Indian Journal of Animal Sciences **84** (2): 140–145, February 2014/Article

Bioinformatics and functional analysis of proteins in serum of early pregnant buffaloes

A K BALHARA[1], A K MOHANTY[2], MEENAKSHI GUPTA[3], MANU JAMWAL[4], SUNESH[5], S K PHULIA[6] and INDERJEET SINGH[7]

Central Institute for Research on Buffaloes, Hisar, Haryana 125 001 India

Received: 3 September 2013; Accepted: 11 October 2013

ABSTRACT

Bioinformatics using different web based resources was done to classify 2–D gel resolved novel/ up- or down-regulated proteins (n=48) detected in serum of early pregnant buffaloes and find out interactions among them. Proteins were searched for their gene symbols on the ExPASy and the NCBI web portals and then PANTHER software was used to classify them. These proteins were primarily involved in regulatory, catalytic, immune regulation, cell differentiation and transporter functions. Among molecular functions, majority of proteins was involved in either catalytic or binding activities (29.5% each); receptor activity (13.6%), transporter activity (9.1%), enzyme regulator (6.8%) and structural molecule activity (6.8%) besides ion-channel and antioxidant activities. Based on biological functions, maximum (17.6%) proteins were found involved in metabolic function, followed by cellular processes (15.7%). Almost 10.8% proteins were involved in cell communications and 9.8% in transport functions, the rest were involved in immune and developmental processes. To these proteins 23 pathways were assigned, most predominantly Wnt signalling, inflammation - mediated by chemokine and cytokine signalling, and blood coagulation pathways. Other proteins were involved in as many as 20 pathways. A single layer extension to make connections for preparing protein-protein interaction map and network analysis of these bovine proteins, revealed that 9 of these proteins were interacting with each other in 6 different ways - interactions affecting molecular transportation, level of expression, regulation of expression, binding of proteins, interactions resulting in direct regulation and direct regulation inhibition and those leading to modification of the proteins. These findings suggested that the proteins identified with up-regulation, down-regulation or specific appearance at a particular stage during early buffalo pregnancy, primarily act interdependently in the functions involving cellular growth and differentiation.

Key words: Bioinformatics, Buffalo, Early pregnancy, Serum

Detection of protein interactions and function are key to identifying protein networks and understanding cellular processes. It is well accepted that genes in all eukaryotes mostly have same basic function, and therefore the Gene Ontology Consortium proposed uniform vocabulary to tag genes with functions (GO; Ashburner *et al.* 2000). GO based classification of proteins in this way offers mechanism to comprehend the function in other unrelated species. Software is available that classifies genes by their functions, using published scientific experimental evidence and evolutionary relationships to predict function even in the absence of direct experimental evidence. Proteins can be classified according to molecular function, biological process, cellular components and pathways. The interactions and functions of the proteins, specifically identified in 2-D gels of early pregnant buffalo serum are reported in this paper so that these can be identified in relation to biological processes involved in early embryonic development and progeny.

Present address: [1]Scientist (balharaak@gmail.com), [6]Senior Scientist (sphulia@gmail.com), Buffalo Physiology and Reproduction Division, [5]Scientist (balharas@gmail.com), AKMU, [7]Director (inderjeet.dr@gmail.com). [2]Principal Scientist (ashokmohanty1@gmail.com), [4]Senior Research Fellow (manujamwal25@gmail.com), Animal Biotechnology Centre (ABTC), National Dairy Research Institute, Karnal 132001. [3]Professor and Head (meenakshi415@gmail.com), Department of Veterinary Physiology and Biochemistry, LLR University of Veterinary and Animal Sciences, Hisar 125004.

MATERIALS AND METHOD

Functional analysis: The 48 proteins identified by MASCOT search in spots from 2D gels were taken for functional analysis. Gene symbols were generated for classification of identified proteins based on the protein analysis through evolutionary relationships (PANTHER) classification system available at http://www.pantherdb.org.

