Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sex-Biased Expression of Genes Allocated in the Autosomal Chromosomes: Lc-ms/ms Protein Profiling in Healthy Subjects

Hayder Ahmed Giha (gehaha2002@yahoo.com)
Arabian Gulf University College of Medicine and Medical Science-previous https://orcid.org/0000-0001-6343-1949
Rabab Asghar Abdulwahab
University of Bahrain College of Health Sciences
Jaafar Abbas
Gulf Medical and Diabetes Center
Zakia Shinwari
King Faisal Specialist Hospital and Research Center
Ayodele Alaiya
King Faisal Centre Health Science Library: King Faisal Specialist Hospital and Research Center

Research Article

Keywords: Sex/gender, Sex-biased gene expression, Proteomics, LC-MS/MS, Bioinformatics, Bahrain

Posted Date: March 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-353208/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Sex and gender has large impact in human health and disease prediction. Men differ from women by a limited number of genes in Y chromosome while their phenotypes differ markedly. In this study, serum samples from healthy Bahraini men and women were analyzed by LC-MS/MS. Bioinformatics databases were used for proteins/peptides (PPs) identification and their gene localization. Results revealed that, the PPs which differed significantly ($p<0.05$ ANOVA) in abundance with fold change (FC) of ³1.5, were twenty, 11 were up-regulated in women (up to 8-folds), and 9 were up-regulated in men but with much lower FC, however, all PPs are encoded by genes located in autosomal chromosomes, indicative of sex-biased gene expression. The only PP related to sex, the sex hormone-binding globulin, was up-regulated in women. The remaining PPs were involved in immunity (6), lipid metabolism (5), gene expression (3) connective tissue (3), and others. The identified PPs were discussed within their physiological and pathological context, in relation to sex e.g. Apo-B100 (pathological cholesterol) was unregulated in men while the inflammatory/immunity related PPs, including Alpha-1-acid glycoproteins, were up-regulated in women. Finally, we propose proteomic as an ideal complementary tool for study of the molecular basis of the sex-biased gene expression.

Introduction

Biological and physiological differences exist between men and women [1]. The classical molecular differentiation between the two sexes is the genomic makeup and to some extent the transcriptomic typing [2, 3]. Proteomic differences between the two sexes independent of the genome can be explained in part by the, epigenetic modification, alternative splicing, post-translation modifications and etc., [4]. However, the influence of the sex hormones in gene expression can't be ignored. Although the human genome is expected to constitute between 20,000 - 25,000 genes [5], in our setting proteomic analysis revealed more than 2 millions proteins and peptides (PPs) [6]. A new and expanded central dogma composed of OMICs platforms, including; genome, trasnscriptome and proteome is needed to link the genotypes with the phenotypes. In addition to their normal physiological roles, the PPs are central players in disease etiology, pathology and complications as well as they are utilized as diagnostic and prognostic biomarkers [7].

In the era of personalized and precision medicine, sex (biological determinant of men and women) and gender (environmental and cultural determinant) stands early during diversification of people into categories and subcategories to reach the individuals. Thus, an initial step towards personalization of medicine would be setting the molecular differences between healthy men and women. The inherent susceptibility of men and women to disease is known to vary considerably e.g. men are more prone to infections, coronary heart disease (CHD) while women are more prone to autoimmune and inflammatory diseases and etc., [8, 9, 10, 11]. Also, it is well known that women have more vigorous innate immunity [9], and humoral immunity to antigenic challenges [12], which accelerate pathogen clearance but can lead to increased frequency of immunological disorders such as autoimmune or inflammatory diseases [10, 13]. Furthermore, it has been observed that sex/gender affect the responses to vaccination and its outcome [14].

Proteomics an up-and-coming discipline emerged from the human genome project, and became indispensable platform for better understanding of the genomic and transcriptomic data [15]. It is an approach for proteins/peptide detection and quantitation based on the knowledge about protein sequences, which facilitates protein structure and function prediction. The LC-MS/MS have revolutionized the proteomic analysis [16]. The used techniques and software are regularly upgraded for generation and analysis of huge data [17]. The proteomic analysis is superior to the other biochemical and antibody-based approaches in the ability to detect small differences in PPs levels between study materials, and have high accuracy and sensitivity in PPs identification [18].

Studies of the differences between men and women are mostly in terms of susceptibility to specific disorder, drug response/interaction or variations of biochemical markers [19], but very few studies that have focused on the differences between the two sexes in the healthy status [20]. Sex-specific physiological events or disorders e.g. pregnancy and lactation and breast and prostatic tumors, were organ specific with different markers and susceptibility factors [21].

This data is part of a larger study of the proteomic changes in T2DM, where several proteins were identified to be differentially expressed in T2DM, mostly were up-regulated [6], however, the influence of gender cannot be corrected for by the traditional proteomic statistical analysis. In this study we compared the protein profile of the healthy men and women whom were included in the main study.

Materials And Methods

Study subjects

A subset of samples was obtained from 6 health subjects (3 males / 3 females, age and sex-matched) selected from a larger number of Bahraini volunteers involved in study of T2DM biomarkers, between September 2014 and February 2015 (Table 1). Written informed consent was obtained from each study subject before blood collection. The study design followed the Helsinki declaration terms. The study was approved by the Research and Ethics Committees of the Arabian Gulf University (AGU) and Salmaniya Medical Complex Hospital (SMC), Manama, Bahrain.

Sample collection: Approximately 10 ml of venous blood sample was collected from each study subject, 10-12 hours after overnight fasting, into different collection tubes (Thermo Fisher Scientific, Massachusetts, U.S.A.) and centrifuged (3000 x g for 10 min) to collect the plasma and serum.

Biochemical tests

Glycemic profile: The fasting blood glucose (FBG) level was analyzed by Clinical Analyzer ROCHE COBAS INTEGRA 800 (Rotkreuz, Switzerland) and the results were expressed as, mmol/L. The glycated hemoglobin (HbA1c) was measured by Clinical Chemistry analyzer (Beckman Coulter AU, USA). The HbA1c to total Hb ratio (HbA1c %) was calculated.

Serum lipid profile: The lipids profile parameters; total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides (TGs), were measured in the biochemistry laboratory at SMC, using Clinical Analyzer ROCHE COBAS INTEGRA 800 (Rotkreuz, Switzerland).

