International Journal of Biosciences (IJB)
ISSN: 2220-6655 (Print) 2222-5234 (Online)
Vol. 1, No. 3, p. 20-28, 2011
http://www.innspub.net

RESEARCH PAPER OPEN ACCESS

Investigation of molecular biology of leukaemia inhibitory factor (LIF): A Bioinformatics approach

Sanuwarul Kabir[1], Abdullah-Al-Emran[2*], Zinnia Naoshin[3], Farzana Ahmed[4]

[1]*Department of Genetic Engineering and Biotechnology, University of Dhaka, Bangladesh*

[2]*Departmet of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Santosh,Tangail, Bangladesh*

[3]*Department of Biochemistry, Primeasia University, Bangladesh*

[4]*Department of Mathematics and Natural Science, BRAC University, Dhaka, Bangladesh*

Received: 7 May 2011
Revised: 14 May 2011
Accepted: 15 May 2011

Key words: Leukemia inhibitory factor, bioinformatics, molecular properties, orthologs, paralogs.

Abstract

Leukaemia inhibitory factor (LIF) is a glycoprotein with a variety of functions in different organ systems. It has great impact on embryogenic stem cell as it is required to maintain the differentiated state of a cell. In our study we tried to find out various molecular features of LIF in human such as conserved domains, important transcription factors, sequence motifs, protein motifs, conserved miRNA binding site and single nucleotide polymorphisms with the help of bioinformatics tools. We also tried to generate a generalized downstream signaling pathway for LIF. These findings could be used to control the expression of LIF for treatment and better drug designing.

*Corresponding Author: Abdullah-Al-Emran ✉ emran_geb@yahoo.com

Introduction

The leukemia inhibitory factor (LIF) is a heavily glycosylated single chain polypeptide (Metcalf, 1992). The reported molecular weights of LIF range from 38,000 to 67,000 Daltons. Most, if not all, of this heterogeneity can be ascribed to extensive and variable glycosylation, since deglycosylation reduces the molecular weight to 20000 Daltons (Hilton *et al*., 1988). (*et al* must be italic all through the manuscript. Please correct). Intriguingly, although native LIF is heavily glycosylated, glycosylation is not actually required for biological activity, since *E. coli*-derived (nonglycosylated) recombinant human and murine LIF are biologically active both *in vitro* and *in vivo* (Gearing *et al*., 1987 and Metcalf *et al*., 1990). The core polypeptide predicted from cDNA cloning has a predicted molecular weight of -20000 Daltons (180 amino acid residues) and contains a number of potential N and O-linked glycosylation sites (Gearing *et al.,* 1987). LIF is encoded by a unique gene in the murine and human genomes, located at chromosomal band 22q12.1-12.2 in the human (Sutherland *et al.,* 1989) and llAl in the mouse (Kola *et al*., 1990). The LIF mRNA is 4.8 kb in length and is transcribed from a gene spanning 8 kbp.

LIF protein has an important role in reproduction. Several subsequent studies have shown that it is involved in the control of implantation also in human (Dunglinson *et al*., 1996 and Nachtigall *et al*., 1996). LIF is expressed in the human endometrium in a menstrual cycle-dependent manner. Maximal expression patterns were observed on days 19–25 of the menstrual cycle coinciding with time of blastocyst implantation (Arici *et al*., 1995). LIF also enhances blastocyst formation of human embryos and modulates trophoblast differentiation *in vitro*, and thus induces conditions necessary for implantation (Dunglinson *et al*., 1996).LIF is a produced by a number of different cell types *in vitro,* including stimulated T lymphocytes, monocytes, fibroblasts, astrocytes and various tumor-derived cell lines. LIF exerts a broad range of effects on diverse cell types and many of these actions can be mimicked by related cytokines, in particular interleukin 6 (IL-6), oncostatin M (OSM), ciliary neurotrophic factor (CNTF) and cardiotrophin-1 (CT-1) (Vogiags and Salamonsen, 1999).