The energy pathways and metabolism categories were combined into a single category of biological process and functions of classified proteins. Proteins were also classified according to cellular component and protein class. To build a protein-protein interaction map and network analysis of these bovine proteins, protein-protein interaction data were extracted from the Human Protein Reference Database (HPRD; www.hprg.org) and the NCBI databases (National Centre for Biotechnology Information, USA) (www.ncbi.nlm.nih.gov). Only one layer was extended to make the connections.

RESULTS AND DISCUSSION

Functional analysis using bioinformatics: The 48 proteins identified from different databases were further analyzed for Gene ontology (GO) for functional analysis and protein-protein interaction (Table 1).

The clasification of proteins based on (A) molecular function, (B) biological function, (C) cellular components and (D) protein class are given in Fig. 1. Among molecular functions, maximum proteins were equally (29.5% each) involved in either catalytic or binding activities; followed by 13.6% in receptor activity, 9.1% in transporter activity, 6.8% each in enzyme regulator as well as structural molecule activity and only 2.3% proteins each were involved in ion channel and antioxidant activities. Based on the biological function classification, maximum 17.6% proteins were found involved in metabolic function, followed by 15.7% in cellular processes, 10.8% in cell communications, and 9.8% in transport functions. The rest of the proteins were involved in immune system processes (8.8%), system processes (7.8%), developmental processes (6.9%), response to

