dr. Curtis Davey and Dileep Vasudevan).

II.5 Functional Study of HMGA2 in hESCs

II.5.1 siRNA Transfection

HUES-8 cells were seeded at density of 1 x 10^5 in 24-well plate on MEFs at day 0, and incubated with 0.5 ml of antibiotic-free media containing 200 nM HMGA2- specific siRNA (Dharmacon) plus 4 μl of DharmaFECT2 transfection reagent (Dharmacon) for 24 hrs. To achieve a significant siRNA down-regulation rate, a 3-day consecutive transfection activity was carried out for each experiment. The cells were harvested on day 3, and processed for qRT-PCR and Western blot analysis. The same protocol was applied for controls using transfection reagent alone (denoted as mock) or together with 200 nM non-target (nt)-siRNA (Dharmacon).

II.5.2 Real-time RT-PCR

Following listed primers were used and detailed protocol was described at section II.4.2.

β-actin	F	5'-CAC CTT CTA CAA TGA GCT GCG-3'
	R	5'- TGC TTG CTG ATC CAC ATC TGC -3';
hGAPDH	F	5' - CCC ACT CCT CCA CCT TTG AC - 3'
	R	5' - CTC CCC TCT TCA AGG GGT CT - 3'
HMGA2	F	5'-ATG AGC GCA CGC GGT GAG-3'
	R	5'- GGC TTC TTC TGA ACA ACT TG -3'
hNANOG	F	5' - AGT CCC AAA GGC AAA CAA CCC ACT TC - 3'
	R	5' - ATC TGC TGG AGG CTG AGG TAT TTC TGT CTC - 3'
hOCT3/4	F	5' - CTT GCT GCA GAA GTG GGT GGA GGA A - 3'
	R	5' - CTG CAG TGT GGG TTT CGG GCA - 3'
hSOX2	F	5' - ATG CAG CGC TAC GAC GTG A - 3'
	R	5' - CTT TTG CAC CCC TCC CAT TT – 3'
hUTF1	F	5' - AGC AGA TCC GGA AGC TCA TGG G - 3'
	R	5' - TCC TCG GGG ATG CAG GTG - 3'
hIGF2	F	5' - TCC TCC CTG GAC AAT CAG AC - 3'
	R	5' - AGA AGC ACC AGC ATC GAC TT - 3'
hHAND1	F	5' -TGC CTG AGA AAG AGA ACC AG - 3'
	R	5' - ATG GCA GGA TGA ACA AAC AC - 3'
hGATA4	F	5' - CTC CTT CAG GCA GTG AGA GC - 3'
	R	5' - GAG ATG CAG TGT GCT CGT GC - 3'
hSOX17	F	5' - GGC GCA GCA GAA TCC AGA - 3'
	R	5' - CCA CGA CTT GCC CAG CAT - 3'
hSOX9	F	5' - GAA GCT CGC GGA CCA GTA - 3'
	R	5' - CGT TCT TCA CCG ACT TCC TC - 3'
hNestin	F	5' - CGT TGG AAC AGA GGT TGG AG - 3'
	R	5' - GAG CGA TCT GGC TCT GTA GG - 3'
hMIXL1	F	5' - CCG AGT CCA GGA TCC AGG TA - 3'
	R	5' - CTC TGA CGC CGA GAC TTG G – 3'

Table II.2 Primer Sequences for qRT-PCR analysis of hESC and differentiation markers.

II.5.3 Microarray analysis

Mock-, non-target (nt)-siRNA- and HMGA2-specific siRNA-treated hESCs were subjected to the PIQOR Stem Cell Microarray platform (Miltenyi Biotech GmbH, Germany; http://www.miltenyibiotec.com), using, in each case, untreated hESCs as reference. Total RNA was isolated via standard RNA extraction protocols (Macherey&Nagel, NucleoSpin RNA). RNA samples were quality checked using the Agilent 2100 Bioanalyzer platform (Agilent Technologies). RNA from each sample (0.5 µg) was then used for linear amplification, and subsequently checked again via the Agilent 2100 platform. For each of the three sample pairs, i.e. mock against untreated hESCs, (nt)-siRNA against untreated hESCs, and siRNA against untreated hESCs, duplicate microarray analysis was conducted.

The slides were scanned using a ScanArray^{TM} Lite laser scanner (PerkinElmer), and images were analyzed using ImaGene$^{@}$ software (BioDiscovery Inc.). The two-channel raw signal intensity was normalized using the *Mean Centering* method (Smyth and Speed, 2003), and the normalized Cy5/Cy3 ratios of the four replicates per gene were calculated for each of the six slides. Spots flagged as “low quality” by the image analysis software were eliminated.

Using the replicated spots for each gene, we performed a student-T test to determine the significance of the differences between the average ratio of spots on the siRNA slides and those on mock and (nt)-siRNA slides. The list of genes presented in Table III.1 was then selected based on the following stringent criteria.

(1) The average Cy5/Cy3 ratio of a gene on the siRNA slides must deviate by at least 40% from the corresponding one on the mock and (nt)-siRNA slides.

(2) Second, the P-value of the T-test must be less than 0.05.

We wrote a PERL script to conduct the above procedures for selecting these genes. In addition, we used the *TreeView* software [http://rana.lbl.gov/EisenSoftware.htm] to generate heat map bars displaying the average ratio for each gene on each of the six slides (Microarray analysis was done by Dr. Li Jinming). For each of the genes that passed our criteria, we visually checked the corresponding heat map bars to confirm that the average ratios are consistent between the duplicated slides, and that the average ratios are substantially different between siRNA slides and the two control slides. Inconsistent ones were eliminated, and this ultimately led to the list of genes presented in Table III.1. The complete list of two-channel raw signal intensities plus Cy5/Cy3 ratios, the heat maps for selected genes, the proportion of changes between average ratios on control and test slides, and the P-values can be retrieved from **Supplementary Data** at: http://vhp.ntu.edu.sg/hmga2/pd.htm.

II.5.4 Proliferation assay

hESCs were seeded at density of 3 x 10^4 in MEF-layered 96 well plates at day 0. On day 1, cells were incubated with 0.2 ml of antibiotic-free media containing 200 nM HMGA2- specific siRNA plus 1.6 μ1 of DharmaFECT2 reagent (Dharmacon Corp.) for 24 h. This was repeated on day 2 and 3. WST-1 (Roche) was added to the media at a final concentration of 10% on day 1–4, and plates were incubated at 37°C for 30 min. At 24 h intervals, absorbance (440 nm) was recorded by a standard photoplate reader, using media (cell-free) as reference. The assay was performed in triplicates with (nt)-siRNA as controls.

This colorimetric assay is based on the cleavage of the tetrazolium salt WST-1 to formazan by mitochondrial dehydrogenases found in viable cells. Expansion in the number of viable cells results in an increase in the overall activity of the mitochondrial dehydrogenases in the sample, thus increasing the amount of formazan dye produced by these cells which can then be quantified by measuring the absorbance of the dye solution.

The mean values of each sample were plotted against the days of reading, and comparison was done between si-NT and siHMGA2 transfected cells.

II.5.5 Immunostaining in EB

HUES7 cells were cultured for EB formation as described in section II.3.3. At 21st day after formation, EBs were collected by centrifugation at 600 g for 1 min, washed once with PBS and trypsinized to form single cells. After neutralization, cells were again washed with PBS, and fixed with 500 µl of 2% Paraformaldehyde/PBS for 20 min, washed with 3 ml of PBS, and permeabilized with 0.1% Triton X-100/PBS for 5 min. After another 3 ml of PBS was added to wash the cells, 5 x 10^5 cells were resuspended in 500 µl of 3% BSA/PBS for blocking for 30 min. Cells were then incubated with goat polyclonal antibody against HMGA2 (S15, Santa Cruz) at the dilution of 1:200 in 500 µl blocking buffer for 1.5 hrs. After 3 washes with PBS, cells were then incubated with secondary anti-goat FITCconjugated antibody (Santa Cruz) at a dilution of 1:50 in 500 µl blocking buffer for 30 min. After another 3 washes, cells were analyzed with flow cytometry.

A minimum of 10,000 total events were recorded per sample and analyzed using CELLQuest software (Becton Dickson). Briefly, dot plots of side scatter (SSC, indicative of cell granularity) versus forward scatter (FSC, indicative of cell size) were constructed. This was used to gate live, intact cells (R1) from cell debris at the bottom left and cell aggregates at the top right of the dot plot**.** The percentage of FITC-positive cells was then elucidated by looking at FL1 levels within all R1-gated cells, plotted against cell size (FSC). HUES7 cells incubated with only secondary antibodies were used as negative controls, and cells incubated with both primary and secondary antibodies were used as positive controls.

II.6 Functional Study of HMGA2 as an AP/dRP lyase

II.6.1 *in vitro* lyase assays

DNA cleavage assays employing a 5 kb supercoiled plasmid (pCMV-Int) were performed with either full length HMGA2 or a peptide comprising AT-hook 3 (KRPRGRPRK). A 50 μl reaction contained either 9.18 nM abasic or normal DNA substrate (control) was incubated in a buffer, made up of 25 mM Tris-HCl, pH 7.5; 2 mM EDTA; 100 mM $NaCNBH_3$, with 4.92 μM HMGA2 at 37°C for 40 min. NaCl (100 mM) was used in place of $NaCNBH_3$ in controls. Reactions were quenched by 2% SDS and heated at 65^{o}C for 15 min. QIAGEN QIAquick® PCR Purification Kit was used to purify the samples and eluted in 25 μl TE buffer (pH 7.4) containing 0.5% SDS each. Cleavage was analyzed through agarose gel electrophoresis. To detect HMGA2-DNA complexes, DNA was subsequently transferred to a Nylon membrane. The presence of HMGA2 was detected by immunoblotting using anti-HMGA2 polyclonal antibodies (S15, Santa Cruz) at a 1:500 dilution, and bovine anti-goat HRP-conjugated secondary antibodies (Santa Cruz) at a 1:5,000 dilution.

II.6.2 Generation of HMGA2 cell lines

The hHMGA2 c-DNA was cloned from mRNA, isolated from hES2 cells (Li et al, 2006), with a histidine affinity tag affixed at the C-terminus by RT-PCR using primers:

HMGA2Kpn1F 5'- CGG GGT ACC ATG AGC GCA CGC GGT GAG GGC -3'

HMGA2HisEcoR1R 5'- GGA ATT CTT ACT AGT GGT GGT GGT GGT GGT GAG CGC TGT CCT CTT CGG CAG ACT C -3.

The expression vector pEF1-hHMGA2-His-Neo was constructed by inserting the hHMGA2 coding sequence into pEF1-mycA-Neo (Invitrogen) using *Kpn*I and *EcoR*I (NEB) (Figure II.5). DNA constructs were sequenced.

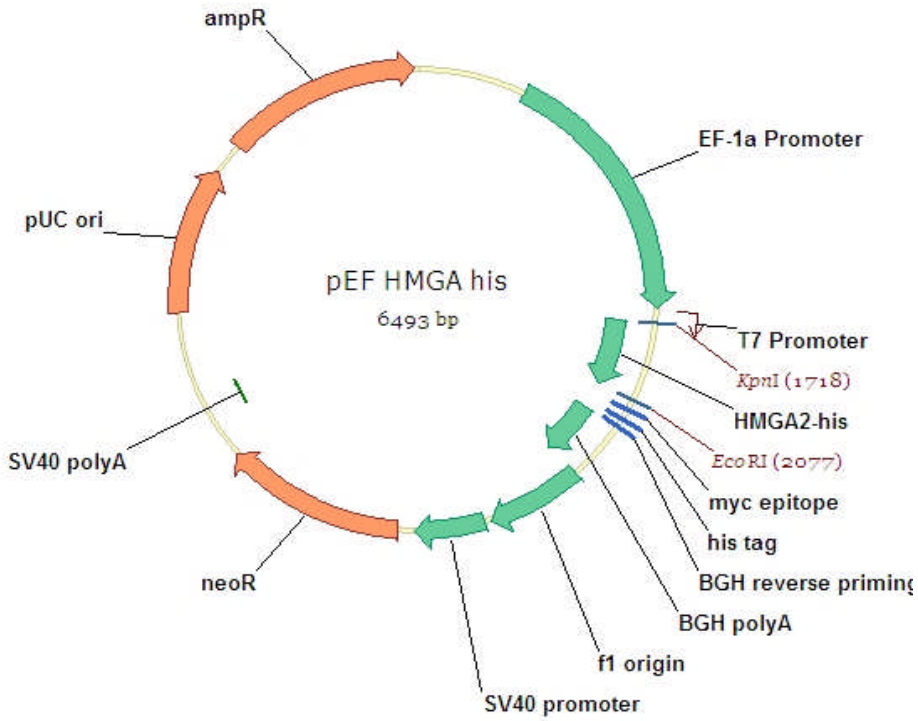


Figure II.5 Cloning map of pEF2-hHMGA2-His-Neo.

hHMGA2-His coding region was cloned by PCR using pET17b-HMGA2-His plasmid as template, with sequences of KpnI and EcoRI sites included in the primers. The fragment was then cloned into pEF1-MycA-Neo vector using KpnI and EcoRI. The cloning procedure was designed with Vector NTI. *Stop codons were designed in HMGA2-His sequence, so the His tag originally in the vector was not expressed.

pEF1-hHMGA2-His-Neo was transfected into HeLa (ATCC No. CCL-2) cells using lipofectamine (Invitrogen) and into A549 (ATCC No. CCL-185) cells via electroporation.

Electroporation:

A549 cells were passaged the day before electroporation. 1 x 10^7 cells per transfection experiment were collected, neutralized, and re-suspended in 400 µl of RPMI 1640 (Gibco) containing 30 µg of plasmid pEF1-HMGA2-His. This DNA/cell suspension was subjected to an exponential pulse (300 V, 950 µF, 300 ὃ) in a 2 mm cuvette using a GenePulser (Biorad). Floating cell debris was then removed and the remaining cell solution was transferred into a 10 cm dish with culture media and allowed to recover under normal culture conditions for 24 hrs. Fresh media was changed then and selection pressure was applied 48 hrs later.

Lipofectamine:

HeLa cells were seeded on a 10 cm cell culture dish and allowed to grow until about 90% confluent. 20 µg of plamid pEF1-HMGA2-His was diluted with 1.5 ml of OPTI-MEM (Gibco) media while 60 µl of Lipofectamine 2000 (Invitrogen) was dissolved in another 1.5 ml OPTI-MEM media. 5 min later, the two were mixed and allowed to be incubated at room temperature for at least 20 min (not more than 1 hr) before being added to the cells in 12 ml of DMEM media. Fresh media was changed after 6 hrs from the addition of the transfection mixture, and cells were allowed to grow for 48 hrs under normal condition before selection pressure was applied.

Neomycin resistant cells were selected with 500 µg/ml of G418. Individual surviving colonies were picked manually into 96-well plates using micro-pipette under reverse phase light microscope (Carl Zeise). Cells from different colonies were expanded and tested for HMGA2 expression by Western blotting analysis. Transfected cell lines A549 1.3, 1.5, 1.6 and HeLa P2, P8, P19 were chosen for following experients.

II.6.3 Trapping of HMGA2 on genomic DNA

For the *in vitro* trapping of HMGA2 on genomic DNA isolated from cells under low pH challenge, A549 cells were first treated on culture dishes with PBS (pH2), or physiological PBS as control, for 10 min at 37 ˚C, followed by a wash with physiological PBS. Cells were harvested in the same buffer and genomic DNA was purified with DNeasy Blood & Tissue Kit (Qiagen). 3 μg of HMGA2 was incubated with 1 μg of DNA in a buffer containing 25 mM Tris-HCl, pH 7.5; 2 mM EDTA and 100 mM $NaCNBH_3$ at 37°C for 40 min. DNA was purified with QIAquick® PCR Purification Kit (Qiagen) and spotted onto PVDF membrane for HMGA2 detection by immunoblotting.

For *in vivo* trapping, adherent transfected 1.6 cells were first treated with PBS at pH 2 for 30 min at 37˚C, followed by a wash with physiologic PBS containing 100 mM $NaCNBH_3$ (Sigma), or 100 mM NaCl as a control. Cells were harvested in the same buffers and genomic DNA was purified with DNeasy Blood & Tissue Kit (Qiagen). Purification involved SDS-based lysis with a short protease treatment for 10 min at 70°C. $NaCNBH_3$ or NaCl was present during elution, and the same amount of genomic DNA was spotted onto a PVDF membrane for HMGA2 detection.

II.6.4 Cell survival assays

II.6.4.1 Annexin-V-PI staining

Low pH

To challenge cells with low pH, A549 and its transfected cell lines were treated with PBS (pH2) at 37°C for 6 min. HeLa and the respective transfected cell lines were challenged at pH3. Cells were allowed to recover under normal culture conditions for 24 hrs before harvest.

Hydroxyurea

Cells were treated with 100 mM Hydroxyurea (Sigma) in normal culture media under culture conditions for 48 hrs before harvest.

Paclitaxel

A549 cell lines were treated with 50 nM paclitaxol (Sigma) in normal culture media for 48 hrs before harvest, while HeLa cell lines were treated with 20 nM paclitaxel for the same period.

Cisplatin

A549 cell lines were treated with 20 μM cisplatin (Sigma) in normal culture media for 48 hrs before harvest, and HeLa cell lines were exposed to 20 μM cisplatin for 24 hrs.

MMS

A549 cell lines were treated with various concentration of MMS (Sigma) in normal

culture media for 1 hr followed by a recovery period for 48 hrs before harvest. In some experiments, we added 5 mM BA to the cell culture media throughout the recovery period. HeLa cell lines were treated with 4 mM MMS.

To test for survivability, cells were trypsinized and harvested, including those from the culture media and wash buffer. Cells were pelleted at room temperature by centrifugation at 600 g for 3 min, washed once with PBS, and stained with Annexin-V fluorescence and Propidium Iodide (PI) using Annexin-V-FLUOS Staining Kit (Roche) according to the manufacturer's protocol.

Flow cytometry

Cells were analyzed by FACS Calibur (Becton Dickson). A minimum of 10,000 total events were recorded per sample and analyzed using CELLQuest software (Becton Dickson). Briefly, dot plots of side scatter (SSC, indicative of cell granularity) versus forward scatter (FSC, indicative of cell size) were constructed. This was used to gate live, intact cells (R1) from cell debris at the bottom left, and cell aggregates at the top right of the dot plot**.** The percentage of PI-positive cells was then elucidated by looking at FL2 levels within all R1-gated cells, plotted against FL1 levels which elucidated Annexin-V-positive cells. Cells with negative staining for both Annexin-V and PI were counted as live cells; those with positive staining for both markers were qualified as necrotic cells; and those with positive staining for Annexin-V, but negative staining for PI, counted as apoptotic cells. Non-stained cells were set as negative controls (Figure II.6).

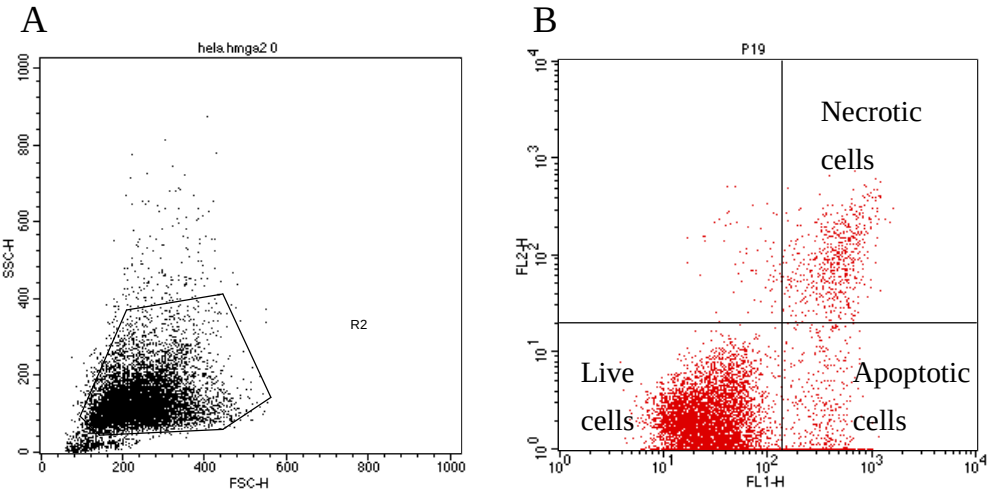


Figure II.6 Illustration of the analysis principle of FACS.