LIF was initially known for its ability to induce differentiation of Ml myeloid leukaemia cells into macrophage like cells (Tomida *et al.,* 1984) and to suppress embryonic stem cell differentiation (Smith *et al.,* 1988). The poly-functionality of LIF led to its discovery and rediscovery by workers in quite separate disciplines: as a factor preventing differentiation commitment in normal embryonic stem cells (differentiation inhibitory activity (DIA))(Williams *et al*., 1988); as a factor able to switch signaling of autonomic nerves from an adrenergic to cholinergic mode (Yamamori *et al*., 1989) ; as an hepatocyte stimulating factor able to stimulate the production of acute phase proteins (hepatocyte stimulating factor 3) (Baumann and Wong, 1989); and as an inhibitor of lipoprotein lipase, blocking lipid transport to adipocytes (Mori *et al.,* 1989). The pleiotrophic and redundant functioning of cytokines such as LIF has been partly explained since the molecular characterization of the cytokine receptor systems. The activation of the LIF receptor (LIFR) complex (Figure 1) requires hetero-dimerisation of the two low affinity components LIFR and gp130 (Gearing *et al.,* 1991) and this combination generates a high affinity binding site (Gearing *et al*., 1992 and Taga *et al*., 1992). Interestingly, the LIFR and gp130 trans-membrane proteins have also been shown to serve as receptors for OSM, CNTF and CT-1, while an active IL-6 receptor complex requires gp130 (Murakami *et al*., 1993). The shared activities of the LIF receptor family suggest that other members may also be important in regulating blastocyst development and implantation.

Since Leukaemia Inhibitory Factor (LIF) has important role in embryonic stem cell differentiation and in reproduction; therefore aim of this study was to identify the orthologs, conserved motif, Transcription factor binding sites, SNPs, miRNA target of LIF and suggesting a possible pathway of LIF gene by the different tools of Bioinformatics. These findings will be helpful to understand the regulations of LIF.

Materials and methods

Nucleotide sequence and protein sequence were collected from NCBI (www.nibi.nlm.nih.gov) (Accession number:- gene-NG_008721, protein-NP_002300) and the SNP data from NCBI SNP database. Conserved Domain was found out from NCBI (http://www.ncbi.nlm.nih.gov/cdd). Orthologs were found using BLAST program of NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). TF-binding sites were scanned using Alibaba from TRANSFAC database(http://www.generegulation.com/pub/databases.html). Sequence motifs were analyzed using MEME (http://meme.sdsc.edu) and phosphotyrosine motif from NCBI. Conserved miRNA binding site was analyzed using target scan tool (http://www.targetscan.org/) Pathway was determined using KEGG pathway (http://www.genome.jp/kegg/pathway.html).

Table 1. List of orthologs.

Species Name	Identity (%)
Gorilla (*Gorilla gorilla*)	99
Macaque (*Macaca mulatta*)	99
Marmoset (*Callithrix jacchus*)	96
Mouse Lemur (*Microcebus murinus*)	92
Dog (*Canis familiaris*)	92
Dolphin (*Tursiops truncatus*)	90

Results

Looking for Conserve Domain and Orthologs

Conserved domains in proteins play a crucial role in protein interactions, DNA binding, enzyme activity, and other important cellular processes. Protein domains are often conserved across many species, and as such, they offer an interesting dataset in how genomes maintain them with relationship to other conserved domains, as well as to proteome size (Malek and Haft, 2001). So conserve domain of LIF was searched and found that conserved domain of LIF protein is 157 nucleotides long (Fig. 2).

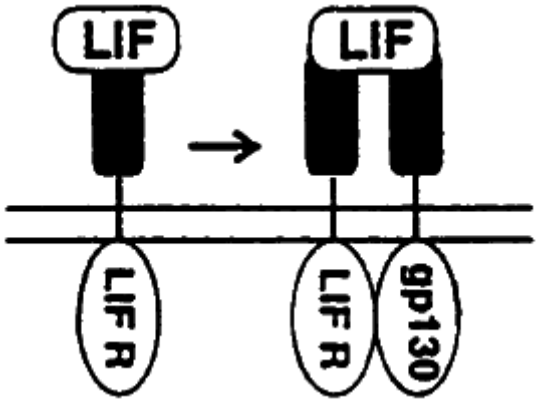

Fig. 1. Model of the formation of the LIF receptor complex. (Modified from Davis *et al*. 1993).