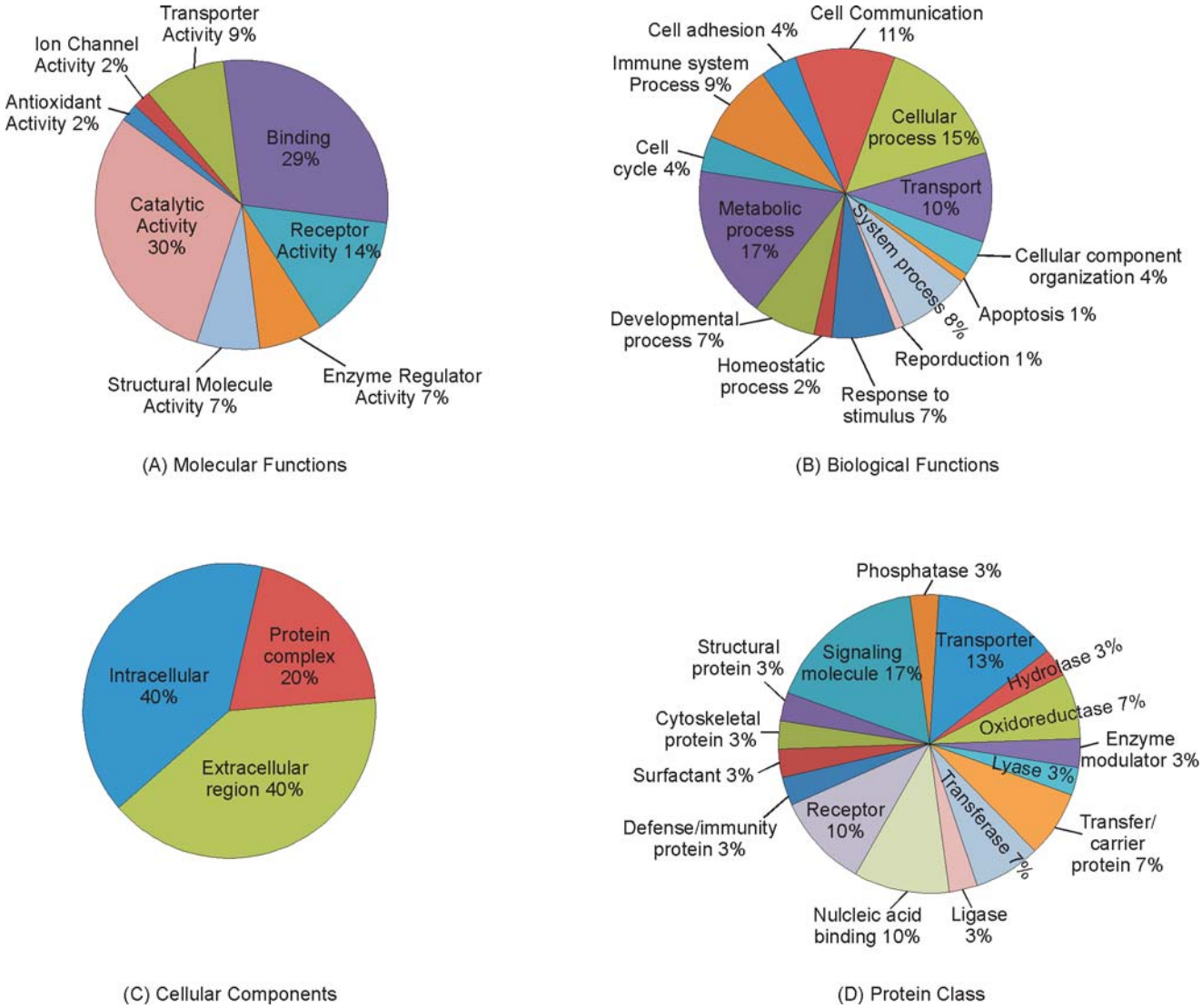


Fig. 1. Classification of identified proteins based on particular characteristics

Table 1. Proteins corresponding to differentially expressed 2-D gel spots of early pregnant buffaloes

Spot No.	Identified protein	Gene name/ symbol	Species
2	Serum albumin precursor	ALB	*Bos taurus*
4	Apolipo-protein A – II	APOA2	*Bos taurus*
5	Synaptojanin-1 (EC 3.1.3.36) (synaptic insitol-1,4,5-trisphosphate 5-phosphatase 1 p150)	SYNJ1	*Bos taurus*
6	Keratin, Type 1 cytoskeletal 10	KRT10	*Bos taurus*
8	1 phosphatidylinisitol-4, 5-biphosphate phosphodiesterase	PLCB1	*Bos taurus Monodelphis*
10	Small nuclear ribonucleoprotein-associated protein B'	SNRPB	*domestica*
13	Serum Albumin	ALB	*Bos indicus*
14	Serum albumin precursor	ALB	*Bos taurus*
17	Ferrochelatase	FECH	*Bos taurus*
18	Protein regulator of cytokinesis 1	PRC1	*Bos taurus*
21	Hypothetical protein	Not available	*Bos taurus*
22	Hypothetical protein (Receptor-type tyrosine-protein phosphatase-like N)	PTPRN	*Bos taurus*
23	MHC Class I antigen	Ceni-A	*Cervus nippon* (sika deer)
25	Collagen alpha-1(II) chain	COL2A1	*Bos taurus*
26	Serotransferrin	TF	*Bos taurus*
27	Immunoglobulin heavy chain variable region	Not found	*Bos taurus*
28	Tuberous sclerosis 2 protein (TSC2)	TSC2	*Bos taurus*
29	Putative glucose transporter isoform 1 (EC 6.3.2.19)	GLUT1	*Bubalus bubalis*
31	Complement C1q subcomponent subunit B	CIQB	*Bos taurus*
33	Cannot be retrieved	Not known	Not known
34	5'-nucleotidase, cytosolic IB	NT5CIB	*Bos taurus*
35	Serum albumin precursor or serum albumin	ALB	*Bos taurus*
36	Complement component C3d	Not available	*Bos taurus*
37	DNA polymerase alpha catalytic subunit	POLA2	*Bos taurus*
38	Interleukin – 27 subunit- alpha (IL27RA) protein	IL27RA	*Bos taurus*
40	Acyl CoA Synthetase	ACS	*Bos taurus*
41	IgG 3 heavy chain constant region	IGHG3	*Bos taurus*
42	Hypothetical protein similar to Acyl-CoA dehydogenase very long chain specific mitochondrial precursor (VLCAD) isoform	Not available	*Canis famililiaris*
43	unclassified	Not known	Not known
44	Glutathione S-transferase	GSTA3	*Bos taurus*
45	Cyclic nucleotide-gated cation channel beta-1	CNGB1	*Bos taurus*
46	Transforming growth factor beta-1 precursor	TGFB1	*Bos taurus*
47	Hypothetical protein	A1A4K4_F2	*Bos taurus*
48	Peroxidasin homolog	PXDN	*Homo sapiens*
50	Protocadherin-12	PCDH12	*Homo sapiens*
51	Hydroxysteroid (17-beta) dehydrogenase 8 (FabG-like protein)	HKE6	*Bos taurus*
53	SLAIN motif-containing protein 2	SLAIN2	*Bos taurus*
54	U4/U6 small nuclear ribonucleoprotein Prp3 protein or PRPF3 protein	PRPF3	*Bos taurus*
55	6-phosphogluconolactonase	PGLS	*Bos taurus*
56	Apoliloprotein B	APOB	*Diceros bicornis* (Black rhinoceros)
57	Complement C3	C3	*Bos taurus*
58	Complement C3	C3	*Bos taurus*
59	Von Willebrand factor	vwf	*Crocuta crocuta* (Spotted Hyena)
60	MHC Class I heavy chain	Not available	*Bos taurus*
62	Apolipoprotein A-I precursor	APOA1	*Bos taurus*
63	Hypothetical protein	CKM	*Bos taurus*
64	Origin recognition complex subunit 3	ORC3	*Bos taurus*
65	Hypothetical protein	Not available	*Bos taurus*