A. Cells acquired were plotted by its granularity against size, and intact cells were gated. B. Gated cells from A were ploted by the fluorescence given by PI against the fluorescence given by Annexin V staining. Cells population with negative staining for both Annexin-V and PI (lower left) were considered live; those with positive staining for both markers (upper right) were qualified as necrotic cells; and those with positive staining for Annexin-V, but negative staining for PI (lower right), counted as apoptotic cells.

II.6.4.2 Colony-Forming assay

2 x 10^5 A549 and 1.3 cells were treated with 2 mM MMS at 37°C for 1 hr and recovered overnight under normal culture conditions with or without 5 mM BA. Cells were trypsinized, and 500 cells of each sample were seeded in a 10 cm tissue culture dish and grown under normal culture conditions for 3 to 5 days without being disturbed until colonies could be recognized under the microscope. Each survival colony should represent a single actively dividing cell survived from the challenges whether or not applied previously. Colony numbers were counted under the microscope and compared among samples that were treated differently. Each

experiment was done in duplicates.

II.6.4.3 MMS challenge with transient HMGA2 transfection

A549 or HeLa cells were seeded at density of 5 x 10^4 in 24-well plate before 1 μg of pEF1-HMGA2-His plasmid was transfected. 48 hrs after transfection, cells were treated with various amount of MMS (0.5 and 1 mM for A549 cells; 0.2 and 0.5 mM for HeLa cells) for 1 hr. Cells were then recovered for 24 hrs before being trypsinized. 2000 cells from each sample were counted and plated in a 10 cm culture dish and left untouched until single cell colonies were formed and become large enough to count. CMV-EGFP cells were used as controls in these experiments, and transfection efficiency was determined before MMS treatment by flow cytometry. Each experiment was performed in triplicates.

Plasmids carrying mutated and truncated HMGA2 were also cloned by colleague Dr. Heike Summer, and these plasmids were then used for transient transfection into A549 cells in the same experiment, to compare with those carrying wild type HMGA2 (Figue II.7).

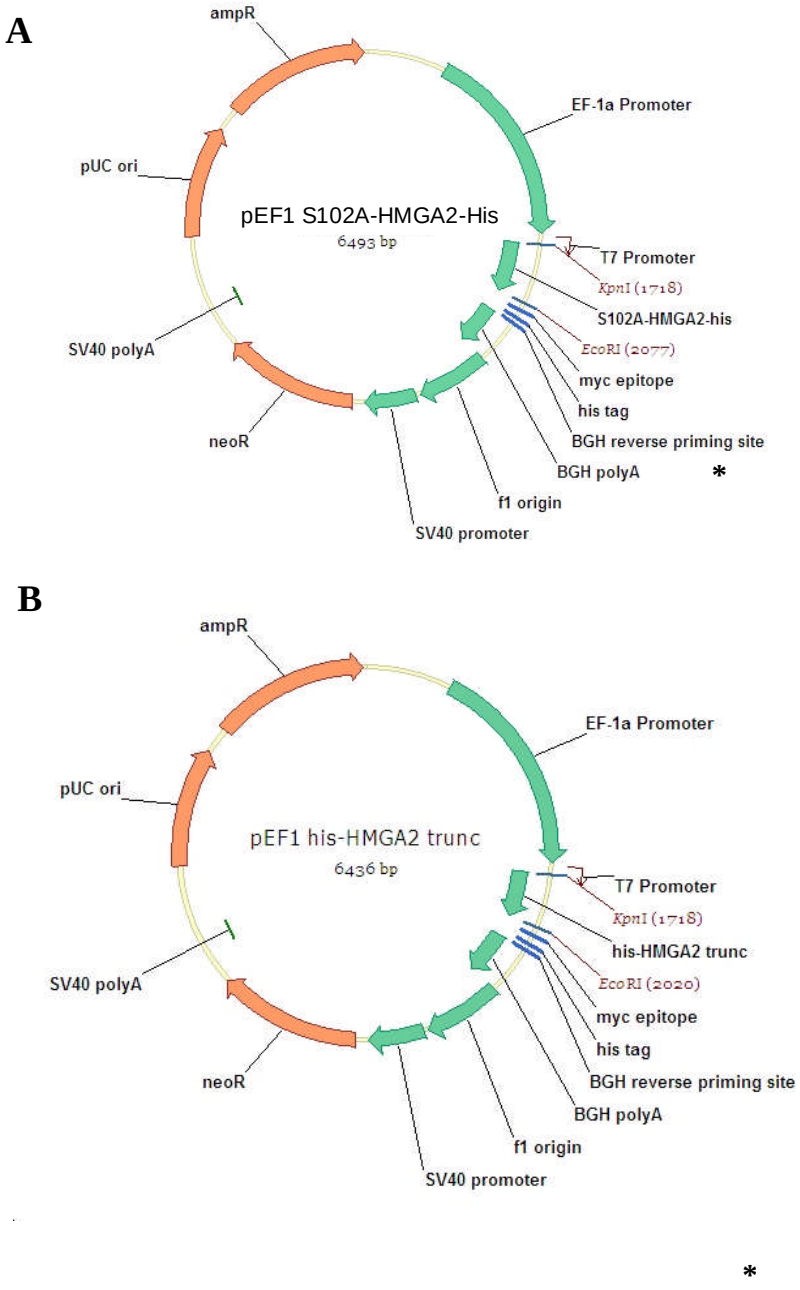


Figure II.7 Cloning maps of vector carrying A. mutated and B. truncated HMGA2.

Mutation and Truncation of HMGA2 coding region were achieved via PCR with respective primers. Fragments were the inserted into pEF1-MycA-Neo vector using KpnI and EcoRI. *Stop codons were designed in HMGA2-His sequence, so the His tag originally in the vector was not expressed.

II.6.5 Protein co-affinity precipitation

2 x 10^7 of each A549, 1.3, HeLa and P2 cells were collected and lysed with 0.5 ml native lysis buffer containing 50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 0.1% NP-40, 2 mM DTT, 5% Glycerol and protease inhibitors (Roche). Sonication was applied to each lysate sample at output power of 2 watts for 15 x 5 sec in iced water bath. Each sample was topped up to 1 ml with lysis buffer and centrifuged at 16,100 g for 10 min at room temperature, and the supernatant was incubated overnight with rolling at 4°C with 50 µl Ni+ sepharose beads (GE Healthcare) which was pre-incubated with lysis buffer containing 1% BSA. 50 µl of supernatant was kept as cell lysate controls. The beads were span down at 100 g for 2 min at room temperature and washed with lysis buffer for 4 x 10 min with rolling. Elution was done with 50 µl of buffer containing 20 mM NaH_2PO_4, 20 mM Na_2HPO_4, 500 mM NaCl and 500 mM Imidazole, pH 8.0.

The eluted material, together with the cell lysate controls were then run on SDS-PAGE and later subjected to western blotting analysis for HMGA2 and APE1. Detection of HMGA2 was done as described above, and detection of APE1 was using rabbit polyclonal antibody against APE1 (Santa Cruz) followed by HRP-conjugated goat anti rabbit secondary antibody.

II. 6.7 Construction of pEF1-HMGA2-His-IRES-EGFP-Neo vector

The vector was constructed by cutting off IRES-EGFP sequence from pCMV-IRES-EGFP-Puro, and inserting into pEF1-HMGA2-His-Neo using restriction enzyme EcoR1 and Xba1 (Figure II.8).

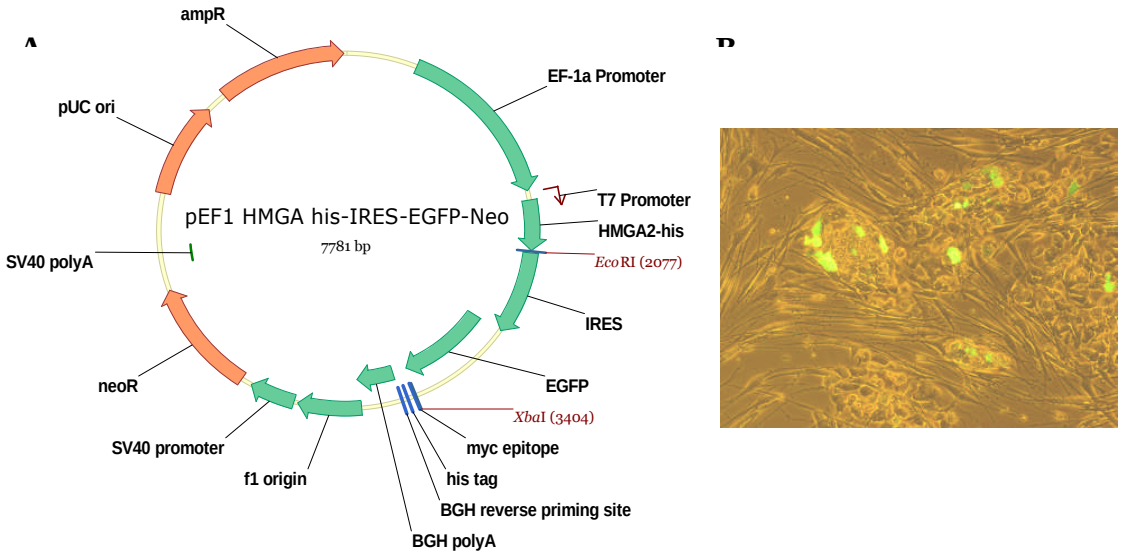


Figure II.8 Construction of pEF1-HMGA2-His-IRES-EGFP-Neo vector.

A. Cloning map of pEF1-HMGA2-His-IRES-EGFP-Neo vector. Vector was constructed by cutting off IRES-EGFP sequence from pCMV-IRES-EGFP-Puro, and inserting into pEF1-HMGA2-His-Neo using restriction enzyme EcoR1 and Xba1.
B. Fluorescent microscopic image on HUES8 cells 48 hrs after transfection with pEF1-HMGA2-His-IRES-EGFP-Neo vector (10X Magnification).

This plasmid can be a useful tool for generating HMGA2 overexpressing clones since the selection of successfully transfected cells can be done via fluorescent microscopy together with Neomycin treatment.

III RESULTS

III.1 Expression and Function of HMGA2 in hESCs

III.1.1 Pluripotency of hESCs

HUES7, HUES8 and hES2 cells were regularly checked for pluripotency when they were in culture. The most commonly used was Alkaline Phosphotase staining. Pluripotent hESCs would give purple color colonies whereas differentiated hESCs and MEFs appear colorless (Figure III.1)

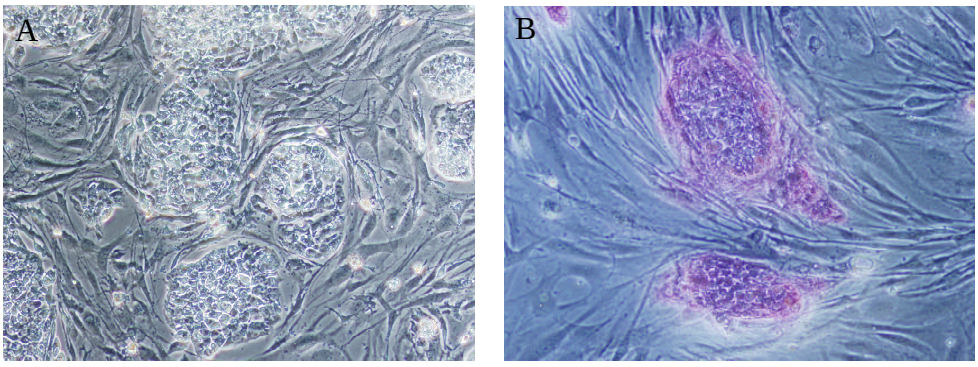


Figure III.1 AP staining on hESCs for pluripotency examination.

A. hES colonies on MEF; B. Pluripotent hESCs give purple colonies upon AP staining, whereas MEFs appeare colorless. Briefly, the cells were grown up to 50% in confluency before fixed with 90% Methanol/10% Formalin (Merck) for 1min, washed with PBS, and incubated with freshly prepared mixture of Napthol/Fast Red Violet/water at a 2:1:1 ratio for 15 min in the dark. Magnification 10x.

More stringent pluripotency examinations were carried out once a month on hESCs in culture. Immunostaining with both nuclear and surface pluripotency biomarkers were utilized (Figure III.2).

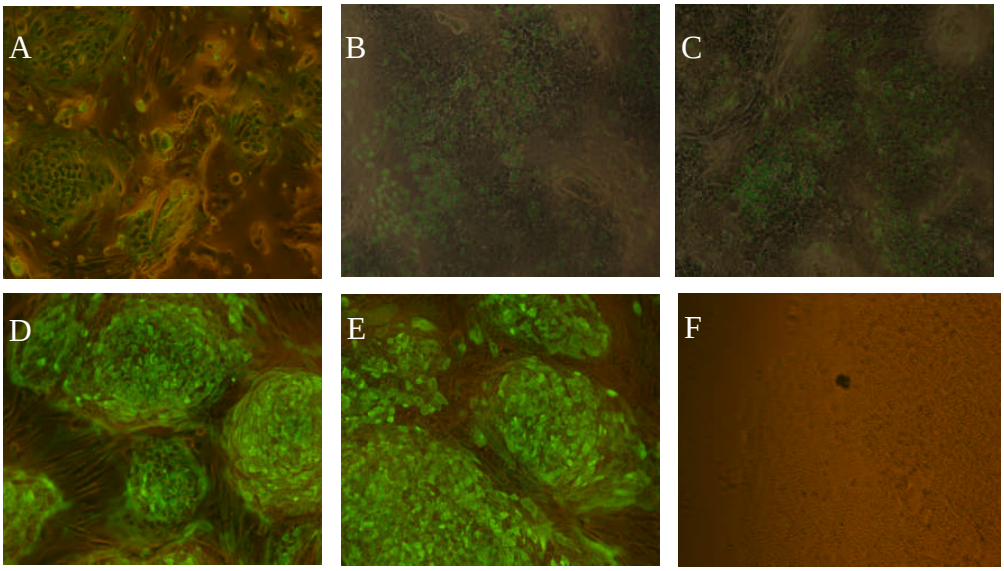


Figure III.2 Pluripotency examinations of HUES-8 cells after 2 weeks of culture by immunostaining (10X magnifications).

A. SSEA-4; B. Oct3/4; C. Sox2; D. TRA-1-60; E. TRA-1-81; F. SSEA-1 (negative control). Oct4 and Sox2 are two well established regulatory genes that highly expressed in hESCs. Also, pluripotent hESCs stain positively for surface markers TRA-1-60, TRA-1-81, SSEA-3 and SSEA-4, but negatively for SSEA-1 (Henderson et al., 2002; Carpenter et al., 2004). This profile is similar to human embryonic carcinoma and human ICM cells.

III.1.2 HMGA2 highly expressed specifically in hESCs

At mRNA level

The study on HMG proteins in hESCs was first started by analyzing mRNA levels for 9 core HMG genes including isoforms that result from alternative splicing. By using RT-PCR, the expression levels in hES cell line hES2 were compared with those in primary human foreskin cells (F3) and three cancer cell lines from different origin

(A549, HeLa, and MCF7). A careful inspection of product intensities, with β-actin taken as reference, indicated that mRNAs coding for HMGA2 and its isoform HMGA2b (v2) are present at higher levels in hES2 cells in comparison with other cell lines (Figure III.3).

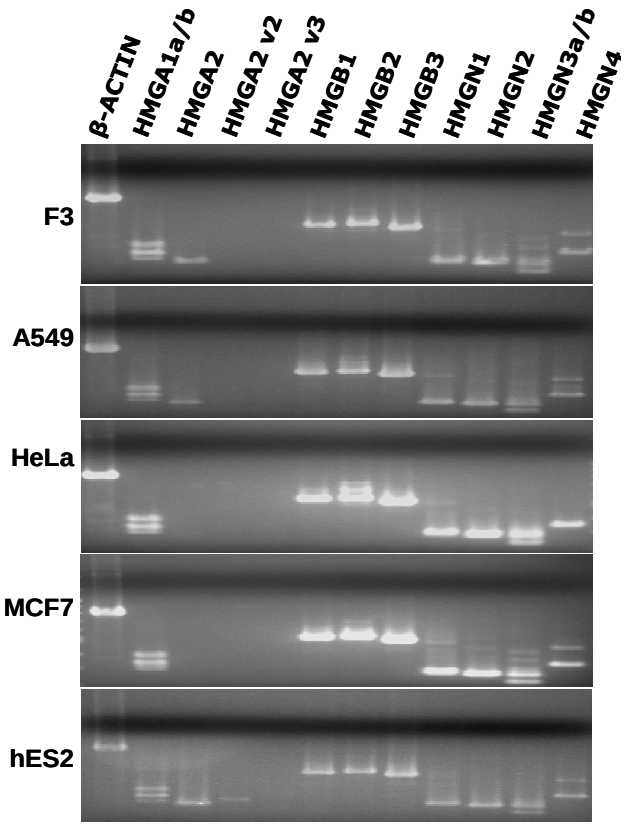


Figure III.3 Comparative analysis of HMG expression in hES2 cells and other human cell lines.

Equal amount of purified total RNA was reverse transcribed and subsequently amplified using specific primers as probes for 9 main HMG genes and three isoforms, as indicated. β-actin was served as control. PCR products were separated through agarose gel electrophoresis and visualized via ethidium bromide staining.

Note: HMGA2 v3 was later withdrawn from NCBI database.

At protein level

In order to verify and quantify the observed differences at the protein level, Western blotting analyse was performed to determine the amount of HMGA2 protein in whole cell extracts prepared from a fixed number of cells. With the HMGA2 signal from F3 cells set as reference (set as 1), it is found that the amount of HMGA2 protein is between 8 and 25 fold higher in hES2 cells than that in other cell lines (Figure III.4). We reproduced this result with a second hES cell line, HUES-8, obtained from a different source (data not shown). We could not directly detect the protein isoform HMGA2b in this analysis either because its molecular mass is only 0.4 kD smaller than HMGA2 and, thus, difficult to separate from HMGA2, or because its amount is below the detection limit.

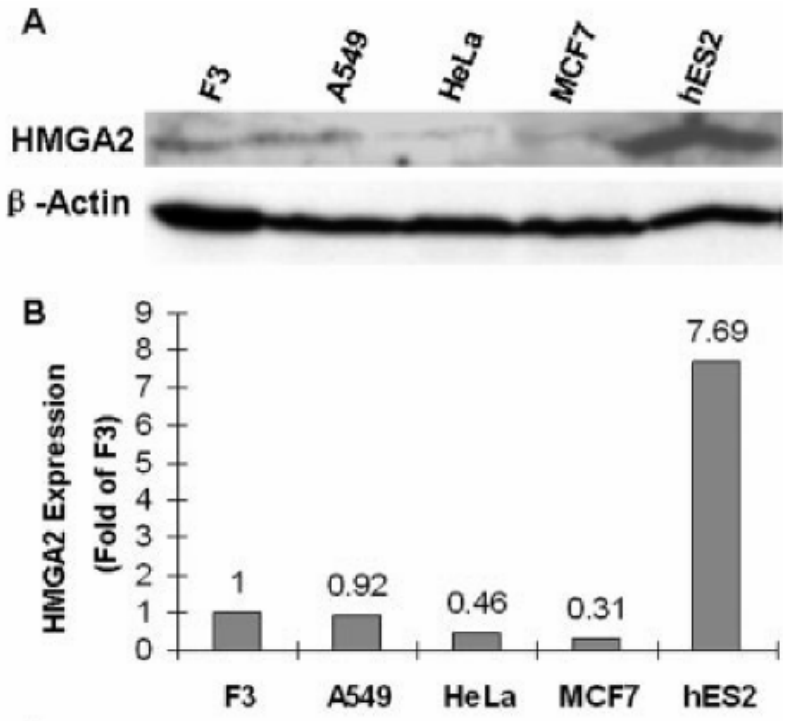


Figure III.4 Analysis of HMGA2 protein expression in different cell lines.