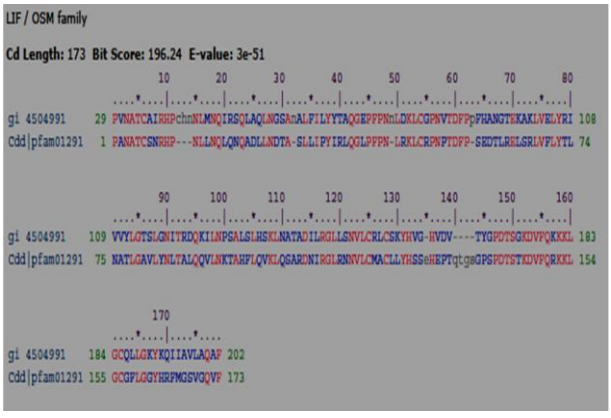

Fig. 2. Conserved Domain of LIF Protein.

Orthologs and paralogs are two fundamentally different types of homologous genes that evolved, respectively, by vertical descent from a single ancestral gene and by duplication. Orthology and paralogy are key concepts of evolutionary genomics (Koonin 2005).

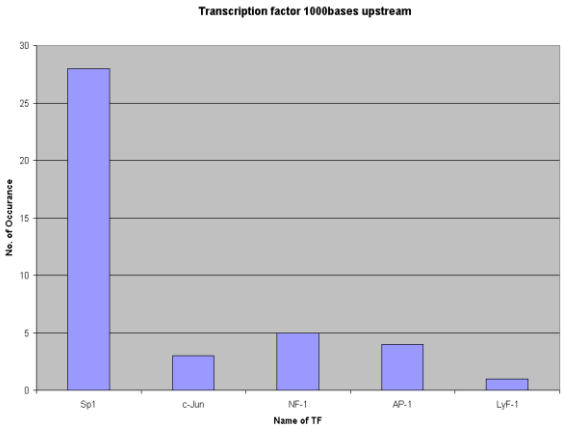

Fig. 3a. Transcription factors of LIF gene at 500 bases upstream region.

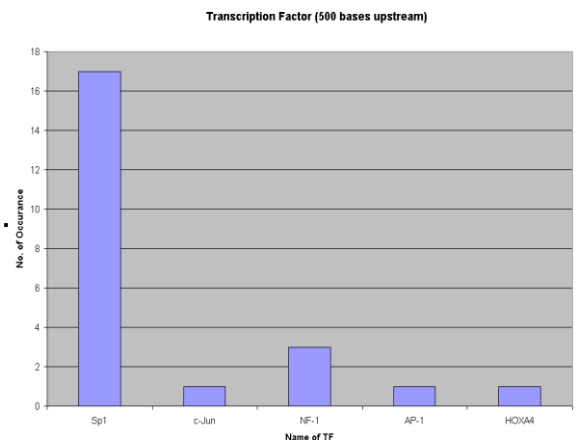

Fig. 3b. Transcription factors of LIF gene at 1000 bases upstream region.

In evolutionary terms, robust identification of orthologs is essential because otherwise any evolutionary scenarios are bound to be meaningless. Orthologs typically retain the same, ancestral function, which makes transfer of functional information within a set of orthologs generally reliable. The evolutionary basis of such conservation of function among orthologs appears fairly obvious (Koonin *et al.*, 2001). Considering this fact orthologs were searched and the list of orthologs is given Table 1 showing 90% or more identity. Orthologous sequences provide useful information in taxonomic classification and phylogenetic studies of organisms. The pattern of genetic divergence can be used to trace the relatedness of organisms. In this study no paralogs of LIF was found.

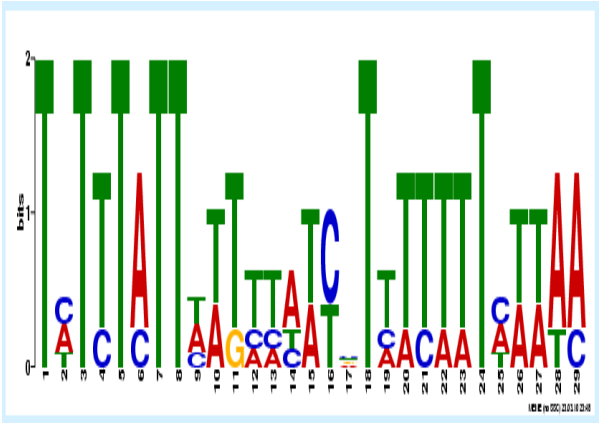

Fig. 4a. Motif of LIF Gene.