Protein profiling using LC-MS/MS

The protein profile analysis was done by LC-MS/MS as previously described [22, 23], in brief;

Serum preparation - protein depletion: Pierce Albumin/IgG Removal kit (Thermo Scientific, U.S.A.) was used to remove the major subclasses of gamma globulin (IgG) and human serum albumin (HSA) from serum in order to overcome the serum protein complexity (depleted serum). The unprocessed serum (non-depleted serum), was also used. Thereafter, the exact quantity of proteins in serum was determined in order to calculate the amount of sample needed for analysis.

In-solution protein digestion: The sera samples, both non-depleted and depleted, were diluted 1:1 with 0.1% *Rapi*Gest™ SF (Waters, UK). Thereafter, the sera were pooled into 2 sets of samples, 3 healthy males and 3 healthy females. The total load for each pooled sample for analysis was 100 µg of protein/peptide (PPs) in a final volume of 25 µl. Then, the PPs in the pooled samples were denatured in a Thermo-mixer R (Eppendorf, Hamburg, Germany) at a speed of 8 x g at 80˚C for 15 min. At the end, the PPs were digested by Trypsin at a 1:50 ratio (Promega Corporation, Madison, WI, USA). The PPs digest were then analyzed using Synapt G2 MS (Waters, Manchester, UK).

Protein identification: As we reported before [6], the 1-dimensional Nano Acquity liquid chromatography coupled with tandem mass spectrometry on a Synapt G2 instrument (Waters, Manchester, UK) was used for label-free quantitative expression protein profiling. For ESI Mass Spectrometry analyses, the instrument settings were optimized on the MassLynx tune page. A protein digest of 3-µg was loaded on the column and all the samples were spiked with yeast alcohol dehydrogenase (ADH, P00330) as an internal standard to the digest to give 200 fmol per injection for absolute quantitation. An Acquity sample manager was used for injection of the analyzed samples.

The samples were analyzed in duplicate runs and data were acquired using the Mass Lynx programs (version. 4.1, SCN870; Waters) operated in resolution and positive polarity modes. The acquired MS data were background subtracted, smoothed and de-isotoped at the medium threshold. Progenesis QI V2.0 (QIfp) for proteomics (Nonlinear Dynamics/Waters, UK) was used for automated data processing and database searching. The generated peptide masses were searched against in UniProt species-specific protein sequence database using the Progenesis QI V2.0 (QIfp) for proteomics, for protein identification and quantification (Nonlinear Dynamics/Waters, UK).

Data analysis and informatics

Proteomics data: Progenesis QI V2.0/TransOmics Informatics (Waters Scientific, Manchester, UK) software was used to process and search the data using the principle of search algorithm as previously described [23, 24]. The data were filtered to show only statistically significant differences ($p < 0.05$, ANOVA) coupled with a change in proteins' abundance by 1.5 fold or more (Maximum Fold Change - MFC ≥ 1.5). Additionally, the absolute quantification was performed using ADH as an internal standard to give an absolute amount of each identified protein/peptide [6].

Bioinformatics: For gene localization, PPs identification and classification, sites of expression in human tissues, biological and molecular functions, and related genetic disorders, several proteomic, genomic and informatics data bases were searched. These included: https://www.uniprot.org; https://www.ncbi.nlm.nih.gov/gene; https://www.genecards.org/; https://www.omim.org/ and http://www.unicarbkb.org/ (for glycated proteins only). Gene expression databases were approached through https://www.uniprot.org/uniprot/. Other information was obtained from published data through the https://www.ncbi.nlm.nih.gov /pubmed and google as non-professional site [6].

Results

Description of the differentially expressed proteins in healthy men and women:

As seen in (Table 2), a total of 20 PPs were found to be differentially expressed between men and women. Using the proteomic data bases, most of the identified PPs are expressed in several tissues and organs, however, the following organs/tissues are the predominant producers for the identified PPs; the liver (9 PPs), the immune cells, mostly lymphocytes (5 PPs), gastrointestinal tract (GIT) including the appendix, oral cavity and gallbladder (4 PPs), female genital organs (3 PPs) and the central nervous system, specifically, the corpus colosum (1 PP). Worth noting, some PPs were expressed predominantly by more than one tissue. Based on function, the PPs were broadly classified into the following categories: a. Lipid metabolism (5 PPs) b. Immunity, (4 PPs), c. Connective tissue (3 PPs), d. Gene expression (3 PPs), e. Acute phase proteins (2 PPs), f. Transport proteins (1 PP), g. Signal transduction (1 PP), h. Unknown (1 PP). However, many PPs could be assigned for more than one class (source e.g. using Reactome pathway database https://reactome.org or Panther Classification System http://www.pantherdb.org/)

Abundance of up-regulated PPs in healthy women and men:

Figure 1, shows the differentially expressed PPs, the PPs which were upregulated in women (A) and the ones which were upregulated in men (B). The differences in the abundance of the PPs in the sera between the two sexes was rated by the maximum fold change (MFC). The PPs which were increased by more than 3 folds i.e. MFC >3 were mostly up-regulated in women. These PPs in a decreasing order were: Apolipoprotein C-II (8.6 folds), Homeobox protein Hox-D13 (6.0 folds), Apolipoprotein A-II (4.0 folds), Isoform 4 of Coiled-coil domain-containing protein 17 (3.6 folds), and Alpha-1-acid glycoprotein 2 (3.4 folds). However, in men only the Isoform 4 of Fibronectin was increased by more than 3 folds compared with women (MFC 3.02). For the fold changes and the p values see (Table 2 and Fig. 1).

Identification of the differentially expressed genes - Bioinformatics data extraction

Bioinformatics was the tool used to determine the differentially expresses genes of the identified PPs, we used the accession numbers obtained from LC-MS/MS analysis. The genetic background of the identified PPs was obtained by navigation through Protein knowledgebase – UniProtKB, site: https://www.uniprot.org/uniprot/. All identified PPs were exclusively expressed by genes (gn) located in the autosomal chromosomes (Ch), as follows: Ch 1 (4 gn), Ch 2 (3 gn), Ch 9 (3 gn), Ch 14 (3 gn), Ch 17 (2 gn), and Ch [3, 12, 19, 21 and 22] each (1 gn). The genes, their Ch and the gene locus in the Ch are shown in (Table 2).

