Short Communication

Computational interrogation of cis-regulatory elements of genes that are common targets of luteotropin and luteolysin in the primate corpus luteum

Padmanaban S. Suresh [a],*, Thejaswini Venkatesh [b]

[a] Centre for Biomedical Research, Vellore Institute of Technology University, Vellore, India
[b] Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India

ARTICLE INFO

ABSTRACT

The rapid recent increase in microarray-based gene expression studies in the corpus luteum (CL) utilizing macaque models gathered increasing volume of data in publically accessible microarray expression databases. Examining gene pathways in different functional states of CL may help to understand the factors that control luteal function and hence human fertility. Co-regulation of genes in microarray experiments may imply common transcriptional regulation by sequence-specific DNA-binding transcriptional factors. We have computationally analyzed the transcription factor binding sites (TFBS) in a previously reported macaque luteal microarray gene set ($n = 15$) that are common targets of luteotropin (luteinizing hormone (LH) and human chorionic gonadotropin (hCG)) and luteolysin (prostaglandin (PG) $F_2\alpha$). This in silico approach can reveal transcriptional networks that control these important genes which are representative of the interplay between luteotropic and luteolytic factors in the control of luteal function. Our computational analyses revealed 6 matrix families whose binding sites are significantly over-represented in promoters of these genes. The roles of these factors are discussed, which might help to understand the transcriptional regulatory network in the control of luteal function. These factors might be promising experimental targets for investigation of human luteal insufficiency.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The corpus luteum (CL) is a transient endocrine gland that develops following ovulation and is an important contributor for the synthesis and secretion of progesterone from the ovary (reviewed by Niswender et al., 2000). The secreted progesterone (P$_4$) is required for the establishment and maintenance of pregnancy in higher mammals. The normal physiology of reproductive cycles depends upon the fully functional CL, and luteal phase dysfunction may lead to premature regression and thereby causing infertility (Devoto et al., 2009). Understanding the molecular signature of luteal tissue in response to luteotrophic support (luteinizing hormone (LH) and human chorionic gonadotropin (hCG)) during different functional states and after different in vivo experimental manipulation is crucial for evaluating the normal physiology and pathophysiology of human reproduction. In this regard, various studies have been conducted in clinically relevant non-human primate models and recent advances in microarray technology and the emergence of the Affymetrix™ Gene Chip® Rhesus MacaqueGenome Array has facilitated to investigate the gene expression profile globally (reviewed by Bishop et al., 2011). In recent times, utilizing the high-throughput methodologies, many interesting reports deciphered the transcriptome of

luteal tissue during different functional stages, in response to LH, hCG, and prostaglandin (PG)F$_2$ α (reviewed by Bishop et al., 2011). These studies compared genome wide changes in the non-human primate luteal tissue and identified important key regulatory luteotrophic factors and luteolysins. There is an increase in the repository of these high-throughput expression data in publicly available data bases such as the Gene Expression Omnibus (GEO). Mining of these data bases and further analysis may enrich the available knowledge on the impacted gene pathways of luteal function. Expressions of genes are tightly regulated by transcription factors that interact with DNA cis-regulatory modules (Spivakov et al., 2012). Transcription factors preferentially bind to specific sequence patterns (transcription factor binding sites). These cis elements or transcription factor binding sites are usually short in length and mostly present in the promoter region of genes (Spivakov et al., 2012). Elucidation of the relationship between transcriptional factors, transcriptional binding sites and their control genes which are differentially regulated in luteal tissue during different functional states is of great importance. In view of the above, determination of transcriptional factor binding sites over-representation in a data base containing differentially regulated genes in theca cells (untreated PCOS (polycystic ovarian syndrome) vs untreated normal) of the ovary has been reported earlier (Sarkar and Maitra, 2008). Priyanka et al. (2009) reported a list of co-regulated genes that are common targets of both luteotropin (LH, hCG) and luteolysin (PGF$_2$α) in CL of bonnet monkey by GeneChip Rhesus Macaque Genome Array. Furthermore, it is established from this study that the cross talk between luteotropic and luteolytic pathways plays an important role in the control of luteal gene expressions that are critical for its function and its demise (Priyanka et al., 2009). Computational identification of transcription factor binding sites over-representation in these co-regulated genes may help to find the common regulatory mechanism of CL function. In silico analysis of cis-acting elements of these targets (co-expressed genes) can be promising targets for experimental investigation of human luteal phase dysfunction. In the above context, we sought to investigate computationally the transcription factors over-representation in a set of genes that are common targets of both luteotropin and luteolysin in bonnet monkey CL as described by the group of Priyanka et al. (2009).