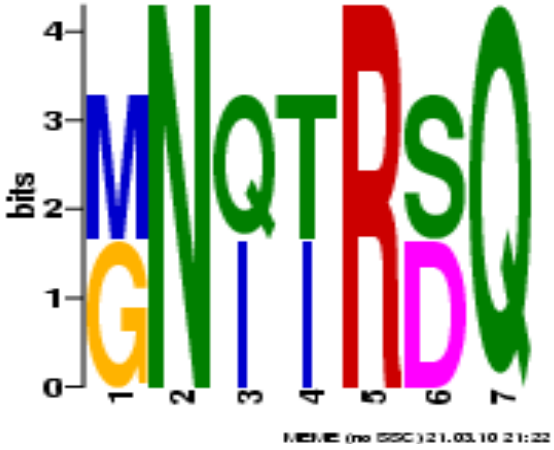

Fig. 4b. Motif of LIF protein.

Searching Transcription Factor (TF) binding sites

Understanding the regulatory networks of higher organisms is one of the main challenges of functional genomics. Gene expression is regulated by transcription factors (TF) binding to specific transcription factor binding sites (TFBS) in regulatory

regions associated with genes or gene clusters. Identification of regulatory regions and binding sites are prerequisite for understanding gene regulation (Sandve and Finn, 2006). In view of this, possible transcription factor binding sites for LIF was searched in these regions. The search result returned a number of transcription factor binding sites. It was found that in both cases, SP1 has the highest possible DNA binding site to LIF gene. It is figured out by excel sheet whereas other potential transcription factor with the occurrence number is also added (Fig. 3a and 3b).

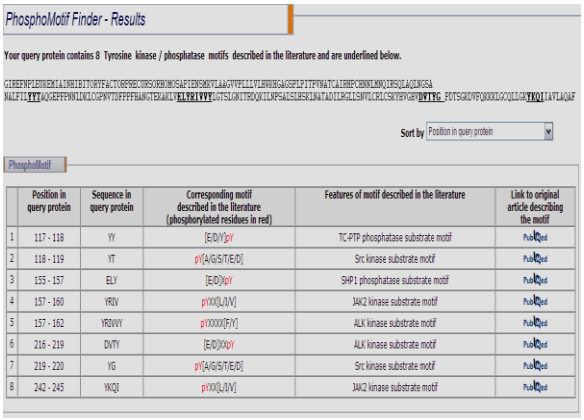

Fig. 4c. Phosphomotif of LIF protein.

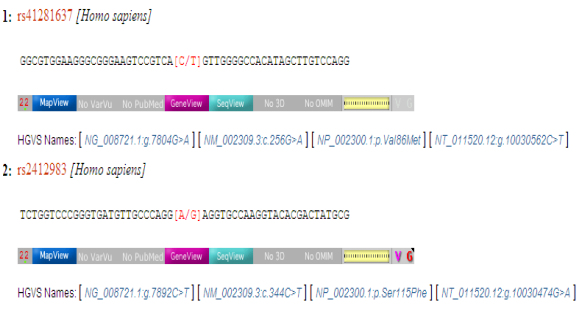

Fig. 5. SNP of LIF protein.

Looking for motifs in the upstream region

Motif is important part and termed as a regulatory element. Sequence motifs are short, recurring patterns in DNA that are involved in important processes at the RNA level, including ribosome binding, mRNA processing (splicing, editing, polyadenylation) and transcription termination (Dhaeseleer 2006). For proteins, a sequence motif is distinguished from a structural motif, a motif formed by the three dimensional arrangement of amino acids. They can be used to perform clustering, family classification, discovery of sub-families in large protein families, gene expression analysis and the study and discovery of homology relations (Sandve and Finn, 2006). Sequence motif was searched for LIF gene (Figure 4a) and LIF protein (Fig. 4b). We also searched for phosphotyrosine motif. In this search, presence of any of over 320 phosphorylation-based motifs curated from the literature in a protein of interest was reported (Fig. 4c).