(A) Western blot analysis of HMGA2 expression using β-actin as internal reference.
(B) Quantification of Western analysis shown in (A), using the signal obtained with F3 human foreskin cell extract as reference (set as 1). Quantification was done using Quantity1 software (Biorad) with densitometry function.

Next, we were interested in obtaining a reliable estimate of the number of HMGA2 molecules present inside an individual hES2 cell. By using extract from a defined number of hES2 cells and known amount of recombinant human HMGA2 protein purified from *E. coli* as standard, our Western analysis revealed that at least 10^5 copies of HMGA2 must be present per stem cell (Figure III.5).

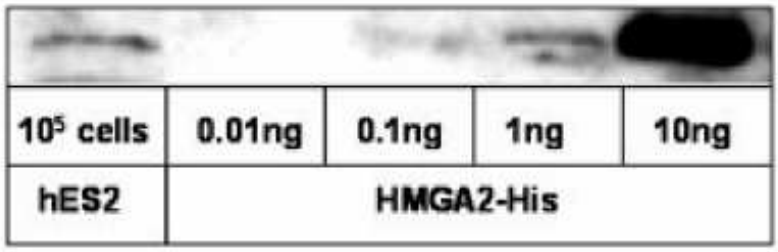


Figure III.5 Determination of HMGA2 copy number in hES2 cells using recombinant purified hHMGA2 protein as standard.

It can be seen that 10^5 cells contains similar amount of HMGA2 protein as 1 ng of HMGA2-His. The two bands were quantified and with the information that the molecular weights for HMGA2 and HMGA2-His are 11.7 and 12.8 respectively, it was calculated that at least 10^5 copies of HMGA2 must be present per stem cell.

III.1.2 Amount of HMGA2 protein further increased in early differentiating hESCs

Because HMGA2 is absent in most adult human tissues, it is interesting to examin whether a high HMGA2 level is maintained if hES2 cells are triggered to differentiate into embryoid bodies (EBs). Taking the cycle difference between β-actin and HMGA2 from F3 cells as reference, our ΔΔ analysis based on the quantitative RT-PCR data surprisingly revealed that the mRNA level of HMGA2 and its isoform HMGA2b is significantly increased at day 7 of EB formation and drops back to or below the level in undifferentiated cells at day 14 of EB formation (Figure III.6). The data also confirmed our earlier conclusion drawn from Figure III.3 that HMGA2 is higher expressed in undifferentiated hESCs compared with other cell lines that were also investigated here.

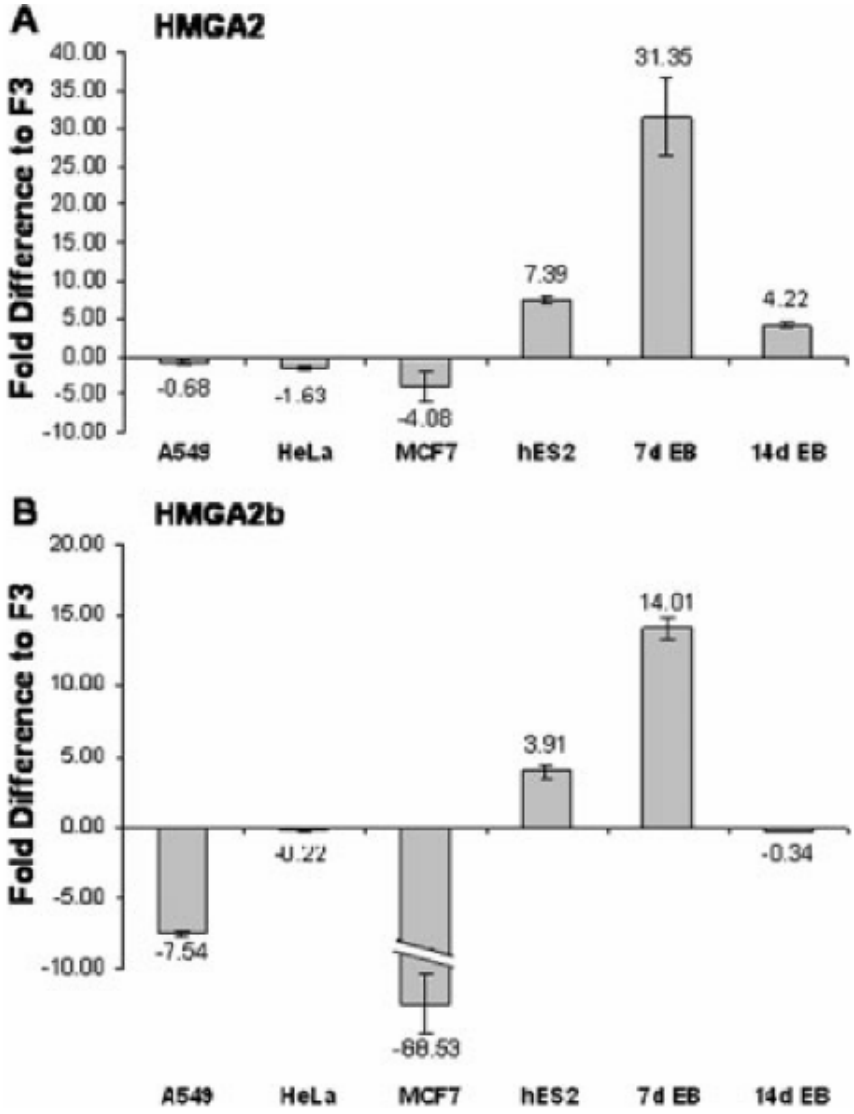


Figure III.6 Quantitative RT-PCR analysis of HMGA2 and HMGA2b expression in hES2 cells and EBs.

Equal amount of total RNA was used as template for reverse transcription with oligo dT as primers. Copy-DNA samples were then amplified in real-time PCR in triplicates, using primer pairs specific for HMGA2 or its isoform HMGA2b. The difference in cycle number between β-actin and HMGA2 (A) or its isoform HMGA2b (B) was determined and used to calculate the fold difference between cell samples, taking cell line F3 as reference.

The upregulation of HMGA2 upon EB formation was confirmed by Western analysis which showed that in 7 days old EBs, the amount of HMGA2 protein is about three-fold higher than that in undifferentiated hESCs (Figure III.7).

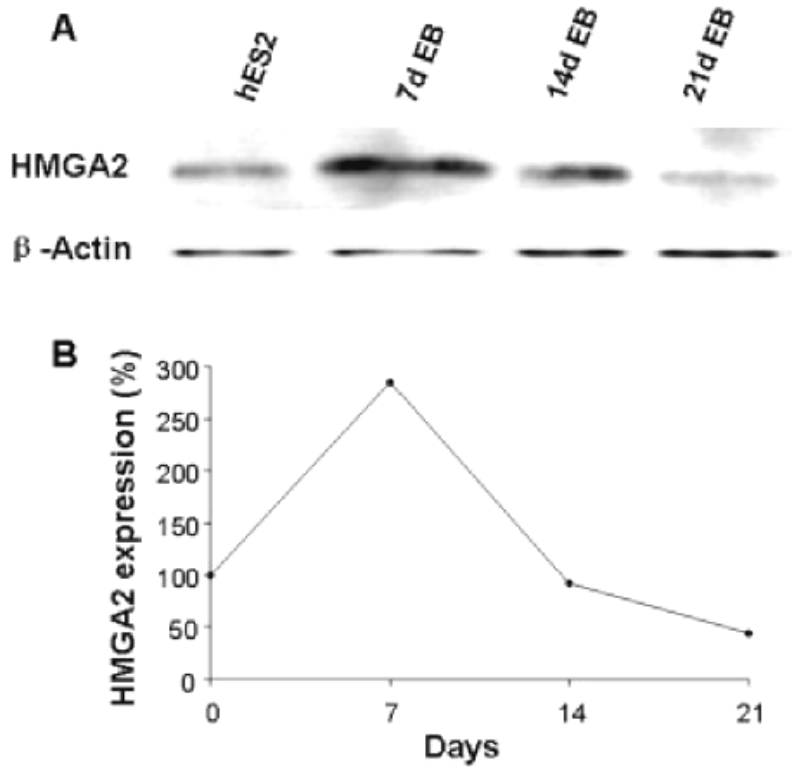


Figure III.7 Western analysis of HMGA2 expression in undifferentiated hES2 cells and during EB formation.

(A) Equal amount of cell extract prepared from the same number of either undifferentiated hES2 cells or EB cells at different time points of differentiation was analyzed by Western blotting using β-actin as internal standard. (B) Quantitative analysis of the result shown in (A), normalized with undifferentiated hES2 cells as 100%.

A very similar modulated expression pattern of HMGA2 during EB formation was also observed with cell line HUES8, except that the up-regulation appeares a few days later during differentiation as compared with hES2 cells (Figure III.8).

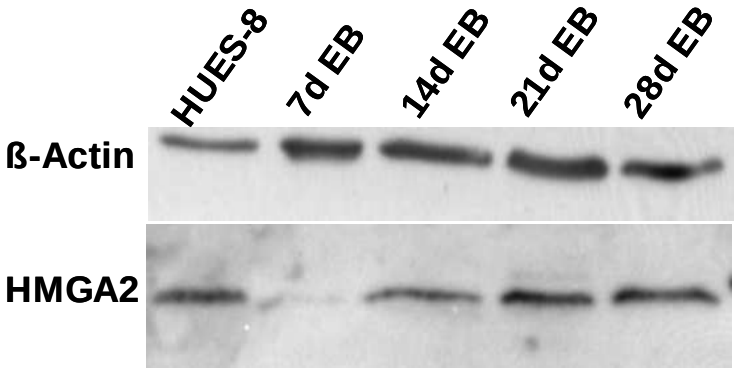
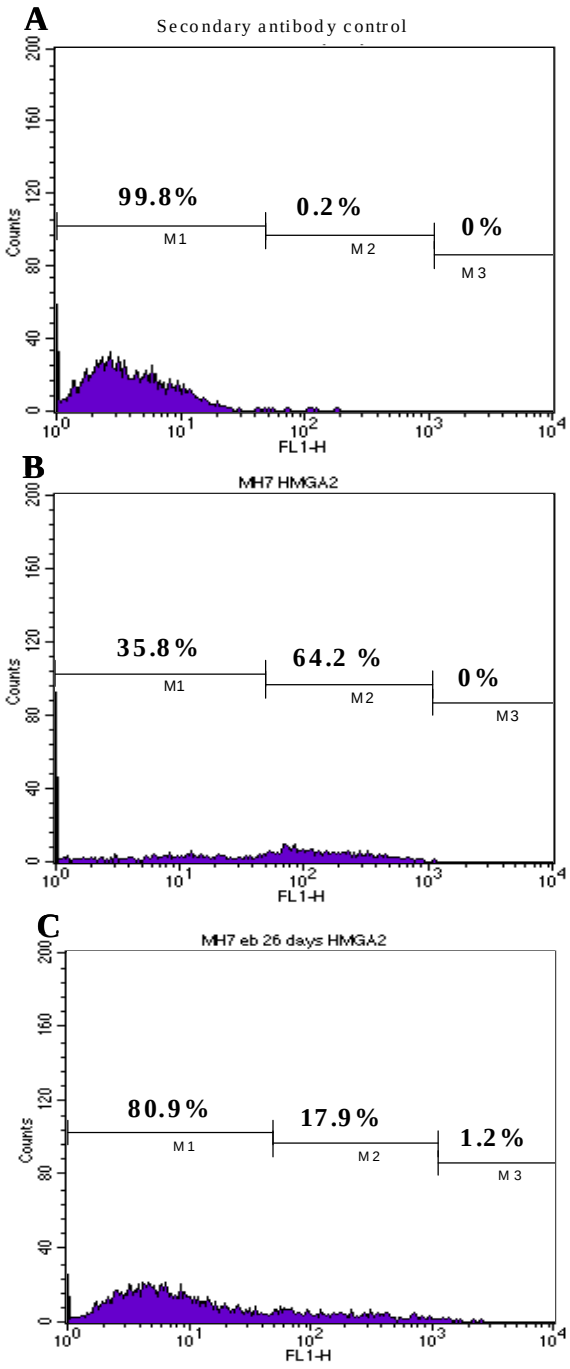


Figure III.8 Western analysis of HMGA2 expression in undifferentiated HUES8 cells and during EB formation.

Different from hES2 cells (Figure III.7), the level of HMGA2 expression in HUES8 cells was first dropped in early EBs and gradually increased until 28 days of EB culture.

III.1.3 HMGA2 expression upregulated in a subset of cell during *in vitro* embryogenesis

To analyze whether the upregulation of HMGA2 in early EB generation is associated with any specific cell differentiation, FACS of HMGA2 expressing cells were carried out on dissociated 21-day-EB generated from HUES7 cells. By comparing with the pluripotent HUES7 cells, it is obvious that HMGA2 is down-regulated in the majority of the cells in 21-day-EB, but a very small subset of cells did show a 10-time upregulated HMGA2 signal. It is very possible then during early embryogenesis, there is a specific cell type marked with very high HMGA2 expression arises transiently (Figure III.9).



Histogram analysis of cell populations with different HMGA2 signal strength in

A). HUES7 cell stained with only FITC conjugated secondary antibody as controls;

B). HUES7 cells cultured on MEFs;

C). Cells dissociated from 7-day-EB generated from HUES7 cells.

HMGA2 is down-regulated in the majority of the cells in 21-day-EB, but a very small subset of cells did show a 10-time upregulated HMGA2 signal Note: the negative HMGA2 staining in HUES7 cells was possibly resulted from MEF cells.

Figure III.9 FACS analysis of different HMGA2 expression in cells of early EB.

III.1.4 Association of HMGA2 protein with hESC chromatin *in vivo* and nucleosome core particle *in vitro*

Our analysis so far revealed that about 10^5 HMGA2 molecules are present inside an individual undifferentiated hESC. Further, in cells from day 7 EBs, this number increased at least three-fold. Given its known nanomolar DNA-binding affinity to AT-rich sequences as small as 4 bp (Cui et al., 2005; Reeves and Nissen, 1990), this finding implies that the majority of HMGA2 molecules should be bound to hESC genomic DNA unless nucleosomal organization severely interferes with HMGA2-binding.

Therefore it will be interesting to investigate the distribution of HMGA2 proteins inside hESCs. We found almost exclusive nuclear HMGA2 immunostaining and an association with both inter and metaphase hESC chromatin as evidenced by the significant overlay with DAPI staining (Figure III.10). Since DAPI has a strong binding preference for A-T bps, the vast majority of HMGA2 molecules inside hESCs seem to be in a tight association with AT-rich chromatin regions during different stages of cell cycle.

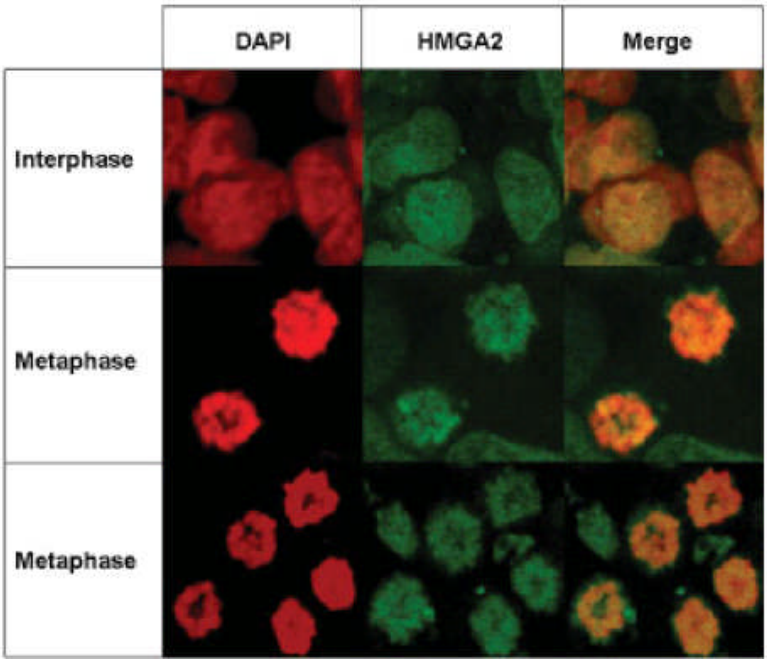


Figure III.10 HMGA2 immunostaining of hESCs at different phases during cell cycle.

hES2 cells were either fixed directly for HMGA2 immunostaining (interphase) and co-staining with DAPI, or first treated with colcemid to arrest cells in metaphase. Confocal images were taken at 100-fold magnification. Note: The cells remain green (HMGA2 positive) and fail to show orange DAPI/HMGA2 overlay in both metaphase merged pictures were cells that were not trapped in metaphase. Therefore, these cells have a much lower DAPI staining intensity compared with the ones in metaphase. In this case then, the contrast setting of the DAPI staining picture had gated these cells out, as shown in the red DAPI staining pictures. However, this did not apply to the HMGA2 staining pictures as the differences in staining intensity between inter- and meta-phase cells were not as great as that of the DAPI staining, thus the interphase cells remained green failed to show orange in the overlay picture.

These results prompted us to probe the interaction of HMGA2 with a nucleosome core particle (NCP), for which the 1.9 Å resolution crystal structure has been solved recently (Davey et al., 2002). The palindromic 147 bp DNA contains 8 AT-tracts of at least 4 nucleotides, of which only 4 appear to be accessible for AT-hook association in the NCP.

Reconstituted NCP was incubated with increasing amount of purified human HMGA2 protein, and resulting complexes were analyzed through electrophoretic mobility shift assay (EMSA). The results showed that NCP becomes saturated with HMGA2 at a molar ratio of 12:1. Further, staining of polyacrylamide gels with Coomassie Blue indicated that, even at the highest molar ratios, HMGA2-binding to NCP does not lead to a significant displacement of histone octamer (Appendix I Figure 6 A, B). This is particularly evident at lower molar ratios between 1:1 and 4:1 when discrete HMGA2-NCP complexes which still contain histones are detectable. In parallel, DNA–HMGA2 interactions were examined using the same DNA segment but not associated with histones. In this case, saturation with HMGA2 was still not reached at a molar ratio of 14:1 (Appendix I Figure 6 C, D). Together, these data showed that more than one HMGA2 molecule can stably associate with a single NCP *in vitro* without displacing histone octamer and, as evidenced by immunostaining, possibly also *in vivo* through AT-tract binding in hESC chromatin. This experiment was done in collaboration with Dr. Dileep Vasudevan and Dr. Curtis Davey, SBS, NTU.h

III.1.5 siRNA -mediated down-regulation of HMGA2 in hESCs

Next, siRNA technology was employed in this study to probe into HMGA2's biological role(s) in pluripotent hESCs and verified first that transfection of HMGA2-specific siRNA significantly reduces the level of HMGA2 protein. With Western blotting analyses, it could be seen that repeated siRNA transfections over a period of 72 h reproducibly led to a 4 to 5 fold decrease compared with untransfected, mock transfected, or non-target (nt)-siRNA transfected hESCs (Figure III.11). These results are consistent with the efficacy of siRNA-mediated knock-down seen in qRT-PCR analyse where the level of HMGA2 mRNA is reduced about 7 fold (Figure III. 12).