Proteins/peptides upregulated in healthy women

As shown in (Fig. 1-A), the PPs which were upregulated in healthy women compared with men (ANOVA <0.05, MFC >1.5); were 11 PPs, 8 were identified in non-depleted sera (Fig. 2 A&B) and 3 in depleted sera (Fig. 3 A&B). The expression profiles of the identified PPs in non-depleted sera (Fig. 2-A) and depleted sera (Fig. 3-A) were strongly differentiating between women and men using correspondence analysis. The Hierarchical Cluster Analysis of the expression profiles of the identified PPs in non-depleted sera are shown in (Fig. 2-B), and that of the PPs in depleted sera are shown in (Fig. 3-B), both differentiate between men and women. Of these PPs, the ones which were expressed predominantly in the liver are 4 PPs, by lymphocytes and bone marrow cells were 3 PPs, and 1 PP in the CNS (corpus callosum), while only 2 PPs were expressed predominantly by females' genital organs. The classes of the PPs up-regulated in both women are men were the lipids metabolism and immunity classes, in women the 2 classes included the following PPs: Apolipoprotein A-II, Apolipoprotein C-II and CD5 antigen-like and Ig mu chain C. The acute phase proteins (Alpha-1-acid glycoprotein 1 and Alpha-1-acid glycoprotein 2) and signal transduction (Multiple PDZ domain protein) classes of PPs, were upregulated exclusively in women. Furthermore, 2 PPs involved in gene expression; the Homeobox protein Hox-D13 and Isoform 4 of Coiled-coil domain-containing protein 17, and one PP (Integrin beta-2) classified as connective tissue and another as transporter (Sex hormone-binding globulin) PP were upregulated also in women.

Proteins/peptides (PPs) upregulated in healthy men

The PPs upregulated in men compared with women were 9 PPs (Fig. 1-B), 7 were identified in non-depleted sera while 2 in depleted sera. Using the correspondence analysis, the expression profiles of the former (Fig. 2-A) and the latter (Fig. 3-A) PPs, were strongly differentiated women from men. As shown in Figs. (2-B and 3-B), the Hierarchical Cluster Analysis of the expression profiles of the 7 PPs in the non-depleted sera and that of the 2 PPs in depleted, respectively, clearly distinguish between men and women. Three of the up-regulated PPs in men; the complement factor H, Ig gamma-1 chain C region and Ig gamma-2 chain C region, belongs to the immunity class of PPs. Two PPs classified under each of the lipid metabolism (Apo-B100 and Isoform SH-iPLA2 of 85/88 kDa calcium-independent phospholipase A2) and connective tissue (Isoform 4 of Fibronectin and Vitronectin) classes. One PP, the Eukaryotic translation initiation factor 4E type 3 (Fragment), is involved in gene expression, specifically translation. However, one PP, the Isoform 2 of Putative golgin subfamily A member 2B, is likely to be allocated to Golgi apparatus but its class was unknown.

Ingenuity Pathway Analysis (IPA)

The Ingenuity Pathway Analysis (IPA) is used for meaningful interpretation of the gene expression data using prior biological knowledge. For networking of the identified PPs between each other and other PPs sharing the same function, and for casual and functional relation to diseases and immunity, the 20 PPs were subjected to IPA, using an all-in-one, web-based software application, from QIAGEN Bioinformatics. Only 14 PPs (7 were up regulated in men and 7 in women) (Table 3), were mapped in the IPA database together with another 18 PPs not identified in this study (Fig. 4). These PPs were found to be implicated in several networks. The gene sources of these proteins were highlighted in grey in (Fig. 4). The cellular localization of some of these PPs included the plasma membrane, cytoplasm, nucleus, and extracellular space (Table 3).

Discussion

The fact that men and women biologically, morphologically and physiologically are different [25], brings about the magnitude and nature of this difference at molecular level. It is well known that, genomically and genetically both sexes posses the same chromosomes and genes except for the sex chromosomes (XY), suggesting that the main differences between both sexes would be proteins encoded by or their synthesis is influenced by proteins/peptides (PPs) encoded by genes on Y chromosome e.g. SRY gene [26]. In this study, proteomic analysis revealed marked differences in levels of 20 PPs, none of which is encoded by genes located in the sex chromosomes, unlike what was observed elsewhere [27]. Furthermore, the PPs up regulation was more evident in women, confirming a recent report showing significant sexual dimorphisms in protein abundance between twin pairs of opposite sex [27], although men have the unique Y-chromosome.

The observation that, all identified PPs are expressed by genes located in the autosomal chromosomes (Table 2), is supported by a previous study reported that the sex differences in part are due to differences in expression of genes not in sex chromosomes [28], what was termed as sex-biased gene expression. Furthermore, three PPs known to be regulators for gene expression were differentially expressed in this study, supporting the sex-biased expression independent of the sex-chromosomes. Two transcription factors, Homeobox protein Hox-D13 and Isoform 4 of Coiled-coil domain-containing protein 17 (CCDC17) were up-regulated in women while a translational factor, Eukaryotic translation initiation factor 4E (eIF4E) type 3 (Fragment), was up-regulated in men. The Hox-D13 is not unexpected to be upregulated in women since the Hox genes family products are known to be transcriptional factors involved in female reproductive system development and function [29]. However, there are no clues that the other two PPs, the CCDC17 and eIF4E, have any specific sex-related role.

Interestingly, only one out of the 20 identified PPs was directly related to sex, the sex hormone-binding globulin (SHBG), which was up-regulated in women. This protein is only rarely investigated in healthy subjects. One study showed that SHBG mRNA is strongly correlating with serum SHBG protein level with

higher levels of the mRNA and protein in women than in men [30]. This report is supporting our proteomic analysis finding.

Of the 20 PPs which were markedly differentially expressed between sexes, 11 were up regulated in women and 9 in men, however, the difference in abundance were more marked in PPs up regulated in women (Fig. 1A&B). Using the same material, we previously detected more than 2 million PPs in human serum [6], the extremely low number of the identified PPs in this study compared to the total number of PPs, can be explained by the set criteria for marked differentially expressed PPs. The system was set to select only markedly differentially expressed PPs (MFC ≥1.5), what might had excluded thousands of PPs differentially expressed between the 2 sexes but with fold change less than 1.5. Worth noting in biological systems function and homoeostasis depend on fine balance of molecules i.e. not necessarily the quantity [31, 32]. However, using proteomic platform the identified PPs were selected based on quantity not function.