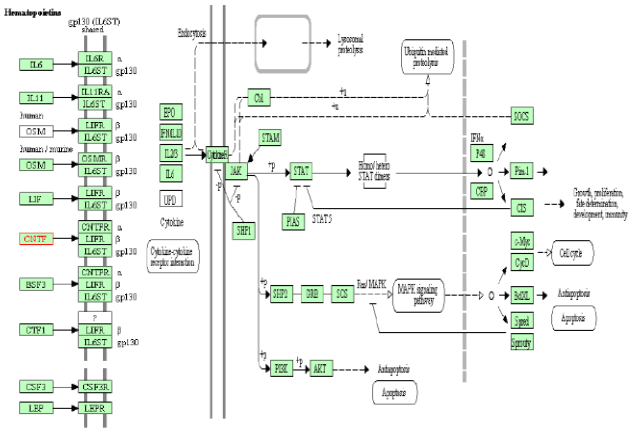

Fig. 6. KEGG pathway of LIF protein.

Analyzing Single Nucleotide Polymorphism (SNP) in human LIF

Single nucleotide polymorphisms provide a rich source of information about the evolutionary history of human populations (Stoneking 2001). They are partly responsible for individual differences in the effectiveness and tolerability of drugs. SNPs have therefore become one of the most important objects of medical research, especially as they could also provide clues to new targets (Chakravarti 2001). In this study SNPs in LIF gene, occurring in human population were tried to find out. Database analysis revealed two naturally occurring SNPs in LIF gene in human population (Fig. 5).

Suggesting possible pathway for LIF protein in human

Although it has been reported that mouse ES cells self-renew under leukemia inhibitory factor (LIF) (Okita and Yamanaka 2006), the intracellular signaling pathway activated by LIF in human ES cells is not yet clear (Yue *et al*., 2010).Considering the fact overall pathway and interacting protein list of LIF was generated using KEGG pathway designed for pathway maps for metabolism and other cellular processes, as well as human diseases; manually created from published materials (Fig. 6).

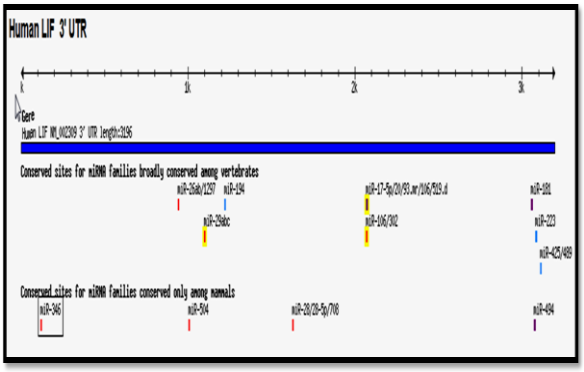

Fig. 7a. Conserved binding site of miRNA at the 3'UTR region of LIF gene.

Identification of possible miRNA target on LIF:

Small, noncoding RNA molecules known as microRNAs (miRNAs) have gained attention as regulatory molecules. These genomically encoded RNAs undergo several modifications before being converted into mature 21–23 base pair transcripts capable of gene silencing. Studies have also demonstrated that specific miRNAs regulate gene expression during germ line development and cellular differentiation (Hatfield, 2005). Using target scan tools conserved miRNA were identified that target the 3′ UTR of mature transcripts. and from multiple sequence alignment it was found that 5'-UACUUGA-3' sequence of LIF transcript at the 3'UTR region is conserved in the vertebrates Three miRNA hsa-mir-26a,hsa-mir-26b, hsa-mir1297 were found to be bind to the conserved region of LIF and thus controlling its function.

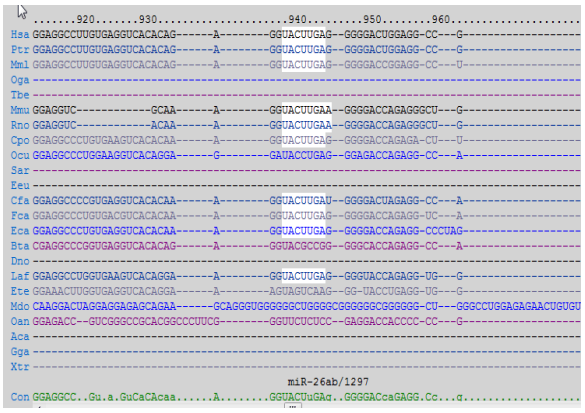

Fig. 7b. Multiple sequence alignment amongst vertebrate LIF gene showing conserved sequence and region.