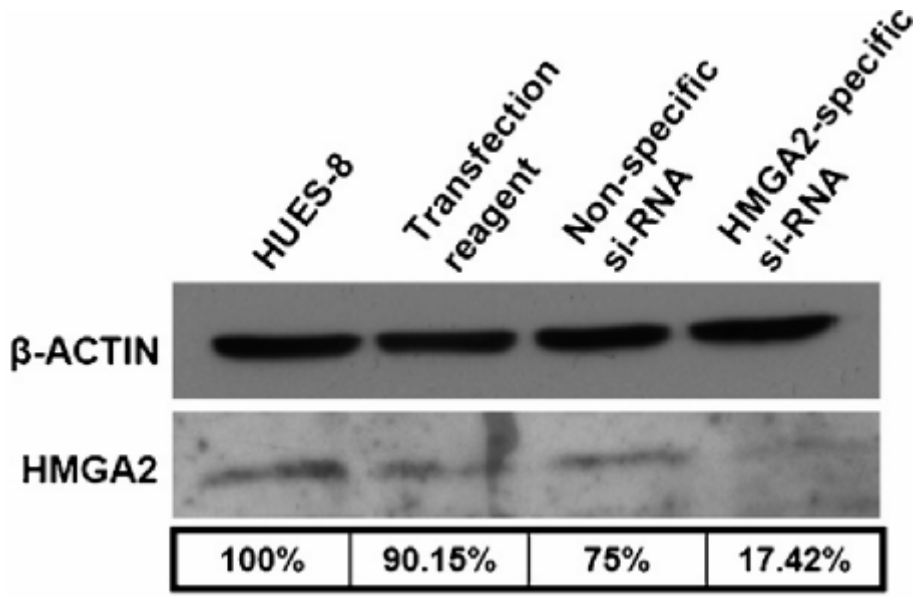


Figure III.11 Western analysis of HMGA2 protein expression in hESCs upon siRNA treatment.

Quantification of the Western analysis was obtained with normalization with HMGA2 expression in untreated HUES-8 cells as 100%. And β-actin was used as internal ref erence.

III.1.6 HMGA2 affects the expression of key pluripotency and differentiation markers

With the establishment of our siRNA protocol which worked to significantly reduce HMGA2 protein levels in HUES8 cells during a period of 72 h, qRT-PCR was used to determine changes in mRNA levels of a selected set of well-known key pluripotency and differentiation markers, using mock- and nt-siRNA-treated hESCs as reference. The results showed that the mRNA level of human undifferentiated transcription factor 1 (hUTF1) is significantly reduced, while the transcript levels for other key pluripotency genes OCT3/4, NANOG, and SOX2 remain nearly unaffected (Figure III. 12). UTF1 is one of the most sensitive pluripotency markers known for ESCs and is quickly down-regulated at the onset of differentiation (Trounson. 2006). It encodes for a transcription factor that interacts with activating transcription factor 2 (ATF2); however, the downstream target genes have not been identified. Interestingly, UTF1 was recently found to be involved in the control of mESCs proliferation (Nishimoto et al., 2005). Hence, our results suggest a link between HMGA2, cell proliferation control, and pluripotency in hESCs (see below).

mRNA levels of well-known markers for all 3 germ layers were also analyzed and it was found that down-regulation of HMGA2 reduces the transcript level for the mesoderm marker gene insulin like growth factor 2 (IGF2) and for the neural stem cell specific marker NESTIN. In contrast, knock-down of HMGA2 significantly increased the levels of transcripts coding for endoderm markers SOX17 and MIXL1 (Figure III.12). These findings suggest that HMGA2 plays a role in maintaining a balance of mesodermal versus endodermal cell differentiation in hESCs. Interestingly, in addition to IGF2's role in mesodermal cell lineage commitment, it has also been associated

with human fetal growth (Kaku et al., 2007), which, again, links HMGA2 with cell proliferation during embryogenesis.

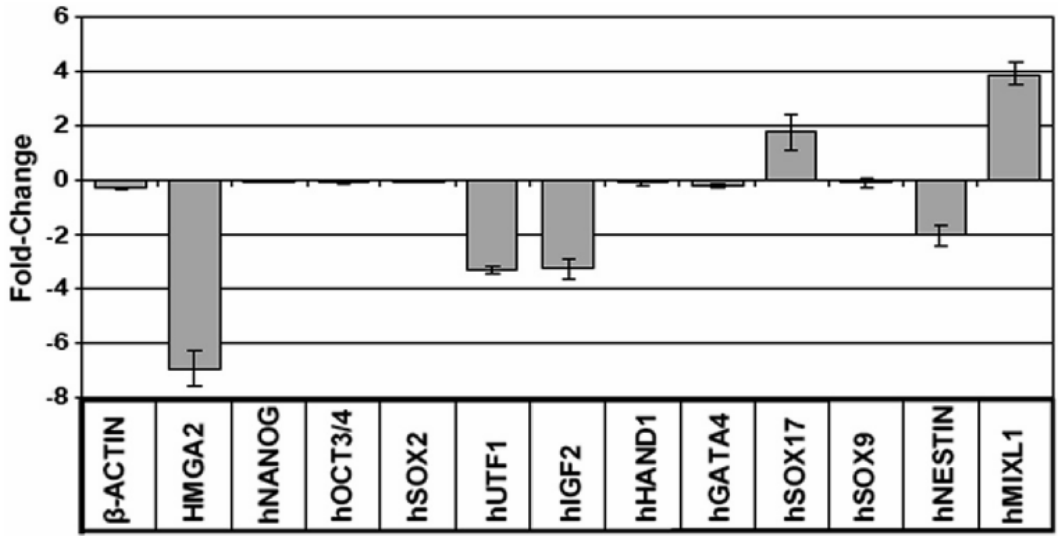


Figure III.12 ΔΔ analysis of qRT-PCR test on the expression levels of puripotency and germ-line markers upon siRNA treatment.

Each reaction was carried out in triplicate and the value was obtained by comparing ΔCt in HMGA2-specific siRNA transfected cells to the average ΔCt in both mock and (nt)-siRNA transfected cells using hGAPDH as internal reference. Error bar represented the standard deviation of the triplicate.

III.1.7 HMGA2, a co-regulator of genes linked to mesenchymal cell differentiation and adipogenesis in hESCs

Next, the PIQOR™ Stem Cell Microarray platform, which contains anti-sense probes for 930 selected stem cell-specific genes, was employed in combination with siRNA. By applying stringent selection criteria (see Section 2.5 for details), our data analysis revealed that 14 genes, including HMGA2, are differentially expressed at a significant level. These genes were listed in Table III.1. Importantly, the results confirmed again that siRNA treatment significantly decreases the level of our target HMGA2 mRNA.

Among the remaining genes listed in Table III.1, 3 code for proteins well known to play an important role in embryogenesis. Expression of LEFTYgenes, LEFTYA and LEFTYB, is increased. These genes are transiently expressed during early hESCs differentiation in EBs. They are known to be controlled by the NODAL signaling pathway, and their expression appears to play a substantial role in mesodermal cell lineage establishment (Dvash et al., 2007). They are also involved in the left-right body symmetry formation and control of gastrulation. Interestingly, both genes are located on human chromosome 1 as direct repeats, and separated by only 50 kb. It is tempting to speculate, therefore, that HMGA2 as DNA architectural factor might be a constituent of the corresponding chromatin domain that harbors the control regions for LEFTY genes and, as such, contributes to their regulation in hESCs. The expression of the third gene, semaphorin 6A-1, is decreased upon siRNA treatment. This gene belongs to a group of genes coding for proteins that are guidance signals in embryogenesis (Klostermann et al., 2000).

It is probably noteworthy that another prominent cell differentiation marker that did not pass our stringent selection criteria, but still ranks at position 53 in our list of 930 genes, is DLK1. Its expression increases (1.22 fold change) as a result of HMGA2 knockdown. The encoded protein, also known as preadipocyte factor 1, contains EGF-like repeats and, intriguingly, is a known inhibitor of adipocyte differentiation (Smas and Sul, 1993). Hence, it is possible that HMGA2's role in adipogenesis includes that as a repressor of DLK1 expression. Other genes in the top 14 list (Table III.1) play roles in the formation of connective tissue (COL6A2, COL16A1) as transcriptional repressor during cell differentiation (NPM1), gap junction formation (GAJ1), and trophoblast differentiation (CALB1).

Gene.ID	NAME	ACC	UniGene	FC	adj.P.Val	B
25938	HMGA2	P52926	Hs.2726	0.472	0	13.795
29198	LEFTB_HUMAN	O75610	Hs.278239	1.745	0	5.698
20559	SEMA6A1	Q9P2H9	Hs.263395	0.735	0.011	2.315
29197	LEFTA_HUMAN	O75611	Hs.25195	1.495	0.015	1.682
2035	CALB1	P05937	Hs.65425	0.69	0.015	1.502
2277	COL16A1	Q07092	Hs.26208	1.565	0.015	1.274
15918	POU5F1	Q9BZV9	Hs.2860	0.759	0.015	1.236
812	GJA1_1	P17302	Hs.74471	0.674	0.021	0.813
32488	NPM1	P06748	Hs.411098	0.796	0.023	0.61
2309	COL6A2_1	P12110	Hs.420269	1.414	0.024	0.466
7873	ODC1	P11926	Hs.443409	0.73	0.024	0.391
12690	S100A11	P31949	Hs.417004	1.259	0.035	-0.02
2344	TIMP1	P01033	Hs.522632	1.453	0.041	-0.256
23248	SOX2	P48431	Hs.816	0.773	0.047	-0.444

Table III.1 Microarray statistics on genes with transcriptional expression level changed upon HMGA2 downregulation

The table columns are: ID: Gene ID; NAME: Gene Name; ACC: SwissProt/TrEMBL accession number; UniGene: UniGene Accession number; FC: Expression fold change between siRNA and mock/ntsiRNA; adj. P: adjusted P-value of empirical Bayes moderated t-test (corrected for multiple testing using Benjamini–Hochberg method); B: B-statistics of empirical Bayes moderated t-test.

Our microarray analysis is largely consistent with our qRT-PCR data. Two genes, IGF2 and NESTIN, which showed differential expression based on the qRT-PCR analysis (Figure III.12) did not appear in our list of selected genes. In the case of IGF2, the array signals were inconclusive and flagged. NESTIN also appears down-regulated in the microarray analysis (0.81 fold change) consistent with the qRT-PCR data. However, it ranked at position 30 in the list due to its corrected P-value of 0.16. Interestingly, the prominent pluripotency markers Oct3/4 and Sox2 appeared at positions 7 and 14, respectively, in our list and are down-regulated (0.76 fold change). While our qRT-PCR analysis also revealed a slight decrease in expression level, it was found not to be statistically significant (Figure III.12). Hence, whether HMGA2 is directly involved in the down-regulation of these two genes or whether this is an indirect effect perhaps due to the induction of cell differentiation remains to be investigated.

III.1.8 HMGA2 regulates hESCs proliferation

Our results showed that HMGA2 is involved in the regulation of UTF1 (Figure III.12) which is known to be critical for ESCs proliferation in mouse (Nishimoto et al., 2005). It has also been demonstrated that mouse embryonic fibroblasts derived from HMGA2 null mice exhibit a significantly slower proliferation rate than wild-type cells (Zhou et al., 1995). Therefore, in order to analyze cell proliferation directly in hESCs, we employed again siRNA technology followed by a colorimetric assay to two hES cell lines, HUES-7 and HUES-8. We transfected hESCs with siRNAs as before and measured cell proliferation using nt-siRNA as control. The results (Figure III.13) showed that after three days of repeated exposure to siRNA, cell proliferation in both cell lines is significantly reduced compared with the control. Using qRT-PCR again,

we confirmed that mRNA levels of both HMGA2 and UTF1 are reduced due to siRNA treatment (data not shown).

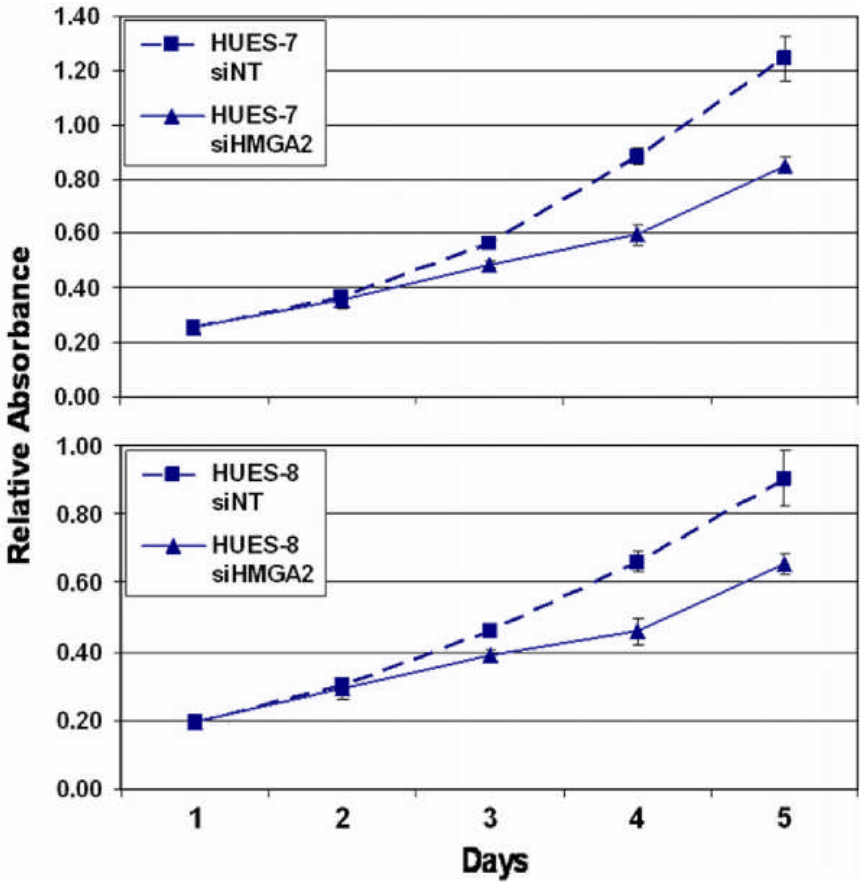


Figure III.13 hESC proliferation affected by HMGA2 down-regulation.

Cells from two hES cell lines (HUES-7 and HUES-8) were seeded at day 0. Twenty-four hours later (day 1), cells were transfected with siRNAs. This was repeated in 24 h intervals (days 2 and 3). Colormetric reagent WST-1 was added to the media and absorbance (440 nm) was recorded each day, using media as reference. The assay was performed in triplicates with transfected (nt)-siRNA as control.

Unfortunately, UTF1 gene was not included in the PIQOR^{TM} Stem Cell Microarray (only contains less than a thousand genes) which did not further strengthen link among HMGA2, UTF1 and hESC proliferation directly. However, there are some other genes which involved in cell proliferation seemed affected on the microarray which

strengthened our proliferation assay.

For example, S100A11, which appeared in our top 14 listed genes, got up-regulated upon HMGA2 siRNA treatment, and studies have suggested that S100A11 was an essential element in the signal transduction for high Ca++- and transforming growth factor (TGF)β-induced suppression of the growth of normal human keratinocytes (Sakaguchi et al., 2003 , 2004 , 2005).

Other than that, the expression level of many proteins involved in cell proliferation signaling (FGF1, FGFR2, VGR2, PKC A/M/N, MAPK6 and AKT) seems down-regulated for 10% to 30%, although the high variation between biological duplicates did not make these in our list of selected genes (analysis from supplementary data of Appendix II).

III.2 Functional Role of HMGA2 as an AP/dRP lyase

III.2.1 HMGA2 has an intrinsic AP lyase activity

The initial screen for AP lyase activities employed supercoiled plasmids which contained an undefined number of AP sites that resulted from depurination during exposure to acidic pH. Incubation with HMGA2 converted the majority of supercoiled AP plasmids into the open circular form, while control plasmids remained supercoiled (Figure III.14A, done by Mr Zhan Lihong from the lab**).** If the observed cleavage of AP site plasmids is due to an associated AP lyase activity, Schiff base intermediates should have been generated by nucleophilic attack of a primary amine group to the aldehyde form of the AP site. Therefore, we tested whether covalent DNA-HMGA2 complexes could be trapped by the reducing agent NaCNBH3 (Verdine and Norman, 2003). HMGA2 was incubated either with AP site plasmids or control substrates under trapping or non-trapping conditions, and reactions were stopped with SDS. After agarose gel electrophoresis, the DNA was transferred to a nylon membrane. Western blotting then revealed that HMGA2 could be trapped during incubation only on AP plasmids, and the lyase could be detected on both the supercoiled and relaxed DNA form, as expected for substrates with multiple AP sites (Appendix III, Figure 1B**).**

Since the three independent AT-hooks of HMGA2 penetrate deeply into the minor DNA grove and are nearly identical in amino acid sequence (Huth et al., 1997), we hypothesized that the primary amine(s) responsible for the lyase activity likely reside within DNA-binding domains. We employed a short peptide comprising the third AT-hook of HMGA2 and incubated it with AP site or control plasmids. As observed with the full length protein, the peptide cleaved only AP site plasmids, albeit with reduced efficiency when compared with HMGA2 (Figure III.14A**).** These results indicate that the nucleophile(s) that attack the C' 1 of deoxyribose at AP sites could reside within AT-hooks.

Since the first two AT-hooks are identical in HMGA1 and HMGA2 proteins, we tested whether the two main HMGA1 isoforms generated by alternative splicing, i.e. HMGA1a and HMGA1b, also exhibit AP lyase activities that can be trapped by NaCNBH3. The results clearly demonstrate that this is the case (Figure III.14B**).**

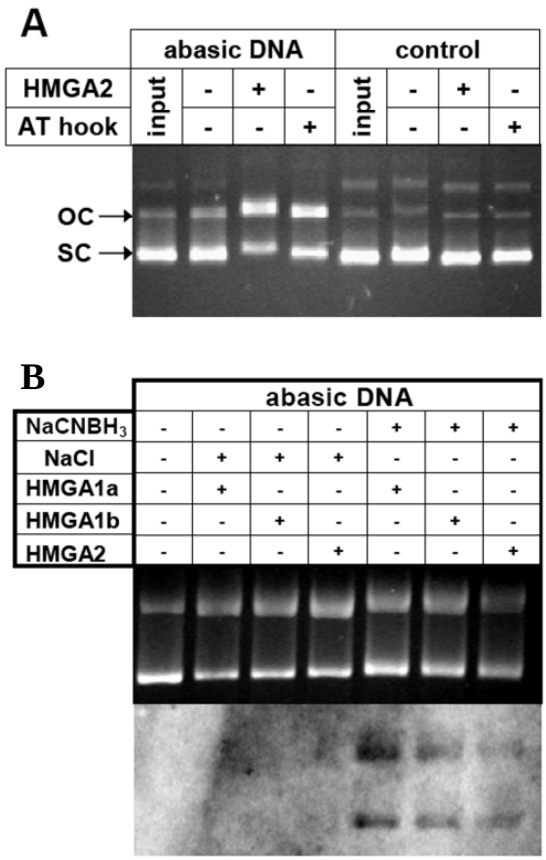


Figure III.14 HMGA2 acts as an AP lyase activity on abasic supercoiled DNA.

(A) Cleavage assay with full length HMGA2 or a peptide comprising the 3rd AT hook with abasic supercoiled (sc) DNA. AP site cleavage converts the topology of sc DNA into the open circular (oc) form. Note that the depurinated input DNA is mostly sc and does not change during incubation without HMGA2 (Done by Mr. Zhan Lihong from the lab). (B) Cleavage and *in vitro* trapping assay with HMGA1a, HMGA1b and HMGA2 on abasic sc DNA. Proteins were detected with anti Penta-His antibody (Qiagen).