In this study, we didn't aim to elaborate on each of the 20 PPs individually, as the informatics databases are provided for that purpose (see materials and methods). Our goal here to, interprets the functional classes of the identified PPs, in terms of physiological and clinical differences between the two sexes. Worth noting, the identified PPs were restricted to a very limited classes of proteins based on their functions (Table 2).

A noticeable and relevant observation regarding the identified PPs, was the up regulation of Apo-lipoproteins, Apo AII and Apo CII in women and that of Apo B100 in men. The former two were the predominant Apo proteins in chylomicron and VLDL, which are not involved in atherosclerosis and coronary heart diseases directly (CHD), while the Apo-B100, the major protein in LDL, is the carrier of pathological cholesterol [33]. It is well known that men are more susceptible to CHD compared to women in the same age before menopause [8]. The sex-biased differential expression of the lipid metabolism class of PPs in this study is consistent with the inherent protection of women against CHD during the reproductive age.

The sex-based immunological dimorphism was known long time ago [34], however, detailed analysis showing the abundance of the immunoglobulins (Ig) chains at transcriptome or proteome level were rarely investigated. In this study, the up regulation of Ig mu chain C region (IgM heavy chain) in women and of Ig gamma-1 &2 chain regions (IgG heavy chains) and complement factor H (the complement alternative pathway) in men, probably reflects the difference between the two sexes in response to antigens. In line with our findings are Saudi and Iranian studies, both showed higher levels of IgM in healthy women compared with men, and higher IgG in men compared with women in the Saudi but not in the Iranian study [35, 36]. Studies of sex difference in the immune responses following infection/vaccination showed that, women across all age groups were able to mount higher immune responses to infections and vaccines than men [14, 37]. In contrast, women are at much higher risk for development of autoimmune diseases e.g., SLE compared to men [38], however, it worth to know if there is a link between the IgM and the increased risk in women for the above-mentioned conditions.

To the best of our knowledge, the differences in levels of the acute phase proteins, alpha-1-cid glycoprotein -1 and 2 (AGP1 and 2), also known as orosomucoid, between men and women were not reported before. In this study both PPs were up-regulated in women. The associations of AGP with rheumatoid arthritis [39] and other autoimmune diseases e.g. Grave's disease [40], together with the high prevalence of these disorders in women [11] are strongly supporting the upregulation of AGP in women observed in this study. The Multiple PDZ domain protein, was the only signal transduction PPs identified in this study, it was up-regulated in women. This protein is known to have the highest expression level in corpus callosum (brain). The up-regulation of the PDZ domain in women worth further investigations because, of the associations of this PP with several and diverse pathologies [41], and it is also involved in control of cell polarity and signal transduction [42, 43].

In this study, 9 out of the 20 PPs were upregulated in men compared with women, but the up regulation in men was not as high as in women in terms of abundance (fold change). Of the noticeable upregulated PPs in men, the 2 connective tissue (CT) proteins, fibronectin and vitronectin, the sex-differences in the levels of both PPs was not reported before. However, both proteins were noticed to be upregulated in T2DM [6, 44]. Furthermore, men and women differ in musculoskeletal system in which CT proteins are major component [45]. The differences in CT proteins between the two sexes have clinical implications e.g. the incidence of the anterior cruciate ligament injury is almost 10 times higher in women compared with men performing the same activity [46], an observation that supports our findings.

The last PP up-regulated in men is the Isoform 2 of Putative golgin subfamily A member 2B, one of the Golgi apparatus proteins, which is probably encoded by a pseudo-gene, (https://www. genecards.org/cgi-bin/carddisp.pl?gene=GOLGA2P5). However, no single article was published about this PP in the Pub Med database. This and other PPs listed in (Table 2), are not discussed here, but worth further investigation.

Finally, as a summary for the above, we further explored the functional characteristics and relatedness of the 20 PPs to disease and other immune mediated disorders using Ingenuity Pathway Analysis (IPA), (Table 3 & Fig. 4). The principle of the IPA was explained previously [47]. Only 14 of the 20 PPs were mapped in the IPA database and were found to be implicated in multiple signaling networks including cell-to-cell signaling and interaction, lipid metabolism, and small molecule biochemistry. The functional annotations of these proteins with others as transporter, transmembrane receptor, and catalysts is referred to in (Table 3), while other functions included implications in different drug agents (data not shown).

In conclusion, 20 PPs were found to be differentially expressed in healthy men and women, 11 up-regulated in women with high fold change, while 9 were up-regulated in men with lower fold change compared to women. Although differentiating between the two sexes, none of the identified PPs is encoded by gene allocated in sex chromosomes. The frequently obvious links between sex, PPs up-regulation, and disease susceptibility e.g. the Apo-B100, men and CHD, are validation for this proteomic data. The association of the remaining PPs with sex and their physiological and pathological roles need further investigations. Finally, the molecular basis of the phenotypic differences between men and women is still unclear.

Declarations

Funding: No funding received

Availability of data and materials: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request

Authors' contributions: HAG, RAA and AA developed the research idea and study design; RAA and JA collected the samples. AA, RAA and ZS conducted the lab work. HAG, AA, and RAA done the bioinformatics and statistical analysis. HAG, RAA and AA wrote the manuscript and all authors contributed to the final version of the manuscript

Ethics approval and consent to participate: In this study, the procedures involving human participants were done in accordance with the ethical standards of the Research Committees of the Arabian Gulf University and SMC (Manama, Bahrain), and with the 1964 Helsinki Declaration and its later amendments. An informed consent was obtained from each subject.

Patient consent for publication: The patients were consented for publication

Competing interests: All authors declare that they have no competing interests

Acknowledgments: The medical doctors and lab staff in SMC (Manama – Bahrain) and the colleagues at the College of Health Sciences of Bahrain University (Manama, Bahrain) and their families and friends are thanked for their contribution in sample collection of our main project. The authors would like also to extend their sincere appreciation to the staff of Proteomics Unit in King Faisal Specialist Hospital and Research Center (Riyadh - Saudi Arabia) for the proteomics lab work.