Discussion

Leukemia inhibitory factor (LIF) is a polyfunctional molecule and act on a broad range of cells. In view of this, better understanding about the molecular properties will be beneficial. For example, the role of LIF in the implantation process is well established. Implantation remains the 'the last great frontier of reproductive biology. Implantation failure is a major reason of infertility in otherwise healthy women. Despite new treatment strategies, particularly in the area of in-vitro fertilization (IVF), only 11–17% of embryos grown *in vitro* achieve successful implantation (Imoedemhe *et al*., 1996 and Van der Elst *et al*., 1996). So detail knowledge about LIF at molecular level may provide an effective approach for therapeutic intervention. In this perspective, the molecular biology of LIF was analyzed using various bioinformatics tools.

In this study conserved domain in LIF was found which is also present in *Gorilla gorilla, Macaca mulatta*, *Callithrix jacchus*, *Microcebus murinus*, *Canis familiaris* and *Tursiops truncates* with 90 – 99 % identity. So this domain is very important and can be used for genetic manipulation.

The highest possible DNA binding site for LIF was found for Sp1 which is known as a human transcription factor involved in gene expression in the early development of an organism. Identification of such sites is not only relevant for locating the promoter of a gene, but they may also allow the prediction of a tissue specific gene expression pattern and responsiveness to known biological signaling pathways. As motif is an important regulatory element, motif was searched for LIF gene, LIF protein and phosphoserine motif. These findings can also be used to control the expression and function of LIF. In search for SNP, two SNPs were found which may be helpful in drug designing. To confirm the role of miRNA for regulation of LIF transcript we identified three miRNA that can bind on the 3'UTR region of LIF at a specific conserved sequence. These miRNA can be targeted for modulating the expression of LIF gene (up regulation or down regulation).

	predicted consequential pairing of target region (top) and miRNA (bottom)	seed match
Position 939-945 of LIF 3' UTR hsa-miR-1297	5' ...GAGGUCACACAGAGGUACUUGAG... 3' GUGGACUUAAUGAACUU	7mer-m8
Position 939-945 of LIF 3' UTR hsa-miR-26a	5' ...GAGGUCACACAGAGGUACUUGAG... 3' UCGGAUAGGACCUAAUGAACUU	7mer-m8
Position 939-945 of LIF 3' UTR hsa-miR-26b	5' ...GAGGUCACACAGAGGUACUUGAG... 3' UGGAUAGGACUUAAUGAACUU	7mer-m8

Fig.7c. A close look at the position and the predicted target of three miRNA on the 3'UTR region of LIF.

LIF plays different roles in human and mouse ES due to structural similarity. Considering this fact, a generalized signaling pathway of LIF in human ES was generated. This pathway would provide valuable information about important proteins for mediating the function of LIF. Knowledge of these proteins will be very useful for therapeutic intervention. Profound investigation of LIF will provide deeper insights into the function and regulation of LIF expression as well as an effective approach for treatment of many health problems like female infertility.

References

Arici A, Engin O, Attar E. 1995. Modulation of leukaemia inhibitory factor gene expression and protein biosynthesis in human endometrium. J. Clin. Endocrinol. Metab. **80,** 1908–1915.

Baumann H and Wong GG. 1989. Hepatocyte-stimulating factor III shares structural and functional identity with leukemia-inhibitory factor. J. Immunol. **143,** 1163-1167.

Chakravarti A. 2001. Single nucleotide polymorphisms: to a future of genetic medicine. Nature **409,** 822-823.

Dunglinson GF, Barlow DH, Sargent IL. 1996. Leukaemia inhibitory factor significantly enhances the blastocyst formation rates of human embryos cultured in serum-free medium. Hum. Reprod. **11,** 191–196

Dhaeseleer P. 2006. What are DNA sequence motifs? Nature Biotechnology **24**, 423 – 425.

Gearing DP, Comeau MR, Friend DJ, Gimpel SD, Thut CJ, McGourty J, Brasher KK. 1992. The IL-6 signal transducer, gp130: an oncostatin receptor and affinity converter for the LIF receptor. Science **255**, 1434–1437.

Gearing DP, Gough, NM, King JA, Hilton DJ, Nicola NA, Simpson RJ, Nice EC, Metcalf D. 1987. Molecular cloning and expression of cDNA encoding a murine myeloid leukemia inhibitory factor (LIF). EMBO J. **6**, 3395-4007.