Furthermore, data obtained with mass spectrometry on trapped HMGA2-DNA complexe strengthen our conclusion that the primary amine(s) responsible for the lyase activity reside in the AT hook (Appendix III, Supplementary Figure 1**).**

Our finding that HMGA1 and HMGA2 can be covalently trapped on AP substrates by NaCNBH3 indicated that the AP lyase-associated cleavage reaction employ β elimination (Verdine and Norman, 2003). Hence, in order to determine the chemical nature of DNA ends generated by HMGAs and the efficiency of AP site cleavage, we employed a ^{32}P labeled double-stranded oligonucleotide which contains a single AP site as substrate (AppendixIII, Figure 2A).

The results show that incubation of purified HMGAs and a peptide of HMGA2 AT-hook 3 with the substrate generates cleavage products which exhibit the same electrophoretic mobility as those produced by endonuclease III, an AP lyase derived from *E. coli* (Appendix III, Figure 2C, lanes 2-6). In addition, products generated by HMGA2 differs from those obtained with *E. coli* enzymes AP endonuclease IV, which produces 3'OH ends, and from those obtained with formamidopyrimidine N-glycosylase (Fpg), which generates 3' phosphates during AP site cleavage (Appendix III, Figure 2C, lanes 7 and 12, respectively). Importantly, when we subsequently reacted ^{32}P-labeled products generated by HMGA proteins with AP endonuclease IV, their mobility was increased. This indicates that the 4-hydroxy-2-pentenal moiety generated by HMGA's AP lyase activity had been removed from the 3' end by endo IV (Appendix III, Figure 2C, compare lanes 3-6 with lane 8- 11). Incubation with increasing protein/DNA ratios and quantification of AP site cleavage furthermore established that all 3 HMGA proteins had similar AP lyase activities which was enhanced compared to the isolated AT-hook (Appendix III, Figure 2D).

III.2.2 HMGA2 is also a dRP lyase

Having established that HMGA2 is an AP lyase, we next examined whether the protein also possesses the related 5'- deoxyribosyl phosphate (dRP) lyase activity. To test this possibility, we employed the same 31-mer oligonucleotide containing a single abasic site. However, the substrate is 3'-labeled and preincubated with endonuclease IV to produce a 5'-dRP moiety at the labeled strand (Appendix III, Figure 3A). In order to stabilize the chemically labile deoxyribosyl phosphate group and increase separation of cleaved from non-cleaved dRP lyase products during electrophoresis, the dRP moiety was reacted with O-benzyl hydroxylamine (BA). The presence of BA adducts leads to retardation during electrophoresis, compared with products that lost the dRP moiety due to a dRP lyase activity (Appendix III, Figure 3B, lanes 3-4). Using this assay, we established that HMGA proteins and the AT-hook 3 peptide efficiently remove the dRP moiety and function, therefore, as a dRP lyase (Appendix III, Figure 3B, lanes 5 to 12). As a control, we tested a linker peptide connecting HMGA2 hooks 1 and 2, and found that it is unable to cleave AP/dRP substrates (data not shown).

Once a genomic AP site is generated, the next step in BER involves AP endonucleases and/or AP/dRP lyases which cleave the DNA backbone at the base lesion. However, the tautomeric open-ring form of deoxyribose produced by DNA glycosylases can react very fast with BA (Appendix III, Figure 3B). It has been shown that BA adducts inhibit DNA cleavage by mammalian AP endonucleases and AP lyases (Liu and Gerson, 2004), and we confirmed that 5 mM BA completely blocks the dRP lyase activity of HMGA2 *in vitro* (Appendix III, Figure 3C).

III.2.3 HMGA2 can be covalently trapped at genomic abasic sites *in vivo*

After demonstrating that HMGA2 is an AP and a dRP lyase, we investigated whether these BER functions play a *bona fide* role in human cervical and lung cancer cells, which contain key oncogenic mutations in *p53* and *KRAS*, respectively. Here, HeLa and A549 cells stable transfected cell lines which constitutively over-expressed HMGA2 were employed. Western blotting analysis showed that HMGA2 falls below the detection limit in extracts prepared from parental cells, but is present in transfected cells. Since these were generated by random genome insertion of the expression vectors, we chose three transfected lines of each parental line for further analysis to eliminate the possibility that a particular phenotype is due to insertional mutagenesis (Figure III.15**).**

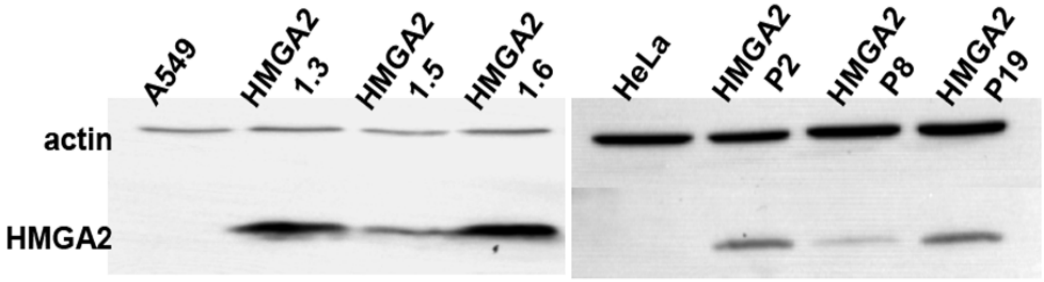


Figure III.15 Western analysis on HMGA2-transgenic cell lines.

HMGA2 transgenic cells (A549 derived 1.3, 1.5 and 1.6; and HeLa derived P2, P8 and P19) showed various expression levels, using β-actin as internal control. It was observed that the relatively low HMGA2 expressing transgenic cell lines (1.5 and P8) have similar HMGA2 expression level as that of hESCs (data not shown).

To investigate whether HMGA2 can be covalently trapped at genomic AP sites using NaCNBH3, parental A549 cells were first challenged with low pH or physiological pH, which is used as control. DNA purified from these cells was then incubated with HMGA2 under trapping conditions, affinity purified, and dot-blotted. The results showed that HMGA2 could only be detected in complex with genomic DNA isolated from cells that were challenged with low pH (Figure III.16A**).** This demonstrates that depurination-introduced AP sites inside the cells can be cleaved by HMGA2 *in vitro.*

In order to directly show that HMGA2 is trapped at AP sites *in vivo*, we employed A549 (1.6) cells and exposed them to low pH. Treated cells are morphologically indistinguishable from untreated controls, and were harvested and lysed at neutral pH under either trapping or non-trapping conditions. Purified DNA was dot-blotted and the results clearly showed that a significant amount of HMGA2-DNA complexes is detectable only with genomic fragments harvested under trapping conditions (Figure III.16B**).** We conclude that cells exposed to low pH challenge must contain a significant number of AP sites that can be cleaved by HMGA2 *in vitro* and *in vivo.*

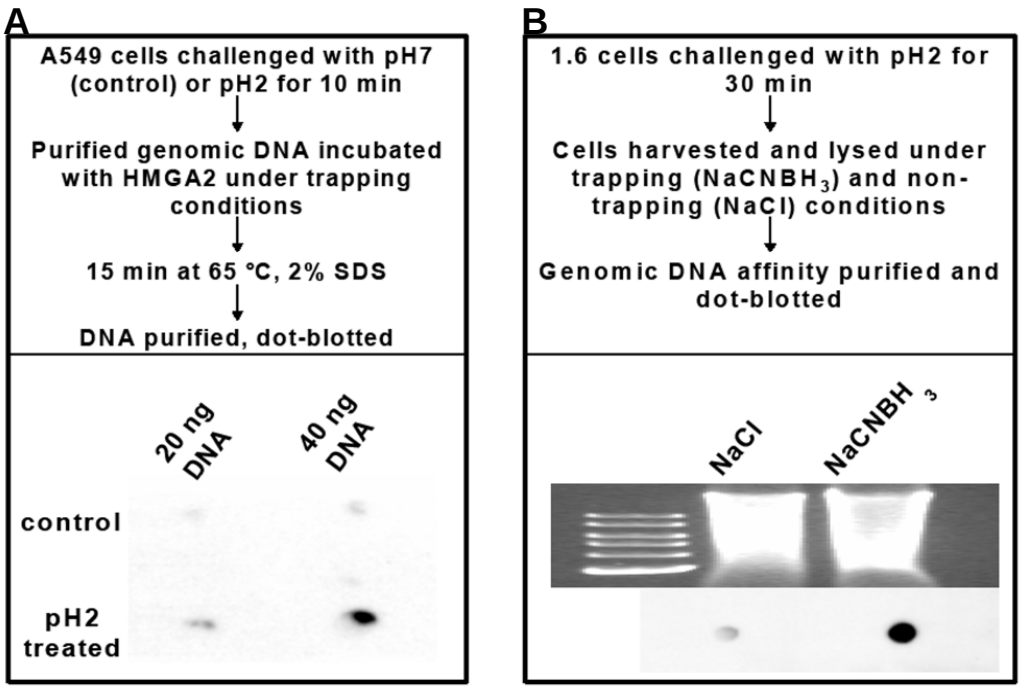


Figure III.16 ***in vitro*** **and** ***in vivo*** **trapping of HMGA2 on AP genomic DNA**

(A) Dot-blot showing ***in vitro*** **trapping of HMGA2 on purified A549 genomic DNA after depurination** ***in vivo*****. The experimental set-up was described in the upper panel.**

(B) Dot-blot showing ***in vivo*** **trapping of HMGA2 on 1.6 cell genomic DNA due to depurination. The experimental set-up was described in the upper panel. A control for the amount of DNA loaded was shown above the blot.**

Next, to analyze the cytotoxic effects that might result from depurination in parental and transfected cells, these cells were challenged with low pH for 6 min. After recovery, FACS analysis of live/dead cells revealed that all transfected cell lines are strikingly more resistant to low pH challenge than parental cells (Figure III.17**).**

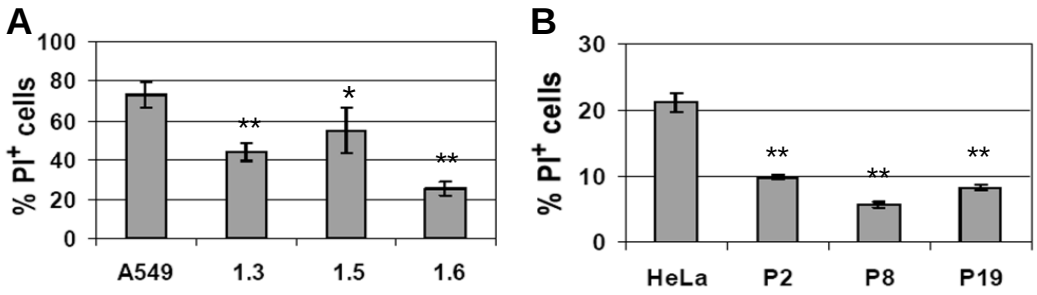


Figure III.17 Protective effect of HMGA2 on cells against low pH challenge

(A) FACS analysis of A549 dead cells (PI^+) determined after 6 min depurination at pH 2 and 24 hrs recovery. (B) FACS analysis of HeLa dead cells (PI^+) determined after 6 min depurination at pH 3 and 24 hrs recovery. The percentage values of dead cells were given by FACS gating of PI^+ events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged. Bars represent mean values with standard deviations obtained from biological triplicates. *P<0.005, **P<0.001

III.2.4 HMGA2 displays compound selectivity to protect cancer cells from genotoxicants

Hydroxyurea (Hu) is a frequently used chemotherapeutic agent for the treatment of proliferative disorders and solid tumors, and is able to induce base oxidation and depurination (Sakano et al., 2001). We exposed parental and transfected cells to Hu for a period of 48 hrs, and determined the fraction of live cells using FACS analysis. Expression of HMGA2 results in significant protection against cell death, leading to a 2 to 8 fold increase in cell survival (Figure III.18**).**

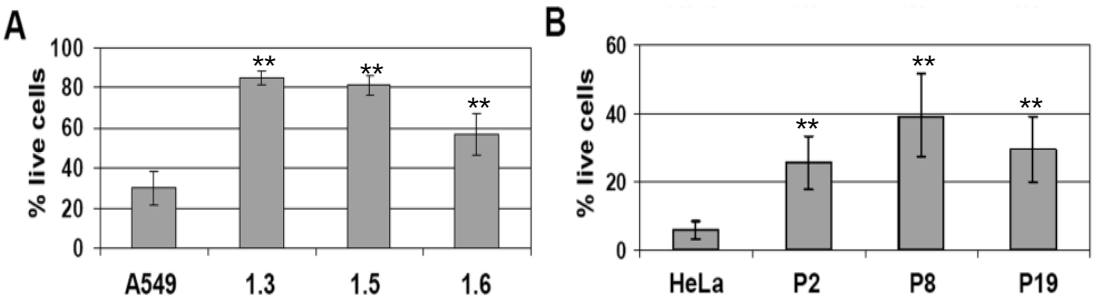


Figure III.18 Cell viability of parental and HMGA2 transgenic cancer cell lines after challenge with Hydroxyurea.

Percentage of live cells was determined by FACS after staining with Annexin-V fluorescence and Propidium Iodide. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged. Cells exposed to 100 mM Hydroxyurea for 48 hrs. **P<0.001

In order to demonstrate that transfected cells are not impaired *per se* in entering apoptosis, we next exposed cells to the chemotherapeutic agent paclitaxel, which is an inhibitor of mitosis and targets tubulin and the Bcl-2 oncoprotein, instead of DNA (Wang et al., 1999). When compared with parental cell lines, none of the HMGA2 transgenic lines displayed a protective phenotype (Figure III.19**).**

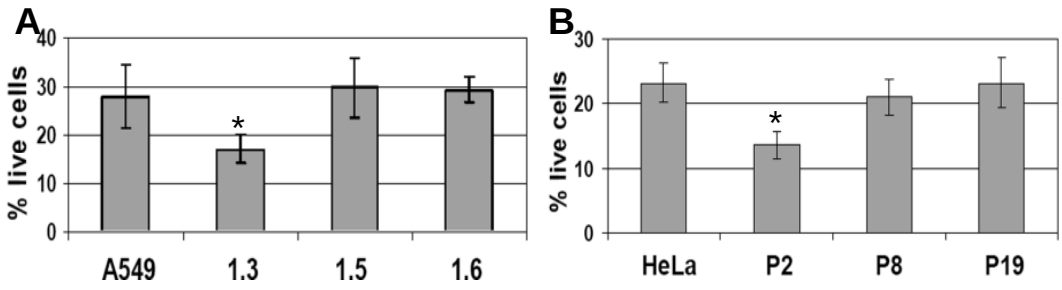


Figure III.19 Cell viability of parental and HMGA2 expressing cancer cell lines after challenge with paclitaxel.

Percentage of live cells was determined by FACS after staining with Annexin-V fluorescence and Propidium Iodide. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged. (A) A549 cells and derivatives treated with 50 nM paclitaxel for 48 hrs. (B) HeLa cells and derivatives treated with 20 nM paclitaxel for 48 hrs. *P<0.005

Expression of HMGA2 was recently shown to increase the cytotoxic effects of DNA double strand breaks induced by certain topoisomerase type II inhibitors and the chemotherapeutic agent cisplatin (Boo et al., 2005). Cisplatin DNA cross-links are mainly repaired by nucleotide excision repair (NER), and not through BER. We therefore tested whether our HMGA2 transgenic lines exhibited a similar phenotype when treated with cisplatin, and the results show that this is indeed the case (Figure III.20**).** Hence, the protective effect of HMGA2 seems limited to BER pathways.

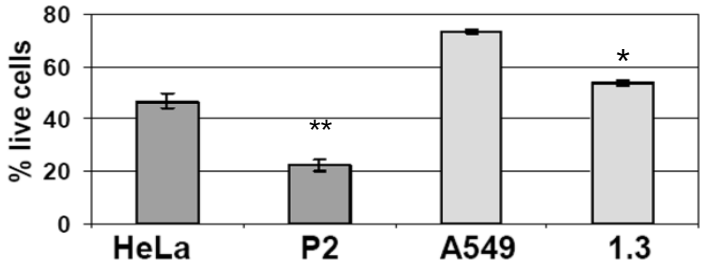


Figure III.20 Cell viability of parental and HMGA2 transgenic cancer cell lines after challenge with cisplatin.

Percentage of live cells was determined by FACS after staining with Annexin-V fluorescence and Propidium Iodide after cisplatin treatment. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged. Bar chart represents mean values with standard deviations indicated and obtained from biological assays in triplicate. *P<0.005; **P<0.001

III.2.5 HMGA2-mediated cellular protection from MMS is blocked by lyase inhibition

Both AP and dRP lyase activities play a central role in the initial steps of BER. The DNA methylating agent methyl methanesulphonate (MMS) has been used frequently to study this particular repair pathway in mammalian cells. MMS produces genomic AP sites through the action of DNA glycosylases, which remove the chemically modified bases. In a first set of experiments, we exposed parental HeLa and the respective transfected cells to 4 mM MMS for 1 hour, followed by 48 hrs recovery before FACS analysis. We included here a second HeLa cell line (EBO) as a control which, like the HMGA2 transfected lines, expresses a neomycin marker but lacks the HMGA2 transgene. The results clearly show that the presence of HMGA2 lead to a significant increase in the number of live cells in all transfected cell lines (Figure III.21**).**

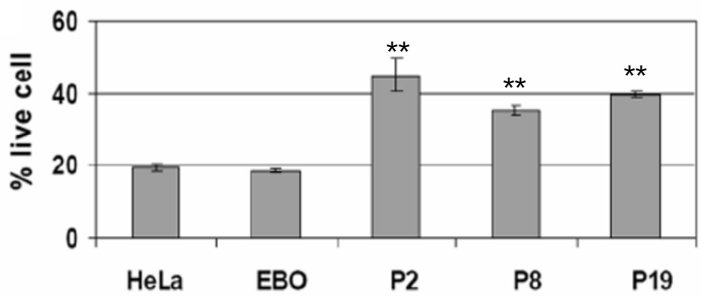


Figure III.21 Cell viability of HeLa parental and HMGA2 transgenic cancer cell lines after challenge with MMS.

Percentage of live cells determined by FACS after challenging HeLa cells and derivatives with 4mM MMS for 1 hr, followed by 48 hrs recovery. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged. EBO cells serve as control of transfected cell lines. **P<0.001.

Likewise, treatment of parental and transfected A549 cells with various MMS concentrations revealed that HMGA2 always confers strong protection against MMS-induced cytotoxicity (Figure III.22A**).** This result was confirmed by colony forming assays after MMS treatment **(**Table III.2**).**

Once a genomic AP site is generated, the next step in BER involves AP endonucleases and/or AP/dRP lyases which cleave the DNA backbone at the base lesion. However, the tautomeric open-ring form of deoxyribose produced by DNA glycosylases can react very fast with MA or BA. Both MA- and BA-adducts efficiently inhibit DNA cleavage by AP endonucleases and AP lyases (Liu and Gerson, 2004). Hence, in order to demonstrate that the protection from MMS-induced DNA damage observed with transfected cells involves HMGA2 lyase activities, we decided to inhibit these early steps in BER by sequential exposure to MMS and BA. The results show that BA alone has no effect on the survival of parental or transfected A549 cells (Figure III.22B**).** However, pre-treatment with 6 mM or 8 mM MMS followed by exposure to BA, not only negates the protection seen with transfected cells, but in fact sensitized these cell lines to MMS treatment (Figure III.22**).** This result was reproduced by colony forming assays **(**Table III.3**).**

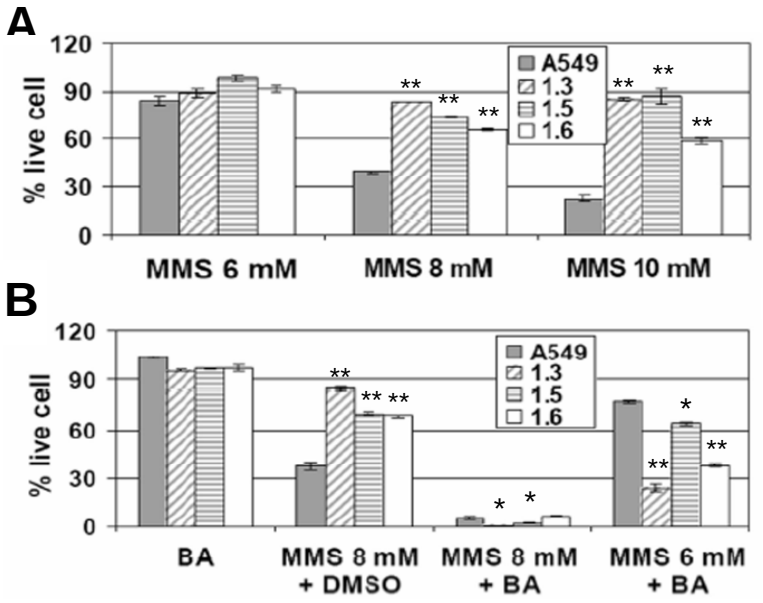


Figure III.22 Cell viability of A549 parental and HMGA2 transgenic cancer cell lines after MMS challenge with/without co-treatment with BA.