stimulus (6.9%), cellular component organization (3.9%), cell cycle (3.9%) and cell adhesion (3.9%), homeostatic process (2%), and 1% each in apoptosis and reproduction. The proportion of extracellular and intracellular proteins was equal at 40% each, while the remaining 20% proteins were classified as reportedly protein complexes.

Classification based on protein class of the identified proteins revealed that the signalling molecules constituted 16.7% of the total identified proteins, transporters 13.3%, receptor and nucleic acid binding proteins 10.0% each, transfer / carrier proteins, transferases and oxidoreductases 6.7% each, and the remaining 29% proteins were equally classified as phosphatases, hydrolases, enzyme modulators, lyases, ligases, defense / immunity proteins, surfactants, cytoskeletal proteins and structural proteins at a frequency of 3.3% each.

For the proteins identified in the present study, 23 pathways were assigned. The identified proteins, most predominantly featured in 3 pathways at a frequency of 7.7% each, viz. Wnt signalling pathway, inflammation mediated by chemokine and cytokine signalling pathway, and the blood coagulation pathways (Fig. 2). Further at 3.8% frequency each, remaining proteins were involved in 20 different pathways.

A protein-protein interaction map and network analysis of the identified proteins revealed that at least 9 proteins are interacting amongst themselves (Fig. 3) - Apolipo-protein A – II, Apolipo-protein A-I precursor, collagen alpha-1(II) chain, Von Willebrand factor, serum albumin, serotransferrin, complement C3, solute carrier family 2, member 1(SLC2A1) or the glucose transporter type 1 (GLUT1) and the tuberous sclerosis 2 protein (TSC2).

Of all the proteins identified, as many as 48 proteins showed homology with proteins from human database. Subsequent classification along similar lines suggested that most of the proteins were involved in regulatory, catalytic, cell differentiation and transporter functions. This is an expected outcome of the present study since early pregnancy is a complex process where both maternal and fetal biomolecules interact and bring necessary adjustments to address immunological, nutritional and developmental demands of the young one, while maintaining sound health of the mother as well. Of the 23 pathways identified for these proteins, highest 7.7% proteins were affiliated to Wnt signalling pathway – an embryogenesis specific pathway (Peifer and Polakis 2000). Other pathways like inflammation mediated by chemokine and cytokine signalling pathway, angiogenesis, insulin/IGF pathway-protein kinase B signalling cascade and heme biosynthesis are general pathways appearing during early pregnancy.