Hilton DJ, Nicola NA, Gough NM, Metcalf D. 1988. Resolution of three distinct factors produced by Krebs ascites cells which have differentiation – including activity on murine myeloid leukemia cell lines. J. Biol. Chem. **263,** 9238 – 9243

Imoedemhe, D.A., Sigue, A.B., Pacpaco, E.L. 1995. A comparative analysis of embryos derived from routine in-vitro fertilization and subzonal microinsemination. Hum. Reprod. **10**, 2970–2975.

Kola I, Davey A, Gough NM. 1990. Localization of the murine leukemia inhibitory factor gene near the centromere on chromosome 11. Growth Factors **2**, 235-240.

Koonin EV. 2005. Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet. **39**, 309-38.

Koonin EV, Galperin Michael Y, Norwell MA. 2003. Sequence-Evolution - Function Computational Approaches in Comparative Genomics, Kluwer Academic Publishers.

Malek JA and Haft DH. 2001. Conserved protein domains are maintained in an average ratio to proteome size. Genome Biology **2**, 80-90.

Metcalf D. 1992. Leukemia Inhibitory Factor-A Puzzling Polyfunctional Regulator. Growth Factors **7**, 169-173.

Metcalf D and Gearing DP. 1990. A myelosclerotic syndrome in mice engrafted with cells producing high levels of leukemia inhibitory factor (LIF). Leukemia **3,** 847-852

Mori M, Yamaguchi K, Abe K. 1989. Purification of a lipoprotein lipase-inhibiting protein produced by a melanoma cell line associated with cancer cachexia. Biochem. Biophys. Res. Commun. **160,** 1085-1092.

Murakami M, Hibi M, Nakagawa N, Nakagawa T, Yasukawa K, Yamanishi K, Taga T & Kishimoto T. 1993. IL-6 induced homodimerization of gp130 and associated activation of a tyrosine kinase. Science **260,** 1808–1810.

Nachtigall MJ, Kliman HJ, Feinberg RF. 1996. The effect of leukaemia inhibitory factor (LIF) on trophoblast differentiation: A potential role in human implantation. J. Clin. Endocrinol. Metab. **81**, 801–806.

Okita K and Yamanaka S. 2006. Intracellular signaling pathways regulating pluripotency of embryonic stem cells. Curr Stem Cell Res Ther. **1(1),** 103-11.

Sandve GK and Finn DF. 2006. A survey of motif discovery methods in an integrated framework. Biology Direct **1,** 11.

Smith AG, Heath JK, Donaldson DD. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature **336,** 688-690.

Stoneking M. 2001. Single nucleotide polymorphisms: from the evolutionary past Nature **409,** 821-822.

Sutherland GR, Baker E, Hyland VJ, Callen DF. 1989. The gene for human leukemia inhibitory factor (LIF) maps to 22q12. Leukemia **3**, 9-13.

Taga T and Kishimoto T. 1992. Cytokine receptors and signal transduction. FASEB Journal **6,** 3387–3396.

Tomida M., Yamamoto-Yamaguchi Y, Hozumi M. 1984. Purification of a factor inducing differentiation of mouse myeloid leukemic M1 cells from conditioned medium of mouse fibroblast L929 cells. J. Biol. Chem. **259,** 10978-10982

Vogiags D and Salamonsen LA. 1999. The role of leukaemia inhibitory factor in the establishment of pregnancy. J Endocrinol. **160(2),** 181-90.

Van der Elst J, van den Abbeel E, Camus M. 1996. Long-term evaluation of implantation of fresh and cryopreserved human embryos following ovarian stimulation with buserelin acetate-human menopausal gonadotrophin (HMG) or clomiphene citrate-HMG. Hum. Reprod. **11,** 2097–2106.

Williams RL, Hilton DJ, Pease S. 1988. Myeloid leukemia inhibitory factor (LIF) maintains the developmental potential of embryonic stem cells. Nature **336,** 684-687.

Yamamori T, Fukada K, Aebersold R, Korsching S, Fann MJ. 1989. The cholinergic neuronal differentiation factor from heart cells is identical to leukemia inhibitory factor. Science **246,** 1412-1416.

Yue Xu, Xiuwen Zhu, Heung Sik Hahm, Wanguo Wei, Ergeng Hao. 2010. Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. Proc Natl Acad Sci U S A. 2010 May 4.

.