(A, B) Percentage of live cells determined by FACS after challenging A549 cells with MMS at the indicated concentrations for 1 hr, followed by 48 hrs recovery. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged Data were normalized to untreated parental and transfected cells (100%). 5 mM BA was added to an additional set of wells during recovery, and DMSO, the solvent for BA, was added to a control set. Note that BA alone does not affect cell viability. *P<0.005; **P<0.001

	A549		1.3	
	control	2 mM MMS	control	2 mM MMS
Experiment 1	120	24	265	64
Experiment 2	284	5	292	121
Experiment 3	304	133	370	329
Experiment 4	312	159	397	379
Mean/SD	255/ 91	80/ 77	306/104	223/154

Table III.2 Colony-Forming assay on A549 and 1.3 cell lines after MMS challenge.

Numbers in the table represent the number of colonies counted at the end of each assay. Cells without any challenges were used as controls.

	A549		1.3	
	control	2 mM MMS	control	2 mM MMS
– BA	312 / 367	159 / 215	397 / 310	379 / 280
+ BA	164 / 324	111 / 193	328 / 306	127 / 72

Table III.3 Colony-Forming assay on A549 and 1.3 cell lines after MMS challenge with/without BA treatment.

Numbers in the table represent the number of colonies counted at the end of each of the two assays performed. Cells without MMS challenge were used as controls.

Furthermore, we confirmed through *in vitro* assays that 5 mM BA completely inhibited the dRP lyase activity of HMGA2 (Appendix III, Figure 3C)**.**

It is noted that the degree of protection against MMS does not directly correlate with HMGA2 levels. This might be due to the fact that even in transfected A549 (1.5) cells which express HMGA2 at a comparatively low level, the intracellular amount of HMGA2 protein might be sufficient to efficiently cope with the number of AP sites generated by 10mM MMS. If so, differences in the protective effect against MMS between low and high level HMGA2 cells might become apparent only at higher MMS doses. Therefore A549 (1.5) and A549 (1.3) cells were exposed to increasing concentrations of MMS and measured cell survival as before. The results clearly reveal that the protective effect directly correlates with HMGA2 expression at MMS concentrations higher than 10mM (Figure III.23**).**

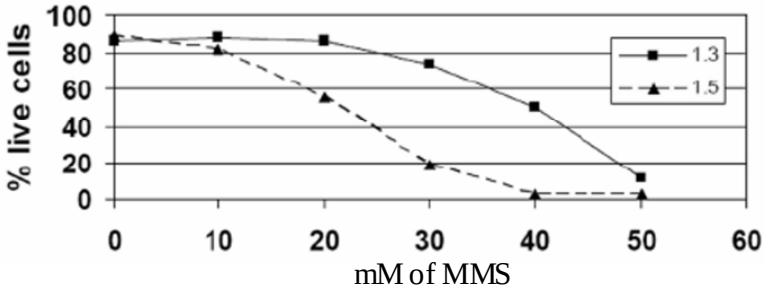


Figure III.23 Comparision of cell viability between A549 1.3 and 1.5 cell lines after increasing MMS challenge.

Percentage of live cells with A549 1.3 and 1.5 cell lines was determined by FACS after challenging with increasing concentration of MMS for 1 hr, followed by 48 hrs recovery. Annexin V and PI were used to stain apoptotic and dead cells, respectively. The percentage values of live cells were given by quaternary FACS gating of Annexin V^-/PI^- events, and the gate was set in FACS dot-blot of negative controls, i.e. cells not being challenged.

To confirm the protective effect of the cells against MMS is contributed by the presence of HMGA2 proteins but not something unique of the transfected cell lines, a colony-forming assay was done with A549 and HeLa cells exposed to MMS challenge after transiently transfected with HMGA2 expressing vector. By comparing to CMV-EGFP transfected cells, it could be seen that HMGA2 transfected cells give at least 20% more colonies after MMS challenge, and the difference gets larger upon the increase in MMS concentration (Figure III.24). In addition, these increases in colony formation are contributed by only 60% to 70% positive transfected cells judged by FACS analysis on CMV-EGFP transfected cells before MMS treatment (data not shown). This results clearly indicate that the presence of HMGA2 protein itself gives rise to the protective effect of the cells against MMS.

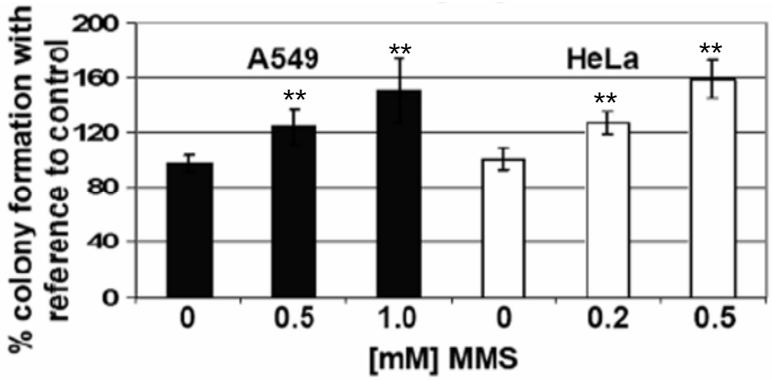


Figure III.24 Percentage of surviving colonies from HMGA2- transfected cells upon treatment of MMS.

HMGA2 transfected cells give at least 20% more colonies after MMS challenge, and the difference gets larger upon the increase in MMS concentration. Cells transfected with control vector serves as reference. **P<0.001

III.2.6 HMGA2 involves in cellular BER through interaction with APE1

With the establishment of HMGA2's AP/dRP lyase activity and its ability to protect cell against genotoxic agents, the possible linkage between these two was then investigated. One obvious possibility is that HMGA2 is part of the cellular BER pathway, and through interaction with other components in the pathway, its lyase activity is involved. It was recently reported that HMGA2 can physically interact with AP endonuclease 1, a major BER component in human cells *in vitro* (Sgarra et al., 2008), and we then tested if they two have interactions inside the cell. Through Ni+ sepharose co-affinity purification method, APE1 signal was detected with enrichment of HMGA2 from 1.3 and P2 cell lysates under native conditions; whereas no signal was detected with the parental cell lines used. The similar APE1 expression levels between the transfected cell lines and their respective parental cell lines show that this observation does not result from the intrinsic expression level (Figure III.25).

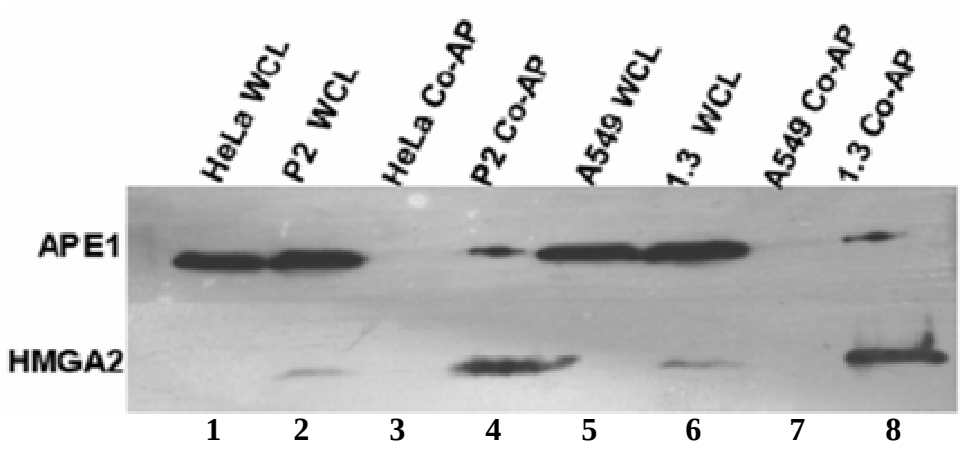


Figure III.25 Western analysis of co-affinity precipitation of cellular APE1 by HMGA2-His.

Native cell lysates of parental (HeLa and A549) and transgenic cells (P2 and 1.3) were subjected to affinity precipitation of His-tagged HMGA2 and its interating proteins using Ni+ sepharose beeds (GE Healthcare), and the elution product were examined for HMGA2 and potential interacting partner APE1 via western blotting. Precipitation products in parental and HMGA2 transgenic cells were loaded in Lane 3, 4, 7, 8. Whole cell lysates (wcl) were used as controls (Lane 1, 2, 5, 6).

IV DISCUSSION

IV.1 HMGA2 in hESCs

IV.1.1 Expression

In this study, we found that HMGA2 shows an expression level in hESCs that differs significantly from those in transformed and untransformed differentiated human cells in culture, up to 24 fold higher. We have also demonstrated that HMGA2 is abundant in hESCs and estimated that there are between 10^5 and 10^6 molecules present in a single hES cell (Li et al., 2006). This leads to 1 HMGA2 molecule per 100 or 10 NCPs, respectively.

The amount of HMGA2 protein further increased about three-fold during early stages of stem cell differentiation, after which it dropped back to or below the level found in undifferentiated cells. Interestingly, a very similar peak in HMGA2 expression was recently reported around day 10 of EB formation in mouse, where it seems to correlate with the appearance of progenitor cells committed to cardiomyogenesis and myogenesis (Caron et al., 2005). Together, these data indicate that among all main HMG proteins, HMGA2 appears to play a significant role in hESC biology.

IV.1.2 Chromosome Binding Properties

These observations and the fact that HMGA2 is a chromatin-associated factor in hESCs led us to analyze in detail the association of this protein with NCP of known atomic structure. It has also been demonstrated recently that HMGA2, using all three AT hook domains, binds to a stretch of 15 AT bps with very high affinity (Ciu et al., 2005). However, depending on the sequence composition of individual DNA cognate sites, it is also possible that one HMGA2 molecule binds with two or only one of its three hooks to a particular DNA segment. This leaves the unbound hook(s) free to interact *in trans* with a potential high affinity site present on a different DNA molecule. Obviously, an important prerequisite for *in trans* binding is a high local concentration of cognate DNA sites.

The binding stoichiometry of HMGA2 at particular sites in a genome is, therefore, an important determining factor for its biological function and likely to be influenced by the higher order structure of chromatin inside a living cell. We found that as many as 12 HMGA2 protomers form stable complexes with a single NCP. An inspection of the crystal structure revealed that the minor groove of at most 4 of the 8 AT-tracts are apparently accessible, serving as potential binding sites for AT-hooks (Davey et al., 2002). This implies that at higher molar ratios, additional HMGA2 molecules must interact with NCP in a way different from the recognition of exposed AT-tracts.

While we cannot exclude at present the possibility that HMGA2-binding to NCP induces conformational changes which expose additional AT-tracts, it is possible that the unusual DNA conformational parameters induced by the binding of the histone octamer to DNA create hitherto unidentified high affinity cognate sequences for HMGA2 (Richmond and Davey, 2003). However, it is also possible that HMGA2 might interact directly with the histone octamer. Such an interaction has been proposed for HMGA1 (Reeves and Nissen, 1993).

Because of the estimated ratio of one HMGA2 AT-hook per 100 nucleosomes in hESCs, it is physiologically more relevant to consider the situation when only one HMGA2 molecule interacts with a single NCP. It is clear from our binding studies that a one-to-one complex between HMGA2 and NCP can form without a displacement of histones. Hence, direct HMGA2-induced nucleosome reorganization appears unlikely in hESCs. We think that HMGA2, through its affinity for short AT-tracts or other unknown cognate sites exposed on nucleosomes (Reeves and Nissen, 1993), serves as a landing platform for other factors.

Interestingly, a recent analysis of potential HMGA molecular partners in human tumor cell lines revealed that HMGA2 may interact with chromatin remodeling/assembly factors (Sgarra et al., 2005). Hence, HMGA2 could indirectly affect the higher-order structure of chromatin fibers in such a way that it renders chromatin more accessible for transcriptional activities or epigenetic modifications. In this respect, it may be relevant that a number of genes involved in cell cycle regulation and transformation, such as cyclin A2, cyclin I, rb1, and BRCA 2 are located on chromosomes 4 and 13. These chromosomes exhibit the highest AT content among all human chromosomes. In addition, it is worth noting that microinjection of HMGA1, which also contains

three AT-hooks, into one-cell mouse embryos induces early activation of gene transcription (Beaujean et al., 2000).

It is therefore likely that a high level of HMGA2 protein in hESCs contributes to the establishment of a transcriptional profile through the definition of a particular hES cell chromatin state and that during early stages of cell differentiation, the observed upregulation of HMGA2 expression leads to changes in that profile. A detailed crystallographic analysis of NCP-HMGA2 complexes and functional studies in hESCs should provide further insight into the interesting question of how the proto-oncogene HMGA2 defines the state of hESC chromatin.

Through the indication of HMGA2 abilities to compete with histone H1; to directly interact with nucleosomes; to bind *in trans* to different DNA segments and to recruit histone acetylase, it is reasonable then to suggest that the protein is able to promote the formation of domains characterized by a more open, accessible chromatin structure. It is clear that this will primarily depend on two factors, the expression level of HMGA2 in an individual cell and the genomic location/clustering of higher affinity cognate sites in the form of AT-rich DNA. It is well established that open chromatin domains are transcriptionally more active. Hence, HMGA2 may in this way contribute to the establishment and/or maintenance of particular transcriptional networks.

IV.1.3 HMGA2 Knockdown in hESCs

With respect to these observations, our FACS analysis also found that during the early differentiation in EBs, HMGA2 is up-regulated in a subset of cells. This strongly suggested that HMGA2 is associated with a specific type of differentiation from ES cell, possibly through its role as a chromatin factor, and regulation for the transcription of certain sets of genes that govern the ESC's differentiation towards a particular direction.

Through siRNA -mediated down-regulation of HMGA2 in hESCs, we found that the expression of key pluripotency markers is down-regulated, including Oct3/4, Sox2 and UTF1. We also analyzed mRNA levels of well-known markers for all 3 germ layers and found that down-regulation of HMGA2 reduces the transcript level for the mesoderm marker gene insulin like growth factor 2 (IGF2) and for the neural stem cell specific marker NESTIN. In contrast, knock-down of HMGA2 significantly increases the levels of transcripts coding for endoderm markers SOX17 and MIXL1. These findings suggest that HMGA2 plays a role in maintaining a balance of mesodermal versus endodermal cell differentiation in hESCs. Interestingly, in addition to IGF2's role in mesodermal cell lineage commitment, it has also been associated with human fetal growth (Kaku et al., 2007), which, again, links HMGA2 with cell proliferation during embryogenesis.

Similiarly, our siRNA approach in combination with the PIQORTM Stem Cell Microarray platform revealed that a number of genes were differentially expressed at a significant level. Many of these genes also are markers of either ESC pluripotency or early differentiation, in line with our findings from RT-PCR results.

Pluripotency Markers Down-regulation

Oct4, Sox2 and UTF1 expression levels were highly down-regulated upon HMGA2 knockdown in hESCs with our qRT-PCR and microarray analysis. While Oct4 and Sox2 are well-known as the standard markers of pluipotency, the transcriptional cofactor undifferentiated transcription factor 1 (UTF1) is also regarded to be very important in defining the hESC pluripotency niche (Bhattacharya et al., 2004; Trounson, 2006). UTF1 expression is rapidly extinguished upon induced differentiation in mouse embryonic carcinoma cells and downregulated in the primitive endoderm of the mouse embryo during the primitive streak stage (Okuda et al., 1998). There are observations showing that endogenous UTF1 is downregulated to nearly undetectable levels, faster than Oct4 or Nanog at the onset of differentiation (Tan et al., 2008). Together, these observations firmly establish UTF1 as a sensitive and reliable pluripotency marker for hESCs, and give a possible explanation for its higher down regulation than that of Oct4 and Sox2 in our qRT-PCR analysis. In combination with these findings, our microarray analysis clearly shows the expression of iconic pluripotency markers Oct4 and Sox2 are significantly down-regulated upon HMGA2 knock-down, suggesting HMGA2 is an important chromosomal factor of which the expression level is crucial for hESC's pluripotent characteristics.

Slow hESCs Growth Rate

Recently, Genomics Institute of Novartis Research Foundation showed that UTF1 might be involved in mESC proliferation (Nishimoto et al., 2005). Our results showed that HMGA2 is involved in the regulation of UTF1, and hESCs with HMGA2 knockdown do have a slower proliferation rate. These together with the report that

mouse embryonic fibroblasts derived from HMGA2 null mice exhibit a significantly slower proliferation rate than wild-type cells (Zhou et al., 1995), possibly suggesting that HMGA2 is associated with the regulation of ES cell growth by regulating the UTF1 expression level.

Mixl1 Up-regulation

The expression of differentiation marker Mixl1, is most affected upon HMGA2 down-regulation in our qRT-PCR analysis. The Mixl1 gene encodes a homeodomain transcription factor that is required for normal mesoderm and endoderm development in mouse. While Mixl1 is required for efficient differentiation of cells from the primitive streak stage to blood, a higher proportion of one endoderm marker E-Cadherin positive cells was detected in embryoid bodies (EBs), differentiated from mESCs with enforced constitutive Mixl1 expression. Also during hematopoietic mesoderm differentiation, Mixl1 expression results in low Flk1 expression and failure in hemoglobinized cells development. It was shown that Mixl1 overexpressing EBs have extensive areas containing cells with an epithelial morphology that express endoderm markers E-Cad, FoxA2 and Sox17, consistent with enhanced endoderm formation. Genetic studies also indicate that Mixl1 can trans-activate the promoters of endoderm markers Gsc, Sox17 and E-Cad, supporting the fact that Mixl1 has a direct role in definitive endoderm formation (Alexander et al., 1999; Pearce and Evans, 1999; Hart et al., 2002; Kofron et al., 2004; Mohn et al., 2003).

Taken together, the level of Mixl1 expression influences the allocation of cells between mesoderm and endoderm during ESC differentiation *in vitro*, with high levels of Mixl1 favoring endoderm formation. In this case, based on our finding that Mixl1

expression is much up-regulated upon HMGA2 knockdown, we could likely conclude that the down-regulation of HMGA2 promotes ESC differentiate into endoderm.

Leftys Up-regulation

Our finding showed that upon HMGA2 knockdown, expression of LEFTY genes, LEFTYA and LEFTYB, is increased. These genes are transiently expressed during early hESC differentiation in EBs. Their expression appears to play a substantial role in mesodermal cell lineage establishment (Dvash et al., 2007). They are known to be controlled by, and also a feedback inhibitor of, the Nodal signaling pathway. Nodal, a member of the TGF-β family of signaling molecules, has been implicated the roles in hESCs pluripotency (Vallier et al., 2004), and mesoderm/endoderm specification during gastrulation.