The results on the protein-protein interaction data from current experiment, described by network of 9 proteins (Fig. 3), strengthen our belief that the significantly affected pregnancy associated proteins are closely related. These interactions (6 types in total) were derieved from the available data on the Human Protein Reference Database (HPRD; www.hprg.org) and the NCBI database (National Centre for Biotechnology Information, USA; www.ncbi.nlm.nih.gov), as described below:

1. Interactions affecting molecular transportation: Complement C3 (Cui *et al*. 2007) and serotransferrin (Martin *et al*. 1998) positively affect molecular transportation of the glucose transporter protein.
2. Interactions affecting level of expression: Serum Albumin affects the expression levels of serotransferrin (Balagopalakrishna *et al*. 1999), Apolipoprotein A-I (Wilcox *et al*. 1991) and Apolipoprotein A-II (Sakai *et al*. 2000). Serotransferrin affects the expression of complement C3 (Tang *et al*. 2001), whereas tuberous sclerosis 2 protein (TSC2) affects the expression level of the glucose transporter protein (Buller *et al*. 2008).
3. Interactions affecting regulation of expression: Expression of serum albumin is regulated by serotransferrin (Worrall *et al*. 1998), collagen alpha-1(II) chain protein (Hsieh-Bonassera *et al*. 2009), apolipoprotein A-I (Zensi *et al*. 2010) and the Von Willebrand factor (Berny *et al*. 2008). Simultaneously, albumin also blocks regulation by serrotransferrin. Apolipoprotein A-II also negatively regulates apolipoprotein A-I expression (Gao *et al*. 2009).
4. Interactions affecting binding of proteins: Serum albumin directly affects binding of apolipoprotein A-I (Hoofnagle and Heinecke 2009).
5. Interactions resulting in direct regulation and direct regulation inhibition: Apolipoprotein A-I and apolipoprotein A-II are involved in this type of interaction with the later showing control over the former protein (Gao *et al*. 2009, Wroblewska *et al.* 2009).
6. Interactions leading to modification of the proteins: Serotransferrin caused protein modification in serum albumin (Shin and Osborne 2003).

It is worth mentioning here that these interactions are deduced as per interactions between identical proteins in human database. Since bovine database is not that voluminous and the buffalo database is almost non-existent, it is not possible to check whether the proteins interact in a similar fashion or not in buffalo.

Gene ontology based functional analysis is an important tool to study functional aspects of the data generated in a proteomics experiment. In the present study, the proteins supposed to be pregnancy influenced were functionally found to be involved in regulatory, catalytic, immune regulation, cell differentiation and transporter functions. Further, the expression of proteins important for embryo development, and hence successful conception, was predominantly observed e.g. expressions of proteins involved in either catalytic or binding activities, receptor activity, transporter