References

1. Federman DD. The biology of human sex differences. Engl. J. Med. 2006; 354:1507–1514.
2. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ. The human transcriptome across tissues and Science. 2015; 348:660–5.
3. Mank JE. The transcriptional architecture of phenotypic dimorphism. Nat Ecol Evol. 2017; 1(1):6. doi: 10.1038/s41559-016-0006. PMID:28812569
4. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics. 2005; (10):1487-502. PMID: 15979981.
5. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004; 431(7011):931-45. PMID:15496913
6. Abdulwahab RA, Alaiya A, Shinwari Z, Allaith AAA, Giha HA. LC-MS/MS proteomic analysis revealed novel associations of 37 proteins with T2DM and notable upregulation of immunoglobulins. Int J Mol Med. 2019; 43(5):2118-2132. doi: 10.3892/ijmm.2019.4127. PMID:30864687
7. Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M. Plasma Proteome Profiling to Assess Human Health and Disease. Cell Syst. 2016; 2(3):185-95. doi: 10.1016/j.cels.2016. 015. PMID: 27135364
8. Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. Lancet1999; 353:89-92.
9. Marriott I, Huet-Hudson YM. Sexual dimorphism in innate immune responses to infectious organisms. Immunol Res 2006; 34(3):177-92. DOI: 10.1385/IR:34:3:177
10. Voskuhl R. Sex differences in autoimmune diseases. Biol Sex Differ 2011; 2(1):1-21. DOI: 10.1186/2042-6410-2-1.
11. Rose NR. Prediction and Prevention of Autoimmune Disease in the 21st Century: A Review and Preview, Am J Epidemiol. 2016; 183(5):403-6. doi: 10.1093/aje/kwv292. PMID:26888748
12. Cook IF. Sexual dimorphism of humoral immunity with human vaccines. Vaccine 2008; 4; 26(29-30):3551-5. DOI: 10.1016/j.vaccine.2008.04.054
13. Okoro NI, Kane SV. Gender-related issues in the female inflammatory bowel disease patient. Expert Rev Gastroenterol Hepatol 2009; 3 (2):145-54. DOI: 10.1586/egh.09.1
14. Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines. Lancet Infect Dis. 2010; 10:338–49.
15. Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev. 1996; (1):19-50, org/10. 1080/02648725. 1996.10647923.
16. Fröhlich T, Arnold GJ. Proteome research based on modern liquid chromatography–tandem mass spectrometry: separation, identification and quantification. J Neural Transm (Vienna). 2006; 113(8):973-94. PMID:16835695
17. Cottrell JS: Protein identification using MS/MS data. J Proteomics. 74:1842–1851. 2011.
18. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature. 2008 455: 1251–1254.
19. Lau ES, Paniagua SM, Guseh JS, Bhambhani V, Zanni MV, Courchesne P, Lyass A, Larson MG, Levy D, Ho JE. Sex Differences in Circulating Biomarkers of Cardiovascular Disease. J Am Coll Cardiol. 2019; 74(12):1543-1553. doi: 10.1016/j.jacc.2019.06.077. PMID:31537263
20. Gianazza E, Miller I, Guerrini U, Palazzolo L, Parravicini C, Eberini I. Gender proteomics Which proteins in non-sexual organs. J Proteomics. 2018a; 178:7-17. doi: 10.1016/j.jprot. 2017.10.002. PMID: 28988882
21. Gianazza E, Miller I, Guerrini U, Palazzolo L, Parravicini C, Eberini I. Gender proteomics II. Which proteins in sexual organs. J Proteomics. 2018b; 178:18-30. doi: 10.1016/j.jprot.2017 .10.001. PMID: 28988880