EBs derived from Lefty overexpressing hESCs show increased expression of neuroectoderm markers Sox1, Sox3, and Nestin. Conversely, they are negative for Sox17, a definitive endoderm marker, and do not generate beating cardiomyocyte structures in conditions that allows mesoendoderm differentiation from wildtype hESCs. EBs derived from Lefty expressing hESCs also contain a greater abundance of neural rosette structures as compared to controls (Smith et al., 2007). Thus, Leftys, by inhibiting Nodal signaling, promotes neuronal specification, indicating a role for this pathway in controlling early neural development of pluripotent cells (Chen and Shen, 2004).

Leftys are also involved in the left-right body symmetry formation and control of gastrulation. Interestingly, both genes are located on human chromosome 1 as direct repeats and separated by only 50 kb. While the overall AT-content of human

chromosome 1 is 58%, the chromosomal region spanning the two genes has a substantially higher AT-content of 69%. Hence, we propose that HMGA2 binds to this region and contributes there to the formation of a particular chromatin structure that is transcriptionally silenced. This domain can then be activated when HMGA2 levels drop during certain stages of differentiation, or as in our case, upon HMGA2 siRNA treatment.

Dlk1 Up-regulation

One interesting gene, of which the expression level is up-regulated with HMGA2 knock-down, is DLK1. During the past years, evidence has accumulated suggesting that dlk1 participates in the control of several differentiation processes, including neuroendocrine differentiation, differentiation of hepatocytes, hematopoiesis, osteogenesis and adipogenesis. It may also be involved in the regulation of cell growth.

Dlk1 has been shown to act as an inhibitor of adipogenesis (Smas and Sul, 1997). It is highly expressed in 3T3-L1 preadipocytes, but its expression is down-regulated during differentiation of these cells into adipocytes. Furthermore, constitutive expression of Dlk1 in 3T3-L1 cells inhibits their adipocyte differentiation, whereas forced down-regulation of Dlk1 by antisense expression enhances adipogenesis of both these and Balb/c 3T3 cells (Smas and Sul, 1993).

The importance of Dlk1 in adipogenesis has been demonstrated *in vivo* through the generation of Dlk1-deficient mice. The mice show a decrease in adipose tissue, and suffer from decreased glucose tolerance, hypertriglyceridemia, and lower insulin sensitivity. On the other hand, they show growth retardation, increased serum lipid metabolites, and develop obesity (Sul et al., 2000).

Therefore, our *in vitro* result which shows HMGA2 is associated with Dlk1's expression in hESCs could suggest that HMGA2 is involved in adipogenesis *in vivo*, and possibly provide the first explanation on the pygmy phenotype of HMGA2 knockout mice (Zhow et al., 1995).

In all, our data identify HMGA2 as a regulator of genes linked to mesodermal cell differentiation, adipogenesis, and cell proliferation in hESCs. They provide the basis for further studies elucidating the molecular mechanisms how these genes are regulated by HMGA2 in pluripotent hESCs, and also during the early stages of human development ex vivo. Furthermore, a better understanding of HMGA2's role in gene regulation may ultimately allow us to direct differentiation of hESCs into particular cell lineages by controlling HMGA2 protein levels.

IV.2 Genotoxicity protection function of HMGA2 as an AP/dRP lyase

IV.2.1 Intrinsic AP/dRP lyase

We have shown in this study that HMGA2 and the homologous HMGA1 proteins have potent intrinsic dRP/AP lyase activities. Furthermore, we presented evidence that the nucleophile(s) that attack AP sites during β elimination reside within AT-hooks. Substitutions within hooks which could compromise or eliminate lyase activities would almost certainly also affect HMGA2 functions as a DNA architectural chromatin factor. In addition, based on the fact that the sequences of AT-hooks are almost identical, substitutions would have to be introduced into all three hooks simultaneously; some of them might also interfere with nuclear import and HMGA2-pRB protein interactions in addition to DNA-binding (Fedele et al., 2006; Cattaruzzi et al., 2007). Hence, it does not seem that a mutational screen is a productive approach to distinguish the lyase activities from the other critical biological functions of HMGA2.

IV.2.2 Genotoxicity Protection

We presented here three lines of evidence that these lyase activities do have important biological functions and play a direct role in conferring resistance against certain DNA-damaging agents. First, HMGA2 was efficiently trapped in a covalent complex with genomic AP sites *in vivo*. This indicated that the enzyme cleaves AP sites within the background of chromatin. Second, subsequent to MMS challenge, treatment with BA sensitizes transfected cells to MMS, thus negating the protection that was observed when cells are exposed to MMS only. Since BA reacts with the deoxyribose moiety at

AP sites and thereby inhibits mammalian AP endonuclease and AP/dRP lyase activities, and due to our finding that HMGA2 interacts with APE 1, we conclude that HMGA2 confers resistance to MMS, and other AP site-inducing genotoxicants, through a direct role in BER that involves its lyase activities. Third, the protective effect directly correlates with the expression level of HMGA2 at elevated MMS concentrations.

Our data showed that HMGA2 confers cellular protection from three different genotoxicants: Hu, MMS, and low pH. It is clear that this is not a consequence of HMGA2's role as chromatin factor and transcriptional regulator which could affect the ability of transfected cells to enter apoptosis, since exposure to the chemotherapeutic agent paclitaxel of these cells revealed no difference in cytotoxicity compared with parental cells. In addition, expression of HMGA2 was recently shown to increase the cytotoxic effects of DNA double strand breaks induced by certain topoisomerase inhibitors and the chemotherapeutic agent cisplatin (Boo et al., 2005). The latter effect is attributed to the inhibition of the nucleotide excision repair (NER) gene ERCC1, which is negatively regulated by HMGA2 (Borrmann et al., 2003). Also, it was hypothesized that HMGA2 masks certain DNA lesions from recognition by the NER machinery (Reeves and Adair, 2005). We confirmed in this study that HMGA2 sensitizes cancer cells to cisplatin, thus further demonstrating that the ability of our transfected cells to enter apoptosis is not compromised. Taken together, a major conclusion from our data is that even transient expression of HMGA2 has a direct supporting role in BER, which involves dRP/AP lyase activities. This leads to resistance of cancer cells against certain genotoxicants. In contrast, the presence of HMGA2 inhibits NER and, thereby, sensitizes cells to a different class of compounds.

The steady state number of AP sites in the genome of mammalian cells is at least 10,000. Most of these sites appear to be processed quickly by APE1, which generates the 5'-dRP moiety (Nakamura and Swenberg, 1999). Its removal by dRP lyase seems to be the rate-limiting step in single nucleotide BER, and the level of dRP lyase activity will, therefore, directly affect the cellular repair capacity and sensitivity against DNA damage that leads to AP sites (Sobol et al., 2000). This scenario is in agreement with our finding that an elevated level of HMGA2 protects cancer cells from DNA-damage-induced cytotoxicity. Mechanistically, BER most likely involves coordinated protein-protein and protein-DNA interactions. It is clear that the level of other BER factors, such as OGG1, APE1, or the scaffold protein XRCC1, might also affect the repair capacity in a situation where there is an abundance of dRP lyase molecules (Almeida and Sobol, 2007; Parsons et al., 2008). In agreement with a recent report on HMGA2 and APE1 interaction *in vitro*, we have demonstrated here that the two proteins also interact with each other inside cells.

Exposure of cells to the chemotherapeutic agent Hydroxyurea (Hu) has been shown to induce base oxidation and depurination (Sakano et al., 2001). However, the main cytotoxic effect of Hu is attributed to the inhibition of the enzyme ribonucleotide reductase, which, in turn, inhibits DNA synthesis during replication and repair (Weinberger et al., 1999). Our results showed that HMGA2 protects cells from Hu-induced cytotoxicity, and it is conceivable that this involves HMGA2's novel activity in AP site processing during BER. However, we were surprised by the magnitude of the protection. This might indicate that HMGA2 could participate in other aspects of genome stability; for example, rescue of stalled replication forks, or the chemical processing of single strand breaks which will also result from Hu treatment.

It is well established that unrepaired AP sites are mutagenic (Loeb and Preston, 1986). Since HMGA2 is found at high levels in many human neoplasias and in pluripotent hESCs, the interesting question emerges whether HMGA2, through its associated lyase activities, contributes to lower mutation rates. From an evolutionary point of view, this is obviously most important for high metabolically active hESCs which replicate their genomes fast and eventually produce germ cells. In fact, recent data showed that genes involved in different types of DNA repair are significantly upregulated in hESCs when compared with differentiated progeny cells (Saretzki et al., 2008).

Our finding that HMGA2 protects cancer cells from certain DNA-damaging agents used in cancer treatment has important implications for disease diagnosis, choice of treatment regimens, and the future development of anti-cancer drugs (Ding et al., 2006). First, we think it will now become even more important than previously suggested (Fusco and Fedele, 2007) to type malignant tumors, and, whenever possible, the replenishing cancer stem cell compartment with respect to the HMGA2 status. This knowledge will impact on the choice of the most suitable chemotherapeutic agents. Since we have shown that HMGA1 isoforms also exhibit dRP/AP lyase activities *in vitro*, we anticipate that the protective effect observed for HMGA2-expressing cancer cells will also be detectable for tumors that show an elevated level of HMGA1 proteins. In that respect, it is noteworthy that HMGA1 was recently shown to belong to a class of proteins that are specifically phosphorylated upon ATM/ATR activation through DNA damage (Matsuoka et al., 2007). It will, therefore, be important to investigate in the future how posttranslational modifications affect the newly discovered lyase activities of HMGA proteins.

It is probably worth to note that HMGA proteins are not the first subfamily of HMG proteins reported bearing lyase activities. Recent report identified the high-mobility group box 1 protein (HMGB1) as specifically interacting with the BER intermediate. Purified HMGB1 was found to have weak dRP lyase activity and to stimulate AP endonuclease activities on BER substrates. Coimmunoprecipitation experiments revealed interactions of HMGB1 with known BER enzymes, and GFP-tagged HMGB1 was found to accumulate at sites of oxidative DNA damage in living cells. HMGB1(-/-) mouse cells were reported slightly more resistant to MMS than wild-type cells, probably due to the production of fewer strand-break BER intermediates. The results suggest HMGB1 is a BER cofactor capable of modulating BER capacity in cells (Prasad et al., 2007).

We have provided first evidence obtained with MMS and BA that a combination therapy which targets HMGA2's role in BER and includes established genotoxicants could prove particularly effective against cancer cells otherwise resistant to chemotherapy. Secondly, anti-cancer therapy targeting HMGA2 will likely not affect normal adult cells which lack detectable HMGA2 levels, but exclusively target HMGA2 expressed in malignant cells. This novel class of anti-cancer compounds could act by interfering with DNA-binding (Fusco and Fedele, 2007) and/or protein-protein interactions in BER involving HMGA2.

V CONCLUSION and FUTURE WORK

In this study, we successfully pointed out the well-known proto-oncogene, HMGA2, plays functional roles in hESCs. We first identified that HMGA2, among all HMG proteins, is highly expressed in hESCs. Its expression level is further up-regulated in a subset of cells in early EBs. With its role as a chromatin factor bearing relatively low binding specificity, high level of HMGA2 contributes to the establishment of a transcriptional profile associated with a particular hESC chromatin state, and during early stages of cell differentiation, the observed upregulation of HMGA2 expression leads to changes in that profile.

The combination of siRNA technology, qRT-PCR and stem cell specific microarray analysis further revealed this role of HMGA2. The change of HMGA2 expression level results in transcriptional changes of many genes, which encode markers for either pluripotency or specific differentiation, in hESCs. In other words, the chromatin status, by which hESCs are largely defined, is changed. In a simpler view, HMGA2 expression level is associated with hESCs' pluripotency and differentiation directions.

However, these conclusions are still not firm enough. There is a missing link on how HMGA2's presence/absence affects the transcription of those genes. Given its close family member HMGA1's function on regulating transcription via enhansosome formation, HMGA2 very much likely applies similar strategy. One good approach that can be applied to investigate on this is ChIP-on-chip analysis, although it is still not perfect because of the potential false-positives created by the formaldehyde cross-linking step.

We also presented in this study a novel role of HMGA proteins, as part of the BER machinery. For the first time, HMGA proteins were reported bearing intrinsic AP/dRP lyase activities, and the active sites likely reside on the positively charged amino acids in their DNA binding domain, AT-Hooks. The lyase activity of HMGA2 can act as part of the BER machinery through interaction with other partners, APE1 for instance, and protect tumor cells from certain genotoxicants. These observations could provide a possible explanation that tumor-inducing cells expressing HMGA2 are more resistant to chemotherapies (Yu et al., 2007), and also provide a possible drug target for cancer treatment. In conjunction with its high expression in hESCs, this novel function of HMGA2 could give hESCs protection against external stress, which is important from an evolutionary point of view.

A good project then can be carried out, and is in the process in our lab, to directly analyze the role of HMGA2 as a lyase in protecting hESCs from genotoxicants. When working on this, an important thing to note is to manipulate HMGA2 expression level in a controlled manner, e.g. Tet-on/off system, because of its association with pluripotency and differentiation.

VI REFERENCES

Abe, N, Watanabe, T, Suzuki, Y, Matsumoto, N, Masaki, T, Mori, T, Sugiyama, M, Chiappetta, G, Fusco, A, Atomi, Y. 2003. An increased high-mobility group A2 expression level is associated with malignant phenotype in pancreatic exocrine tissue. Br. J. Cancer *89*, 2104-2109.

Alexander J, Rothenberg M, Henry GL, et al. casanova plays an early and essential role in endoderm formation in zebrafish. Dev Biol. 1999;215:343-357.

Allinson, SL, Dianova, II, and Dianov, GL. 2001. DNA polymerase β is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts. EMBO. J. *20*, 6919-6926.

Almeida, KH, and Sobol, RW. 2007. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. DNA Repair (Amst.) 6, 695-711.

Amit M, Carpenter MK, Inokuma MS, Chiu CP, Harris CP, Waknitz MA, Itskovitz-Eldor J, Thomson JA. 2000. Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev Biol 227(2):271-8.

Amit M, Itskovitz-Eldor J. Sources, 2006. Derivation, and Culture of Human Embryonic Stem Cells. Semin Reprod Med 24:298-303.

Amit M, Margulets V, Segev H, Shariki K, Laevsky I, Coleman R, ItskovitzEldor J. 2003. Human feeder layers for human embryonic stem cells. Biol Reprod 68:2150–2156.

Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, Groth D, Lehrach H, Burdon T, Adjaye J. 2007. Analysis of oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. Stem Cells 25(2):500-10.

Battista, S, Fidanza, V, Fedele, M, Klein-Szanto, AJ, Outwater, E, Brunner, H, Santoro, M, Croce, CM, Fusco, A. 1999. The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. Cancer Res. 59, 4793-4797.

Benjamini, Y and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc., Ser. B 57, 289–300.

Berner, JM, Meza-Zepeda, LA, Kools, PF, Forus, A, Schoenmakers, EF, Van de Ven, WJ, Fodstad, O, Myklebost, O. 1997. HMGIC, the gene for an architectural transcription factor, is amplified and rearranged in a subset of human sarcomas. Oncogene 14, 2935-2941.

Bhattacharya B, Miura T, Brandenberger R, Mejido J, Luo Y, Yang AX, Joshi BH, Ginis I, Thies RS, Amit M, Lyons I, Condie BG, Itskovitz-Eldor J, Rao MS, Puri RK. 2004 Gene expression in human embryonic stem cell lines: unique molecular signature. Blood; 103(8):2956-64.

Bianchi ME, Agresti A. 2005. HMG proteins: Dynamic players in gene regulation and differentiation. Curr Opin Gen Dev 15:1–11.

Boo, LM, Lin, HH, Chung, V, Zhou, B, Louie, SG, O'Reilly, MA, Yen, Y, Ann, DK. 2005. High mobility group A2 potentiates genotoxic stress in part through the modulation of basal and DNA damage-dependent phosphatidylinositol 3-kinase-related protein kinase activation. Cancer Res. 65, 6622-6630.

Borrmann, L, Schwanbeck, R, Heyduk, T, Seebeck, B, Rogalla, P, Bullerdiek, J, Wisniewski, JR. 2003. High mobility group A2 protein and its derivatives bind a specific region of the promoter of DNA repair gene ERCC1 and modulate its activity. Nucleic Acids Res. 31, 6841-685 1.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122(6):947-56.

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Werning M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349–353.

Bustin M. 2001. Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends Biochem Sci 26:152, 153.

Bustin, M, and Reeves, R. 1996. High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Prog. Nucleic Acids Res. Mol. Biol. 54, 35-100.

Caron, L, Bost, F, Prot, M, Hofman, P, and Binetruy, B. 2005. A new role for the oncogenic high-mobility group A2 transcription factor in myogenesis of embryonic stem cells. Oncogene 24, 6281-6291.

Cattaruzzi, G, Altamura, S, Tessari, MA, Rustighi, A, Giancotti, V, Pucillo, C, Manfioletti, G. 2007. The second AT-hook of the architectural transcription factor HMGA2 is determinant for nuclear localization and function. Nucleic Acids Res. 35, 175 1-1760.

Chambers I, Colby D, Robertson M, et al. 2003 Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell; 113:643–655.

Chen C and Shen MM. 2004 Two Modes by which Lefty Proteins Inhibit Nodal Signaling. Curr Bio 14, 618-624

Chieffi, P, Battista, S, Barchi, M, Di Agostino, S, Pierantoni, GM, Fedele, M, Chiariotti, L, Tramontano, D, Fusco, A. 2002. HMGA1 and HMGA2 protein expression in mouse spermatogenesis. Oncogene 21, 3644-3650.

Di Celo, F, Hillion, J, Hristov, A, Wood, LJ, Mukherjee, M, Schuldenfrei, A, Kowalski, J, Bhattacharya, R, Ashfaq, R, and Resar, LMS. 2008. HMGA2 participates in transformation in human lung cancer. Mol. Cancer Res. 6, 743-750.

Cleynen, I, and Van de Ven, WJ. 2008. The HMGA proteins: a myriad of functions. Int. J. Oncol. 32, 289-305.

Cowan CA, Klimanskaya I, McMahonJ, AtienzaJ, WitmyerJ, ZuckerJP, Wang S, Morton CC, McMahon AP, Powers D, Melton DA. 2004. Derivation of embryonic stem-cell lines from human blastocysts. N EnglJ Med 350:1353–1356.

Cui T, Wei S, Brew K, Leng F. 2005. Energetics of binding the mammalian high mobility group protein HMGA2 to poly (dA-dT)2 and poly (dA)-poly(dT). J Mol Biol 325:629–645.

Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol 319:1097–1113.

Ding, J, Miao, ZH, Meng, LH, and Geng, MY. 2006. Emerging cancer therapeutic opportunities target DNA-repair systems. Trends Pharmacol. Sci. *27,* 338-344.

Dröge, P, and Davey, CA. 2008. Do cells *let-7* determine stemness? Cell Stem Cell *2,* 8-9.

Dvash, T, Sharon, N, Yanuka, O and Benvenisty, N. 2007. Molecular analysis of LEFTY-expressing cells in early human embryoid bodies. Stem Cells 25, 465–472.

Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, Luger K. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol 375:23–44.

Ecochard V, Cayrol C, Rey S, et al. A novel Xenopus mix-like gene milk involved in the control of the endomesodermal fates. Development. 1998;125:2577-2585.

Fedele M, Battista S, Manfioletti G, Croce CM, Giancotti V, Fusco A. 2001. Role of the high mobility group A proteins in human lipomas. Carcinogenesis 22:1583–1591.

Fedele, M, Visone, R, De Martino, I, Troncone, G, Palmieri, D, Battista, S, Ciarmiello, A, Pallante, P, Arra, C, Melillo, RM. 2006. HMGA2 induces pituitary tumorigenesis by enhancing E2F1 activity. Cancer Cell 9, 459-47 1.