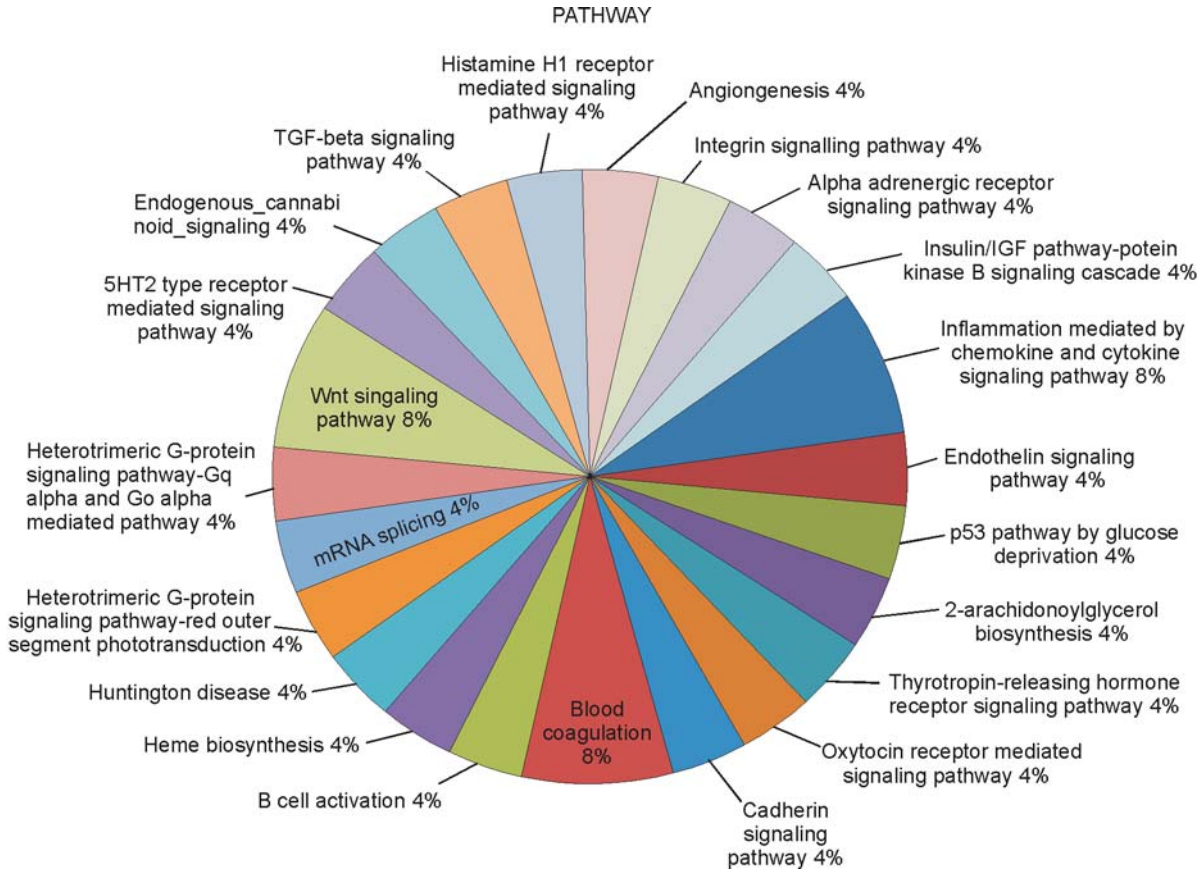
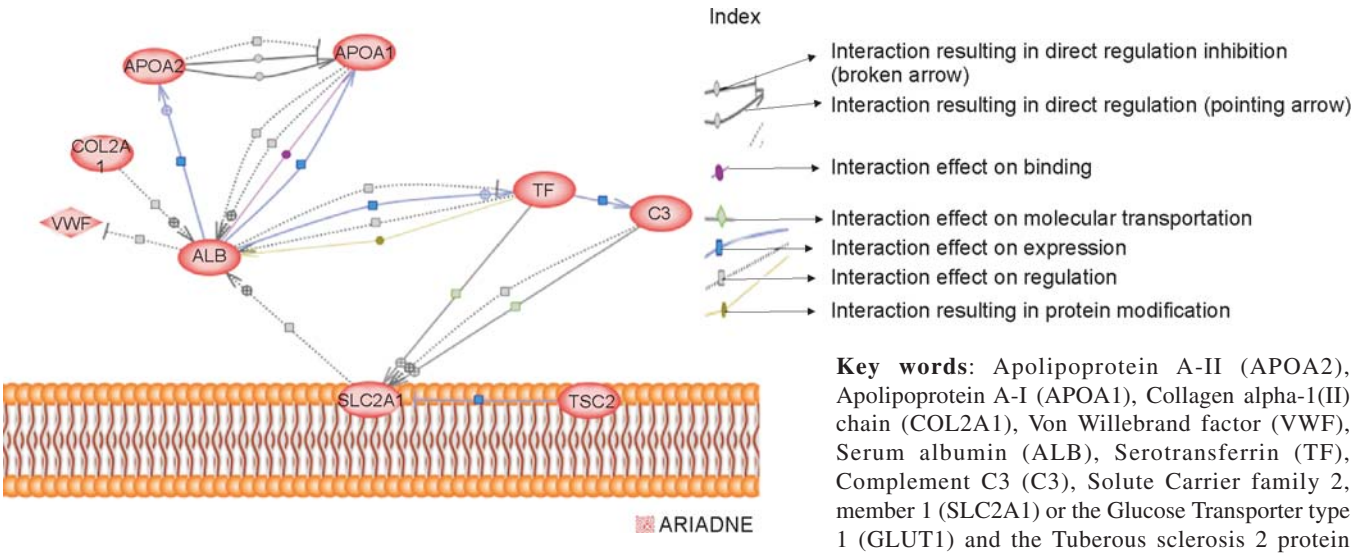