22. Alaiya A, Fox J, Bobis S, Matic G, Shinwari Z, Barhoush E, Márquez M, Nilsson S and Holmberg AR: Proteomic analysis of soft tissue tumor implants treated with a novel polybisphosphonate. Cancer Genomics Proteomics. 2014; 11:39–49.
23. Alaiya A, Assad L, Alkhafaji D, Shinwari Z, Almana H, Shoukri M, Alkorbi L, Ibrahim HG, Abdelsalam MS, Skolnik E, et al: Proteomic analysis of Class IV lupus nephritis. Nephrol Dial Transplant. 2015; 30:62–70.
24. Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV and Geromanos SJ: Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics. 9:1696–1719. 2009.
25. Mealey, L. (2000). Sex differences. NY: Academic Press, **eBook ISBN:**9780080541136
26. Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN, Fellous M. "Genetic evidence equating SRY and the testis-determining factor". Nature. 1990; 348 (6300): 448–50. doi:1038/348448A0. PMID2247149.
27. Witt E, Lorenz M, Völker U, Stangl K, Hammer E, Stangl V. Sex-specific differences in the intracellular proteome of human endothelial cells from dizygotic twins. J Proteomics. 2019; 201:48-56. doi: 10.1016/j.jprot.2019.03.016.PMID:30951907
28. Zhang W, Huang RS, Duan S, Dolan ME. Gene set enrichment analyses revealed differences in gene expression patterns between males and females. In Silico Biol. 2009; 9(3):55-63. PMID: 19795565
29. Du H, Taylor HS. The Role of Hox Genes in Female Reproductive Tract Development, Adult Function, and Fertility. Cold Spring Harb Perspect Med. 2015; 6(1):a023002. doi: 10.1101/cshperspect.a023002. PMID: 26552702
30. Winters SJ, Gogineni J, Karegar M, Scoggins C, Wunderlich CA, Baumgartner R, Ghooray DT. Sex hormone-binding globulin gene expression and insulin resistance. J Clin Endocrinol Metab. 2014; 99(12):E2780-8. doi: 10.1210/jc.2014-2640. PMID:25226295
31. Sies H, Berndt C, Jones DP. Oxidative Stress. Annu Rev Biochem. 2017; 86:715-748. doi: 10.1146/annurev-biochem-061516-045037. PMID:28441057
32. Noviski M, Zikherman J. Control of autoreactive B cells by IgM and IgD B cell receptors: maintaining a fine balance. Curr Opin Immunol. 2018; 55:67-74. doi: 10.1016/j.coi.2018.09. 015. PMID: 30292928
33. Weitgasser R, Ratzinger M, Hemetsberger M, Siostrzonek P. LDL-cholesterol and cardiovascular events: the lower the better? Wien Med Wochenschr. 2018; 168(5-6):108-120. doi: 10.1007/s10354-016-0518-2. PMID: 27770320
34. Butterworth M, McClellan B, Allansmith M. Influence of sex in immunoglobulin levels Nature. 1967;214(5094):1224-5
35. Harfi HA, Godwin JT. Normal serum levels of IgG, IgA, IgM, IgD, and IgE in Saudi Arabia. King Faisal Specialist Hosp Med J. 1985;5:99–105
36. Kardar GA, Shams SH, Pourpak Z, Moin M. Normal value of immunoglobulins IgA, IgG, and IgM in Iranian healthy adults, measured by nephelometry. J Immunoassay Immunochem. 2003; 24(4):359-67. PMID:14677654
37. Klein SL, Marriott I, Fish EN. Sex-based differences in immune function and responses to vaccination. Trans R Soc Trop Med Hyg. 2015; 109(1):9-15. doi: 10.1093/trstmh/tru167. PMID:25573105
38. Fan H, Dong G, Zhao G, Liu F, Yao G, Zhu Y, Hou Y. Gender differences of B cell signature in healthy subjects underlie disparities in incidence and course of SLE related to estrogen J Immunol Res. 2014; 2014:814598. doi: 10.1155/2014/814598. PMID: 24741625.
39. Haston JL, FitzGerald O, Kane D, Smith KD. The influence of alpha1-acid glycoprotein on collagenase-3 activity in early rheumatoid arthritis. Biomed Chromatogr. 2003; 17(6):361-4. PMID:13680845
40. Zimmermann-Belsing T, Feldt-Rasmussen U, From G, Perrild H, Bøg-Hansen TC. Long-term pathologic changes of alpha1-acid glycoprotein (orosomucoid) glycoforms in autoimmune thyroid disease. Autoimmunity. 2002; 35(7):441-7. PMID: 12685872.
41. Facciuto F, Cavatorta AL, Valdano MB, Marziali F, Gardiol D. Differential expression of PDZ domain-containing proteins in human diseases - challenging topics and novel issues. FEBS J. 2012; 279(19):3538-3548. doi: 10.1111/j.1742-4658.2012.08699.x. PMID:22776401
42. Humbert P, Russell S & Richardson H Dlg. Scribble and Lgl in cell polarity, cell proliferation and cancer. BioEssays. 2003; 25, 542–553.
43. Humbert PO, Grzeschik NA, Brumby AM, Galea R, Elsum I & Richardson HE. Control of tumourigenesis by the Scribble/Dlg/Lgl polarity module. Oncogene. 2008; 27, 6888–6907.
44. Abdulwahab RA, Allaith AAA, Shinwari Z, Alaiya A, Giha HA. Association of TATA box-binding protein-associated factor RNA polymerase I subunit C (TAF1C) with T2DM. Gene. 2019; 706:43-51. doi: 10.1016/j.gene.2019.04.076. PMID:31039436
45. Little D, Thompson JW, Dubois LG, Ruch DS, Moseley MA, Guilak F. Proteomic differences between male and female anterior cruciate ligament and patellar tendon. PLoS One. 2014;9(5):e96526. doi: 10.1371/journal.pone.0096526. eCollection 2014. PMID:24818782
46. Gwinn DE, Wilckens JH, McDevitt ER, Ross G, Kao TC. The relative incidence of anterior cruciate ligament injury in men and women at the United States Naval Academy. Am J Sports Med, 2000; 28: 98–102.
47. Krämer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014; 30(4):523-30. doi: 10.1093/bioinformatics /btt703. PMID:24336805

Tables

Table 1. Clinical, biochemical, and demographic characteristics of the study subjects.

Serial number	Sample ID	Clinical diagnosis	Sex	Age, years	Glucose mmol/l	HbA1c (%)	LDLC mmol/l	HDLCmmol/l	TGs mmol/l	T CHOL mmol/l
1.	S4	Healthy	F	42	5.7	ND	3.3	1.5	1.2	5.3
2.	S52	Healthy	M	42	5.1	ND	2.5	0.98	1.4	4.1
3.	S8	Healthy	F	41	5.4	ND	2.5	1.7	1	4.7
4.	S15	Healthy	M	41	4.9	ND	3.1	1.2	0.4	4.5
5.	SR	Healthy	F	47	4.9	ND	3.73	1.6	0.8	5.7
6.	SK	Healthy	M	47	5.5	ND	3.2	0.9	1.7	4.3

LDLC, low density lipoprotein cholesterol; HDLC, high density lipoprotein cholesterol; TGs, triglycerides; T CHOL, total cholesterol; ND, not done; HbA1c, blood glucose and glycated hemoglobin; T2DM, type 2 diabetes mellitus; M, male; F, female.

Table 2. List of up regulated proteins/peptides (PP)s in healthy men (grey shadow) and women, showing the PP category, accession number, gene and gene chromosomal location

	Peptides	Function/type	Accession No	Gene	Chrom
1.	Alpha-1-acid glycoprotein 1	Acute phase protein	P02763	ORM1	9q31
2.	Alpha-1-acid glycoprotein 2	Acute phase protein	P19652	ORM2	*9q32*
3.	Apolipoprotein A-II	lipid metabolism	V9GYM3;P02652;V9GYC1;V9GYE3;V9GYG9;V9GYS1	APOA2 -	1q23.3
4.	Apolipoprotein C-II	lipid metabolism	V9GYJ8;K7ER74;P02655	APOC2 -	19q13.32
5.	Apolipoprotein B-100	lipid metabolism	A8MUN2	APOB	2p24.1
6.	PLA2G6	Lipid metabolism	O60733-2;B0QYE9;O6 0733;O60733-3;O60733 -4	PLA2G6	22q13.1
7.	CD5 antigen-like	Lipid metabolism & Immunity	O43866	CD5L	1q23.1
8.	Ig mu chain C region	Immunity	P01871;P01871-2	IGHM	*14q32.33*
9.	Ig gamma-1 chain C region	Immunity	#	IGHG1	*14q32.33*
10.	Ig gamma-2 chain C region	Immunity	P01859	IGHG2	*14q32.33*
11.	Complement factor H	Immunity	P08603;B1AKG0;P08603-2;Q03591;Q5TFM2	CFH	*1q31.3*
12.	Integrin beta-2	Connective tissue & Immunity	P05107;A0A087WX36;D3DSM0;E5RHE6;E5RIG7;E5RK25;E5RK54;E7EVZ9;J3KNI6	ITGB2	*21q22.3*
13.	Isoform 4 of Fibronectin	Connective tissue	P02751-4	FN1	*2q35*
14.	Vitronectin	Connective tissue	P04004	VTN	*17q11.2*
15.	CCDC17	Gene expression (trsp)	Q96LX7-3	CCDC17	1p34.1
16.	Homeobox protein Hox-D13	Gene expression (trsp)	P35453	HOXD13	*2q31.1*
17.	*eIF4E3	Gene expression (trsl)	C9J7Z6;Q8N5X7;Q8N5X7-2	EIF4E3	*3p13*
18.	Multiple PDZ domain protein	Signal transduction	O75970;B7ZB24;F5H1U9;H0YGQ3;O75970-2;O75970-3;O75970-5	MPDZ	9p23
19.	* SHBG	Transport protein	I3L145;I3L2X4;P04278;P04278-5	SHBG	17p13.1
20.	* GOLGA2B	Unknown (UK)	Q9HBQ8-2;Q9HBQ8	GOLGA2P5 Pseudogene	12q23.1