2. Materials and methods

2.1. Data set

Luteal genes (n = 15) that are common targets of luteotropin (LH or hCG) and luteolysin (PGF2α) were chosen from the previous study of Priyanka et al. (2009). The patterns of expressions of these genes in bonnet monkey CL were well characterized at the whole genome level and qPCR analysis after different treatments (Priyanka et al., 2009). Priyanka et al. (2009) utilized AffymetrixTM RhesusGeneChip® expression array for the analysis of gene expression changes in bonnet monkey CL. There exists 97–99% identity between available select genes of both rhesus and bonnet monkeys (reviewed by Bishop et al., 2011). With the availability of only rhesus macaque (*Macaca mulatta*) genome (Gibbs et al., 2007), in silico analysis of promoter regions of the experimentally validated bonnet monkey luteal genes was performed selecting *M. mulatta* in the commercially available software Genomatix Suite (Genomatix Software GmbH, Munich, Germany; http://www.genomatix.de). This online tool assumes that binding sites for transcription factors should be greater for transcription factors that regulate the co-expressed genes (Cartharius et al., 2005).

2.2. Computational promoter analysis

Analysis was done using the Gene2Promoter application, part of the GenomatixSuite (Genomatix Software GmbH, Munich, Germany; http://www.genomatix.de). Promoter sequences of the 15 genes (*INHA, SCARB1, SCD, INHBA, FST, StAR, MOCOS, CYP19A, A2M, CRHBP, DCN, CTGF, ACVRIB, PKIA,* and *SERPINE1*) were retrieved and promoter sequences were generally assumed as 500 bp upstream and 100 bp downstream of the transcriptional start site. This is the default setting in the Gene2 promoter while analyzing the promoter sequences. Gene symbols were given as input and the following information of each gene such as the mapping with chromosome, contig and positions and mapping quality, promoters and alternative transcripts for each locus were retrieved (http://www.genomatix.de). Retrieved promoter sequences were analyzed for their quality levels. According to the software, the quality levels

Table 1
List of analyzed genes which are targets of both luteotropin and luteolysin in the primate corpus luteum.

Group A genes upregulated by luteotropin and downregulated by luteolysin	Group B genes downregulated by luteotropin and upregulated by luteolysin

INH A
  (Genomatix promoter ID: GXP_1045711)
  Inhibin-α
	A2M
	  (Genomatix promoter ID: GXP_1041782)
	  Alpha-2-macroglobulin

SCARB1
  (Genomatix promoter ID: GXP_1666349)
  Scavenger receptor class B, member 1
	DCN
	  (Genomatix promoter ID: GXP_1043552)
	  Decorin

SCD
  (Genomatix promoter ID: GXP_1696686)
  Stearoyl-CoA desaturase
	CRHBP
	  (Genomatix promoter ID: GXP_1045711)
	  Corticotropin releasing hormone binding protein

INHBA
  (Genomatix promoter ID: GXP_1685998)
  Inhibin beta A
	CTGF
	  (Genomatix promoter ID: GXP_1080627)
	  Connective tissue growth factor

FST
  (Genomatix promoter ID: GXP_2902968)
  Follistatin
	ACVR1B
	  (Genomatix promoter ID: GXP_1951887)
	  Activin receptor type-1B

STAR
  (Genomatix promoter ID: GXP_3515467)
  Steroidogenic acute regulatory protein
	PKIA
	  (Genomatix promoter ID: GXP_1045711)
	  Protein kinase (cAMP-dependent, catalytic) inhibitor α

MOCOS
  (Genomatix promoter ID: GXP_3499434)
  Molybdenum cofactor sulfurase
	SERPINE1
	  (Genomatix promoter ID: GXP_3515805)

CYP19A1
  (Genomatix promoter ID: GXP_2904509)
  Cytochrome P450, family 19, subfamily A, polypeptide 1	