Fig. 2. Percent of identified proteins corresponding different pathways.



Key words: Apolipoprotein A-II (APOA2), Apolipoprotein A-I (APOA1), Collagen alpha-1(II) chain (COL2A1), Von Willebrand factor (VWF), Serum albumin (ALB), Serotransferrin (TF), Complement C3 (C3), Solute Carrier family 2, member 1 (SLC2A1) or the Glucose Transporter type 1 (GLUT1) and the Tuberous sclerosis 2 protein (TSC2)

Fig. 3. Network proposed for 9 of the identified proteins in sera of early pregnant buffalo.

activity, enzyme regulation, metabolic functions etc. Many proteins were seen to be concerned with cell communications, transport functions, immune system processes and developmental processes. These proteins predominantly featured in pathways important for developmemt processes in the embryo including Wnt signalling, inflammation

mediated by chemokine and cytokine signalling, and blood coagulation pathways. Protein- protein interaction map and network analysis of these bovine proteins indicate that probably the expression of these proteins is complex wherin they may be interacting with each other. These interactions affect molecular transportation, levels of expression, binding, modifications, direct regulation and direct regulation inhibition. The early pregnancy in buffaloes does have changes in expression of proteins, but exactly which proteins change is need to be studied further. These are only preliminary studies and further experimentation using modern biotechnological tools like proteomics and bioinformatics will help in the identification of biomarkers for the events in early pregnancy.

ACKNOWLEDGEMNTS

Authors are grateful to Dr R.C. Upadhyay, DCP Division, NDRI Karnal and National Initiative on Climate Resilient Agriculture (NICRA) for providing funds in carrying out this research. The authors are also thankful to the Directors of Central Institute for Research on Buffaloes, Hisar and National Dairy Research Institute, Karnal for providing animals and instruments, respectively, for this study.

REFERENCES

Ashburner M, Ball C A, Blake J A, Botstein D, Butler H, Cherry J M, Davis A P, Dolinski K, Dwight S S, Eppig J T, Harris M A, Hill D P, Tarver L I, Kasarskis A, Lewis S, Matese J C, Richardson J E, Ringwald M, Rubin G M and Sherlock G. 2000. Gene ontology: Tool for the unification of biology. *Nature Genetics* **25**: 25–29.

Balagopalakrishna C, Paka L, Pillarisetti S and Goldberg I J. 1999. Lipolysis-induced iron release from diferric transferrin: Possible role of lipoprotein lipase in LDL oxidation. *Journal of Lipid Research* **40**(7):1347–56.

Berny M A, White T C, Tucker E I, Bush-Pelc L A, Di Cera E, Gruber A and McCarty O J. 2008. Thrombin mutant W215A/E217A acts as a platelet GPIb antagonist. *Arteriosclererosis Thrombosis and Vascular Biology* **28**(2): 329–34.