*PLA2G6: Isoform SH-iPLA2 of 85/88 kDa calcium-independent phospholipase A2. *CCDC17: Isoform 4 of Coiled-coil domain-containing protein 17. *eIF4E3: Eukaryotic translation initiation factor 4E type 3 (Fragment). *SHBG: Sex hormone-binding globulin. *GOLGA2B: Isoform 2 of Putative golgin subfamily A member 2B. trsp: transcription. trsl: translation. #: A0A087WV47;A0A087WYC5; A0A087WYE1; A0A087X010;A0A0A0MS07;A0A0A0MS08; P01781;P01857.

Table 3. List of the 14 identified protein/peptides that were implicated in Ingenuity Pathway Analysis Database. Grey filling indicates PPs up-regulated in men

Symbol	Entrez Gene Name	Location	Family
APOA2	*apolipoprotein A2*	Extracellular Space	transporter
APOB	*apolipoprotein B*	Extracellular Space	transporter
APOC2	*apolipoprotein C2*	Extracellular Space	transporter
CFH	*complement factor H*	Extracellular Space	other
EIF4E3	*eukaryotic translation initiation factor 4E family member 3*	Cytoplasm	translation regulator
FN1	*fibronectin 1*	Extracellular Space	enzyme
IGHG2	*immunoglobulin heavy constant gamma 2 (G2m marker)*	Plasma Membrane	other
IGHM	*immunoglobulin heavy constant mu*	Plasma Membrane	transmembrane receptor
ITGB2	*integrin subunit beta 2*	Plasma Membrane	transmembrane receptor
MPDZ	*multiple PDZ domain crumbs cell polarity complex component*	Plasma Membrane	other
ORM1	*orosomucoid 1*	Extracellular Space	other
PLA2G6	*phospholipase A2 group VI*	Cytoplasm	enzyme
SHBG	*sex hormone binding globulin*	Extracellular Space	other
VTN	*vitronectin*	Extracellular Space	other

