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

A K BALHARA1, A K MOHANTY2, MEENAKSHI GUPTA3, MANU JAMWAL4, SUNESH5, S K PHULIA6 and INDERJEET SINGH7

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.

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 classification 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](image-url)
Table 1. Proteins corresponding to differentially expressed 2-D gel spots of early pregnant buffaloes

<table>
<thead>
<tr>
<th>Spot No.</th>
<th>Identified protein</th>
<th>Gene name/symbol</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Serum albumin precursor</td>
<td>ALB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>4</td>
<td>Apolipo-protein A – II</td>
<td>APOA2</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>5</td>
<td>Synaptotagin-1 (EC 3.1.3.36) (synaptic insitol-1,4,5-trisphosphate 5-phosphatase</td>
<td>SYNJ1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td></td>
<td>1 p150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Keratin, Type 1 cytoskeletal 10</td>
<td>KRT10</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>8</td>
<td>1 phosphatidylinsitol-4, 5-biphosphate phosphodiesterase</td>
<td>PLCB1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Monodelphis</td>
</tr>
<tr>
<td>10</td>
<td>Small nuclear ribonucleoprotein-associated protein B’</td>
<td>SNRBP</td>
<td>domestica</td>
</tr>
<tr>
<td>13</td>
<td>Serum Albumin</td>
<td>ALB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>14</td>
<td>Serum albumin precursor</td>
<td>ALB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>17</td>
<td>Ferrochelatase</td>
<td>FECH</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>18</td>
<td>Protein regulator of cytokinesis 1</td>
<td>PRC1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>21</td>
<td>Hypothetical protein</td>
<td>Not available</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>22</td>
<td>Hypothetical protein (Receptor-type tyrosine-protein phosphatase-like N)</td>
<td>PTPRN</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>23</td>
<td>MHC Class I antigen</td>
<td>Cen-A</td>
<td>Cervus nippon (sika deer)</td>
</tr>
<tr>
<td>25</td>
<td>Collagen alpha-1(II) chain</td>
<td>COL2A1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>26</td>
<td>Serotransferrin</td>
<td>TF</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>27</td>
<td>Immunoglobulin heavy chain variable region</td>
<td>Not found</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>28</td>
<td>Tuberous sclerosis 2 protein (TSC2)</td>
<td>TSC2</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>29</td>
<td>Putative glucose transporter isoform 1 (EC 6.3.2.19)</td>
<td>GLUT1</td>
<td>Bubalus bubalis</td>
</tr>
<tr>
<td>31</td>
<td>Complement C1q subcomponent subunit B</td>
<td>CIQB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>33</td>
<td>Cannot be retrieved</td>
<td>Not known</td>
<td>Not known</td>
</tr>
<tr>
<td>34</td>
<td>5’-nucleotidase, cytolic IB</td>
<td>NTSCIB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>35</td>
<td>Serum albumin precursor</td>
<td>ALB</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>36</td>
<td>Complement component C3d</td>
<td>Not available</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>37</td>
<td>DNA polymerase alpha catalytic subunit</td>
<td>POLA2</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>38</td>
<td>Interleukin – 27 subunit- alpha (IL27RA) protein</td>
<td>IL27RA</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>40</td>
<td>Acyl CoA Synthetase</td>
<td>ACS</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>41</td>
<td>IgG 3 heavy chain constant region</td>
<td>IGHG3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>42</td>
<td>Hypothetical protein similar to Acyl-CoA dehydrogenase very long chain specific</td>
<td>Not available</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td>mitochondrial precursor (VLCAD) isoform</td>
<td></td>
<td>Canis familiaris</td>
</tr>
<tr>
<td>43</td>
<td>unclassified</td>
<td>Not known</td>
<td>Not known</td>
</tr>
<tr>
<td>44</td>
<td>Glutathione S-transferase</td>
<td>GSTA3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>45</td>
<td>Cyclic nucleotide-gated cation channel beta-1</td>
<td>CNGB1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>46</td>
<td>Transforming growth factor beta-1 precursor</td>
<td>TGFBI</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>47</td>
<td>Hypothetical protein</td>
<td>A1A4K4_F2</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>48</td>
<td>Peroxidasin homolog</td>
<td>PXDN</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>50</td>
<td>Protocadherin-12</td>
<td>PCDH12</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>51</td>
<td>Hydroxysteroid (17-beta) dehydrogenase very long chain specific mitochondrial</td>
<td>HKE6</td>
<td>Bos taurus</td>
</tr>
<tr>
<td></td>
<td>precursor (VLCAD) isoform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>SLAIN motif-containing protein 2</td>
<td>SLAIN2</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>54</td>
<td>U4/U6 small nuclear ribonucleoprotein Prp3 protein</td>
<td>PRPF3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>55</td>
<td>6-phosphogluconolactonase</td>
<td>PGLS</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>56</td>
<td>Apolipoprotein B</td>
<td>APOB</td>
<td>Diceros bicornis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Black rhinoceros)</td>
</tr>
<tr>
<td>57</td>
<td>Complement C3</td>
<td>C3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>58</td>
<td>Complement C3</td>
<td>C3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>59</td>
<td>Von Willebrand factor</td>
<td>vwf</td>
<td>Crocuta crocuta</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Spotted Hyena)</td>
</tr>
<tr>
<td>60</td>
<td>MHC Class I heavy chain</td>
<td>Not available</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>62</td>
<td>Apolipoprotein A-1 precursor</td>
<td>APOA1</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>63</td>
<td>Hypothetical protein</td>
<td>CKM</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>64</td>
<td>Origin recognition complex subunit 3</td>
<td>ORC3</td>
<td>Bos taurus</td>
</tr>
<tr>
<td>65</td>
<td>Hypothetical protein</td>
<td>Not available</td>
<td>Bos taurus</td>
</tr>
</tbody>
</table>
Pregnancy associated proteins are closely related. These
signalling cascade and heme biosynthesis are general
angiogenesis, insulin/IGF pathway—protein kinase B
(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,
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) - Apolipoprotein A
– II, Apolipoprotein 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 derived from the available
data on the Human Protein Reference Database (HPRD;
www.hprd.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 serotransferrin.
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 Heincke 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
considered 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
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 development 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 wherein 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.

ACKNOWLEDGEMENTS

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