Fusco, A, and Fedele, M. 2007. Roles of HMGA proteins in cancer. Nat Rev Cancer *12,* 899-910.

Gattas, GJ, Quade, BJ, Nowak, RA, and Morton, CC. 1999. HMGIC expression in human adult and fetal tissues and in uterine leiomyomata. Genes Chromosomes Cancer *25,* 3 16-322.

Gregoire F, Smas CM, Sul HS. Understanding adipocyte differentiation. 1998 Physiol Rev; 78: 783-809.

Grosschedl R, Giese K, Pagel J. 1994. HMG domain proteins: Architectural elements in the assembly of nucleoprotein structures. Trends Genet 10:94–100.

Hart AH, Hartley L, Sourris K, et al. Mixl1 is required for axial mesendoderm morphogenesis and patterning in the murine embryo. Development. 2002;129:3597-3608

Hebert, C, Norris, K, Scheper, MA, Nikitakis, N, and Sauk, JJ. 2007. High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma. Mol. Cancer *6:* 5.

Henderson JK, Draper JS, Baillie HS, Fishel S, Thomson JA, Moore H, Andrews PW. 2002. Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. Stem Cells 20(4):329-37.

Hirning-Folz, U, Wilda, M, Rippe, V, Bullerdiek, J, and Hameister, H. 1998. The expression pattern of the Hmgic gene during development. Genes Chromosomes Cancer *23,* 350-357.

Hirst CE, Ng ES, Azzola L, Voss AK, Thomas T, Stanley EG, Elefanty AG. 2006. Transcriptional profiling of mouse and human ES cells identifies SLAIN1, a novel stem cell gene. Dev Biol 293(1):90-103.

Hoffman LM, Carpenter MK. 2005. Characterization and culture of human embryonic stem cells. Nat Biotechnol 23(6):699-708.

Hough SR, Clements I, Welch PJ, Wiederholt KA. 2006 Differentiation of mouse embryonic stem cells after RNA interference-mediated silencing of OCT4 and Nanog. Stem Cells; 24:1467–1475.

Huth, JR, Bewley, CA, Nissen, MS, Evans, JN, Reeves, R, Gronenborn, AM, Clore, GM. 1997. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. Nat. Struct. Biol. *4*, 657-665.

James D, Noggle SA, Swigut T, Brivanlou AH. 2006. Contribution of human embryonic stem cells to mouse blastocysts. Dev Biol 295(1):90-102.

Kaku, K, Osada, H, Seki, K and Sekiya, S. 2007. Insulin-like growth factor 2 (IGF2) and IGF2 receptor gene variants are associated with fetal growth. Acta Paediatr. 96, 363–367.

Klostermann, A., Lutz, B., Gertier, F. and Behl, C. (2000) The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1) bind to the enabled/vasodilator stimulated phos¬phoprotein-like protein (EVL) via a novel carboxy-terminal zyxin-like domain. J. Biol. Chem. 275, 39647–39653.

Kofron M, Wylie C, Heasman J. The role of Mixer in patterning the early Xenopus embryo. Development. 2004;131:2431-2441.

Kuroda T, Tada M, Kubota H, Kimura H, Hatano SY, Suemori H, Nakatsuji N, Tada T. Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol Cell Biol. 2005; 25(6):2475-85.

Laslett AL, Filipczyk AA, Pera MF. Characterization and culture of human embryonic stem cells. Trends Cardiovasc Med. 2003; 13(7):295-301.

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, ZuckerJP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. 2006. Control of devel-

opmental regulators by polycomb in human embryonic stem cells. Cell 125:301–313.

Li, O., Vasudevan, D., Davey, C.A., and Dröge P. (2006). High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. Genesis *44,* 523-529.

Li, O., Li, J., and Dröge, P. (2007). DNA architectural factor and proto-oncogene HMGA2 regulates key developmental genes in pluripotent human embryonic stem cell. Febs Lett. *518,* 3533-3537.

Liau, S.S., and Whang, E., (2008). HMGA1 is a molecular determinant of chemoresistance to gemcitabine in pancreatic adenocarcinoma. Clin. Cancer Res. *14,* 1470-1477.

Ligon, A.H., Moore, S.D., Parisi, M.A., Mealiffe, M.E., Harris, D.J., Ferguson, H.L., Quade, B.J. and Morton, C.C. (2005) Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas. Am. J. Hum. Genet. 76, 340–348.

Liu, L., and Gerson, S.L. (2004). Therapeutic impact of methoxyamine: blocking repair of abasic sites in the base excision repair pathway. Curr. Opin. Investig. Drug. *5,* 623- 627.

Loeb, L.A., and Preston, B.D. (1986). Mutagenesis by apurinic/apyrimidinic sites. Annu. Rev. Genet. *20,* 20 1-230.

Luger K, Rechsteiner TJ, Richmond TJ. 1999. Preparation of nucleosome core particle from recombinant histones. Methods Enzymol 304:3–19.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald III, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y. et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science *316,* 1160-1166.

Mayr, C., Hemann, M.T., and Bartel, D.P. (2007). Disrupting the pairing between *let-7* and Hmga2 enhances oncogenic transformation. Science *315,* 1576-1579.

McMurray, H., Sampson, E.R., Compitello, G., Kinsey, C., Newman, L., Smith, B., Chen, S-R., Klebanov, L., Salzman, P., Yakovlev, A., and Land, H. (2008). Nature *453*, 1112-1116.

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell 10:105–116.

Meyer, B., Loeschke, S., Schultze, A., Weigel, T., Sandkamp, M., Goldmann, T., Vollmer, E., Bullerdiek, J. (2007). HMGA2 overexpression in non-small cell lung cancer. Mol. Carcinog. ***46,*** 503-511.

Mitsui K, Tokuzawa Y, Itoh H,Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 2003; 113:631-42.

Miyazawa, J., Mitoro, A., Kawashiri, S., Chada, K.K., and Imai, K. (2004). Expression of mesenchyme-specific gene HMGA2 in squamous cell carcinomas of oral cavity. Cancer Res. *64*, 2024-2029.

Mohn D, Chen SW, Dias DC, et al. Mouse Mix gene is activated early during differentiation of ES and F9 stem cells and induces endoderm in frog embryos. Dev Dyn. 2003;226:446-459

Nakamura, J., and Swenberg, J.A. (1999). Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues. Cancer Res. *59*, 2522-2526.

Narita M, Krizhanovsky V, Nunez S, Chicas A, Hearn SA, Myers MP, Lowe SW. 2006. A novel role for high-mobility group A proteins in cellular senescence and heterochromatin formation. Cell 126:503–514.

Nichols J, Chambers I, Taga T, Smith A. 2001 Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. Development; 128:2333–2339.

Nichols J, Zevnik B, Anastassiadis K, et al. 1998 Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell; 95:379–391.

Nishimoto, M., Miyagi, S., Yamagishi, T., Sakaguchi, T., Niwa, H., Muramatsu, M. and Okuda, A. (2005) Oct-3/4 maintains the proliferative embryonic stem cell state via specific binding to a variant octamer sequence in the regulatory region of the UTF1 locus. Mol. Cell. Biol. 25, 5084–5094.

Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 2000; 24(4):372-6.

Okuda A, Fukushima A, Nishimoto M, Orimo A, Yamagishi T, Nabeshima Y, Kuro-o M, Nabeshima Y, Boon K, Keaveney M, Stunnenberg HG, Muramatsu M. 1998 UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra- embryonic cells. EMBO J.; 17(7):2019-32.

Okumura-Nakanishi S, Saito M, Niwa H, Ishikawa F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. J Biol Chem. 2005; 280(7):5307-17.

Olive, P.L., Banath, J.P., Durand, R.E. (1990). Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells using the “comet” assay. Radiat. Res *122,* 86 –94.

Park, S.M., Shell, S., Radjabi, A.R., Schickel, R., Feig, C., Boyerinas, B., Dinulescu, D.M., Lengyel, E., Peter, M.E. (2007). *Let-7* prevents early cancer progression by suppressing expression of the embryonic gene HMGA2. Cell Cycle *6,* 2585-2590.

Parsons, J.L., Tait, P.S., Finch, D., Dianova, I.I., Allinson, S.L., Dianov, G.L. (2008). CHIP-mediated degradation and DNA damage-dependent stabilization regulate base excision repair proteins. Mol. Cell ***29,*** 477-487.

Pearce JJ, Evans MJ. Mml, a mouse Mix-like gene expressed in the primitive streak. Mech Dev. 1999;87:189-192

Pera MF, Reubinoff B, Trounson A. Human embryonic stem cells. J Cell Sci. 2000; 113(Pt 1):5-10.

Prasad, R., Liu, Y., Deterding, L.J., Poltoratsky, V.P., Kedar, P.S., Horton, J.K., Kanno, S., Asagoshi, K., Hou, E.W., Khodyreva, S.N. *et al.* (2007). HMGB1 is a cofactor in mammalian base excision repair. Mol. Cell *27*, 829-841.

Prasad, R., Beard, W.A., Chyan, J.Y., Maciejewski, M.W., Mullen, G.P., Wilson, S.H. (1998). Functional analysis of the amino-terminal 8-kDa domain of DNA polymerase beta as revealed by site-directed mutagenesis. DNA binding and 5'-deoxyribose phosphate lyase activities. J. Biol. Chem. *273*, 11121-11126.

Rao M. Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells. Dev Biol. 2004; 275:269-86.

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene *277*, 63-81.

Reeves, R., and Adair, J.E. (2005). Role of high mobility group (HMG) chromatin proteins in DNA repair. DNA Repair (Amst.) *4*, 926-938.

Reeves R, Nissen MS. 1990. The AT-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. J Biol Chem 265:8573–8582.

Reeves R, Nissen MS. 1993. Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles. J Biol Chem 268:21137–21146.

Reeves R, Wolffe AP. 1996. Substrate structure influences binding of the non-histone protein HMG-I(Y) to free nucleosomal DNA. Biochemistry 35:5063–5074.

Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol. 2000; 18(4):399-404.

Richards M, Tan SP, Chan WC, Bongso A. Reverse serial analysis of gene expression (SAGE) characterization of orphan SAGE tags from human embryonic stem cells identifies the presence of novel transcripts and antisense transcription of key pluripotency genes. Stem Cells. 2006; 24(5):1162-73.

Richmond TJ, Davey CA. 2003. The structure of DNA in the nucleosome core. Nature 423:145–150.

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem. 2005; 280(26):24731-7.

Rogalla, P., Drechsler, K., Kazmierczak, B., Rippe, V., Bonk, U., Bullerdiek, J. (1997). Expression of HMGI-C, a member of the high mobility group protein family, in a subset of breast cancers: relationship to histologic grade. Mol. Carcinog. *19*, 153-156.

Saijoh, Y., Adachi, H., Sakuma, R., Yeo, C.Y., Yashiro, K., Watanabe, M., Hashiguchi, H., Mochida, K., Ohishi, S., Kawabata, 2000. Left-right asymmetric expression of lefty2 and nodal is induced by a signaling pathway that includes the transcription factor FAST2. Mol. Cell 5, 35–47.

Schier, A.F. (2003). Nodal signaling in vertebrate development. Annu. Rev. Cell Dev. Biol. 19, 589–621.

Schier, A.F., and Shen, M.M. 2000. Nodal signalling in vertebrate development. Nature 403, 385–389.

Sakaguchi M., Miyazaki M., Takaishi M., Sakaguchi Y., Makino E., Kataoka N., Yamada H., Namba M., Huh N. H. S100C/A11 is a key mediator of Ca(2+)-induced growth inhibition of human epidermal keratinocytes. J. Cell Biol. 2003;163:825–835.

Sakaguchi M., Miyazaki M., Sonegawa H., Kashiwagi M., Ohba M., Kuroki T., Namba M., Huh N. H. PKCalpha mediates TGFbeta-induced growth inhibition of human keratinocytes via phosphorylation of S100C/A11. J. Cell Biol. 2004;164:979–984.

Sakaguchi M., Sonegawa H., Nukui T., Sakaguchi Y., Miyazaki M., Namba M., Huh N. H. Bifurcated converging pathways for high Ca2+- and TGFbeta-induced inhibition of growth of normal human keratinocytes. Proc. Natl. Acad. Sci. USA. 2005;102:13921–13926.

Sakano, K., Oikawa, S., Hasegawa, K., and Kawanishi, S. (2001). Hydroxyurea induces site-specific DNA damage via formation of hydrogen peroxide and nitric oxide. Jpn. J. Cancer Res. *92*, 1166-1174.

Saretzki, G., Walter, T., Atkinson, S., Passos, J.F., Bareth, B., Keith, W.N., Stewart, R., Hoare, S., Stojkovic, M., Armstrong, L. *et al.* (2008). Downregulation of multiple stress defense mechanisms during differentiation of human embryonic stem cells. Stem Cells *26*, 455-464.

Sgarra R, Rustighi A, Tessari MA, Di Bernardo J, Altamura S, Fusco A, Manfioletti G, Giancotti V. 2004. Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. FEBS Lett 574:1–8.

Sgarra R, Tessari MA, Di Bernardo JD, Rustighi A, Zago P, Liberatori S, Armini A, Bini L, Giancotti V, Manfioletti G. 2005. Discovering high mobility group A molecular partners in tumour cells. Proteomics 5:1494–1506.

Sgarra, R., Furlan, C., Zammitti, S., Lo Sardo, A., Maurizio, E., Di Bernardo, J., Giancotti, V., Manfioletti, G. (2008) Interaction proteomics of the HMGA chromatin architectural factors. Proteomics 8(22): 4721-4732.

Singh, N.P., McCoy, M.T., Tice, R.R., and Schneider, E.L. (1988). A simple technique for quantitation of low levels of DNA damage in individual cells. Exp. Cell Res., *175:*184–191.

Smas, C.M. and Sul, H.S. (1993) Pref-1, a protein containing EGF-like repeats, inhibits adipocyte differentiation. Cell 73, 725– 734.

Smas CM, Sul HS. 1997Molecular mechanisms of adipocyte differentiation and inhibitory action of pref-1. Crit Rev Eukaryotic Gene Expression; 7: 281 -298.

Smith JR, Vallier L, Lupo G, Alexander M, Harris WA, Pederson RA. 2007. Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. Dev Bio 303, 107-117.

Sul HS, Smas CM, Mei B, Zhou L. 1997 Regulation of fat synthesis and adipose differentiation. Prog Nucleic Acid Res Mol Biol; 60: 317 -345.

Sul HS, Smas C, Mei B and Zhou L. 2000 Function of pref-1 as an inhibitor of adipocyte differentiation. Int J of Obesity 24(4), 15-19

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature 336(6200):688-90.

Smyth, G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3 (1).

Smyth, G.K., Michaud, J. and Scott, H. (2005) The use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 21(9), 2067–2075.

Smyth, G.K. and Speed, T.P. (2003) Normalization of cDNA microarray data. Methods 31, 265–273.

Sobol, R.W., Prasad, R., Evenski, A., Baker, A., Yang, X.P., Horton, J.K., Wilson, S.H. (2000). The lyase activity of the DNA repair protein beta-polymerase protects from DNA damage-induced cytotoxcity. Nature *405,* 807-8 10.

Spagnoli FM, Hemmati-Brivanlou A. Guiding embryonic stem cells towards differentiation: lessons from molecular embryology. Curr Opin Genet Dev. 2006; 16(5):469-75.

Subramanian, D., and Griffith, J.D. (2002). Interactions between p53, hMSH2-hMSH6 and HMG I(Y) on holliday junctions and bulged bases. Nucleic Acids Res. ***30,*** 2427- 2434.

Tallini, G., and Dal, C.P. (1999). HMGI(Y) and HMGI-C dysregulation: a common occurrence in human tumors. Adv. Anat. Pathol. *6*, 237-246.

Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006; 126(4):663-76.

Tan SM, Dröge P. 2005. Comparative analysis of sequence-specific DNA recombination systems in human embryonic stem cells. Stem Cells 23:868–873.

Tessari MA, Gostissa M, Altamura S, Sgarra R, Rustighi A, Salvagnao C, Caretti G, Imbriano C, Mantovani R, Del Sal G, Giancotti V, Manfioletti G. 2003. Transcriptional activation of the cyclin A gene by the architectural transcription factor HMGA2. Mol Cell Biol 23:9 104– 9116.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science. 1998; 282(5391):1145-7.

Thuault, S., Valcourt, U., Petersen, M., Manfioletti, G., Heldin, C.H. and Moustakas, A. (2006) Transforming growth factor-beta employs HMGA2 to elicit epithelial–mesenchymal transition. J. Cell. Biol. 174, 175–183.Verdine, G.L., and Norman, D.P. (2003). Covalent trapping of protein-DNA complexes. Annu. Rev. Biochem. *72*, 337-366.

Trounson, A. 2006 The production and directed differentiation of human embryonic stem cells. Endocr. Rev. 27, 208–219

Vallier, L., Reynolds, D., Pedersen, R.A., 2004a. Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. Dev. Biol. 275, 403–421.

Vallier, L., Rugg-Gunn, P.J., Bouhon, I.A., et al., 2004b. Enhancing and diminishing gene function in human embryonic stem cells. Stem Cells 22, 2–11.

Vallier, L., Alexander, M., Pedersen, R.A., 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell. Sci. 118,

4495–4509.

Wang J, Rao S, Chu J, et al. 2006 A protein interaction network for pluripotency of embryonic stem cells. Nature; 444:364–368.

Wang, T.H., Popp, D.M., Wang, H.S., Saitoh, M., Mural, J.G., Henley, D.C., Ichijo, H., Wimalasena, J. (1999). Microtubule dysfunction induced by paclitaxel initiates apoptosis through both c-Jun N-terminal kinase (JNK)-dependent and –independent pathways in ovarian cancer cells. J. Biol. Chem. *274*, 8208-8216.

Wei CL, Miura T, Robson P, Lim SK, Xu XQ, Lee MY, Gupta S, Stanton L, Luo Y, Schmitt J, et al. Transcriptome profiling of human and murine ESCs identifies divergent paths required to maintain the stem cell state. Stem Cells. 2005; 23(2):166-85.

Weinberger, M., Trabold, P.A., Lu, M., Sharma, K., Huberman, J.A., Burhans, W.C. (1999). Induction by adozelesin and hydroxyurea of origin recognition complex-dependent DNA damage and DNA replication checkpoints in saccharomyces cerevisiae. J. Bio. Chem. *274*, 35974-35984.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature336(6200):684-7.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., Song, E. (2007). *Let-7* regulates self renewal and tumorigenicity of breast cancer cells. Cell *131*, 1109-1123.

Zaidi, M.R., Okada, Y., and Chada, K.K. (2006). Misexpression of full-length HMGA2 induces benign mesenchymal tumors in mice. Cancer Res. *66*, 7453-7459.

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol. 2006; 8(10):1114-23.

Zhou, X., Benson, K.F., Ashar, H.R., and Chada K. (1995). Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature 376, 771-774.

VII APPENDIX

1. Li O, Vasudevan D, Davey CA, Dröge P. 2006, High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. Genesis 44(11):523-9.

2. Li O*, Li J*, Dröge P. 2007, DNA architectural factor and proto-oncogene HMGA2 regulates key developmental genes in pluripotent human embryonic stem cells. FEBS Lett. 581(18):3533-7

3. Summer H*, Li O*, Bao Q, Zhan L, Peter S, Sathiyanathan P, Henderson D, Klonisch T, Goodman SD, Dröge P. 2009, HMGA2 exhibits dRP/AP site cleavage activity and protects cancer cells from DNA-damage-induced cytotoxicity during chemotherapy. Nucleic Acids Res. 37(13):4371-84.

4. Pfannkuche K, Summer H, Li O, Hescheler J, Dröge P. 2009, The high mobility group protein HMGA2: a co-regulator of chromatin structure and pluripotency in stem cells? Stem Cell Rev Rep. 5(3):224-30.

* Equal contribution