Buller C L, Loberg R D, Fan M H, Zhu Q, Park J L, Vesely E, Inoki K, Guan K L and Brosius F C. 2008. A GSK-3/TSC2/mTOR pathway regulates glucose uptake and GLUT1 glucose transporter expression. *American Journal of Physiology Cell Physiology* **295**(3): C836–43.

Cui W, Paglialunga S, Kalant D, Lu H, Roy C, Laplante M, Deshaies Y, Cianflone K. 2007. Acylation-stimulating protein/C5L2-neutralizing antibodies alter triglyceride metabolism in vitro and in vivo. *American Journal of Physiology, Endocrinology and Metabolism* **293**(6): E1482-91.

Gao X, Yuan S, Jayaraman S and Gursky O. 2009. Differential stability of high-density lipoprotein subclasses: effects of particle size and protein composition. *Journal of Molecular Biology* **387**(3): 628–38.

Hoofnagle A N and Heinecke J W. 2009. Lipoproteomics: using mass spectrometry-based proteomics to explore the assembly, structure, and function of lipoproteins. *Journal of Lipid Research* **50**(10):1967–75

Hsieh-Bonassera N D, Wu I, Lin J K, Schumacher B L, Chen A C, Masuda K, Bugbee W D and Sah R L. 2009. Expansion and redifferentiation of chondrocytes from osteoarthritic cartilage: cells for human cartilage tissue engineering. *Tissue Engineering Part A* **15**(11): 3513–23.

Martin L B, Shewan, Millar C A, Gould G W and James D E. 1998. Vesicle-associated membrane protein 2 plays a specific role in the insulin-dependent trafficking of the facilitative glucose transporter GLUT4 in 3T3-L1 adipocytes. *Journal of Biological Chemistry* **273**(3):1444–52.

Peifer M and Polakis P. 2000. Wnt signalling in oncogenesis and embryogenesis- a look outside the nucleus. *Science* **287**(5458): 1606–09.

Sakai T, Jin FY, Kamanna VS and Kashyap M L. 2000. Albumin inhibits apolipoprotein AI and AII production in human hepatoblastoma cell line (Hep G2): additive effects of oleate-albumin complex. *Atherosclerosis* **149**(1): 43–49.

Shin D J and Osborne T F. 2003. Thyroid hormone regulation and cholesterol metabolism are connected through sterol regulatory element-binding protein–2 (SREBP-2). *Journal of Biological Chemistry* **278**(36): 34114–118.

Tang S, Lai K N, Chan T M, Lan H Y, Ho S K and Sacks S H. 2001.Transferrin but not albumin mediates stimulation of complement C3 biosynthesis in human proximal tubular epithelial cells.*American Journal of Kidney Diseases* **37**(1): 94–103.

Wilcox H G, Frank R A and Heimberg M. 1991. Effects of thyroid status and fasting on hepatic metabolism of apolipoprotein A-I. *Journal of Lipid Research* **32**(3): 395–405.

Worrall S, de Jersey J, Wilce P A, Seppa K, Hurme L and Sillanaukee P. 1998. Comparison of carbohydrate-deficient transferrin, immunoglobulin A antibodies reactive with acetaldehyde-modified protein and acetaldehyde-modified albumin with conventional markers of alcohol consumption. *Alcoholism, Clinical and Experimental Research* **22**(9): 1921–26.

Wroblewska M, Kortas-Stempak B, Szutowicz A and Badzio T. 2009. Phospholipids mediated conversion of HDLs generates specific apoA-II pre-beta mobility particles. *Journal of Lipid Research* **50**(4): 667–75.

Zensi A, Begley D, Pontikis C, Legros C, Mihoreanu L, Büchel C and Kreuter J. 2010. Human serum albumin nanoparticles modified with apolipoprotein A-I cross the blood-brain barrier and enter the rodent brain. *Journal of Drug Targeting* **18**(10): 842–48.