Figures

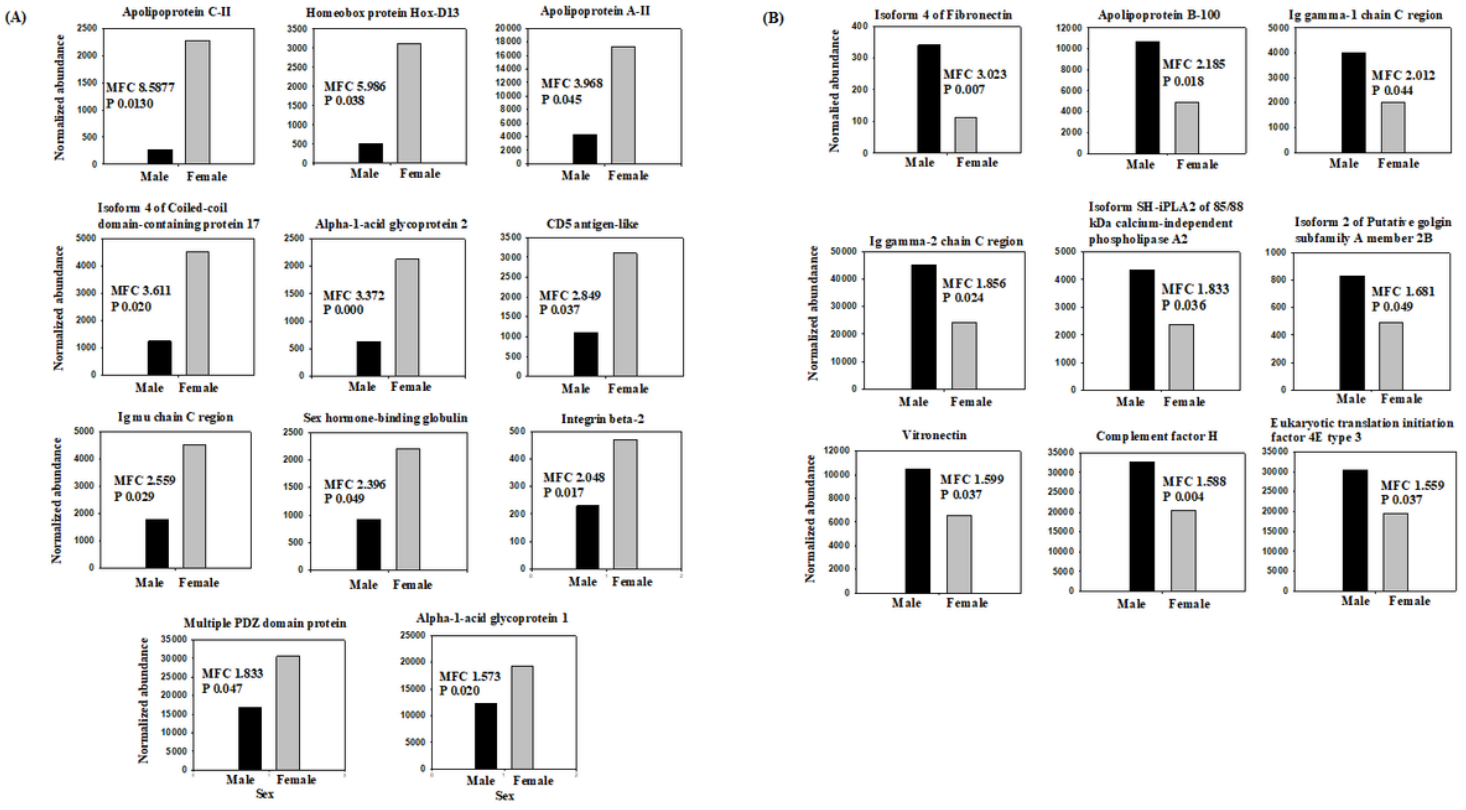

Figure 1

Histograms of the normalized abundance of differentially-expressed proteins/peptides (PPs) identified in healthy males (black) and females (grey). Figure 1-A, shows PPs upregulated in women compared with men, Figure 1-B. shows PPs upregulated in men compared with women. The P-values (ANOVA) and magnitude of change (maximum fold change - MFC) are indicated for each PP in each figure. The proteins were identified using label-free quantified liquid chromatography tandem mass spectrometry on Synapt G2 analysis – LC-MS/MS. Progenesis QI for Proteomics software was used for data analysis. Note: there are no bar errors as samples are usually pooled in LC-MS/MS

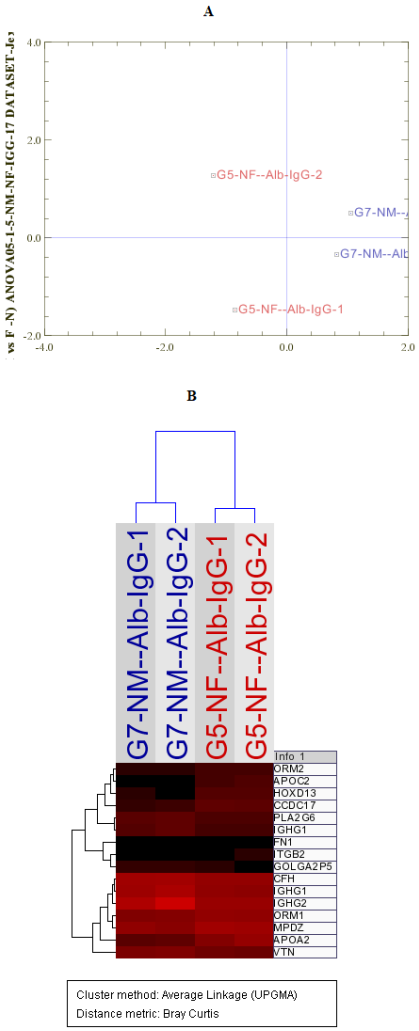

Figure 2

(A) Correspondence analysis (CA) plot of healthy subjects non-depleted serum using the expression dataset of 15 identified proteins/peptides (PPs) that were significantly differentially expressed (P<0.05 – ANOVA, MFC (minimum fold change) ≥1.5-) between men and women. The expression profiles of the identified PPs distinctively differentiate women from men using correspondence analysis. The letters in the background are the accession numbers of all the identified PPs in the analysis. [The image was generated using J-Express Pro V1.1 software program (java.sun.com)]. (B) Unsupervised Hierarchical Cluster Analysis of the expression profiles of non-depleted serum samples using 17 PPs that differ significantly (P<0.05 – ANOVA, MFC ≥1.5) between men (blue) and women (red). The dendrogram was generated using the Bray Curtis Correlation distance metric and an average linkage clustering method from the J-Express Pro V1.1 software program (java.sun.com). Note: one PP (IGHG1) had 2 accession numbers (subtypes of the same PP) is shown in the fig as two different PPs.

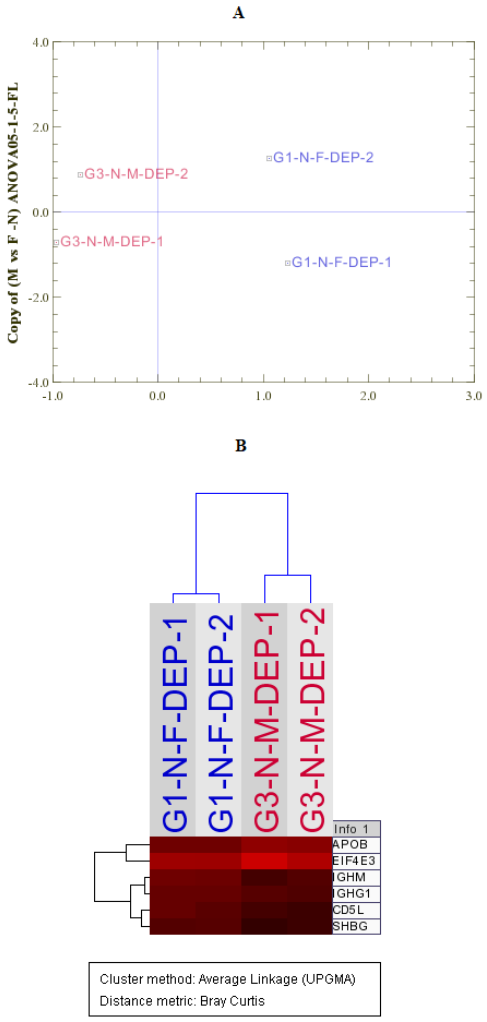

Figure 3

(A) Correspondence analysis (CA) plot of depleted serum samples from healthy subjects using the expression dataset of 5 identified proteins/peptides (PPs) that were significantly differentially expressed (P<0.05 – ANOVA, MFC ≥1.5- minimum fold change) between men and women. The expression profiles of the identified PPs clearly differentiate women from men using correspondence analysis. [The image was generated using J-Express Pro V1.1 software program (java.sun.com)]. (B) Unsupervised hierarchical Cluster Analysis of the expression profiles of non-depleted serum samples using 5 proteins that differ significantly (P<0.05, analysis of variance; maximum fold change ≥1.5) between men (red) and women (blue). The dendrogram was generated using the Bray Curtis Correlation distance metric and an average linkage clustering method from the J-Express Pro V1.1 software program (java.sun.com). Note: one PP (IGHG1) was up-regulated in both non-deleted (Fig. 2) and in depleted sera.

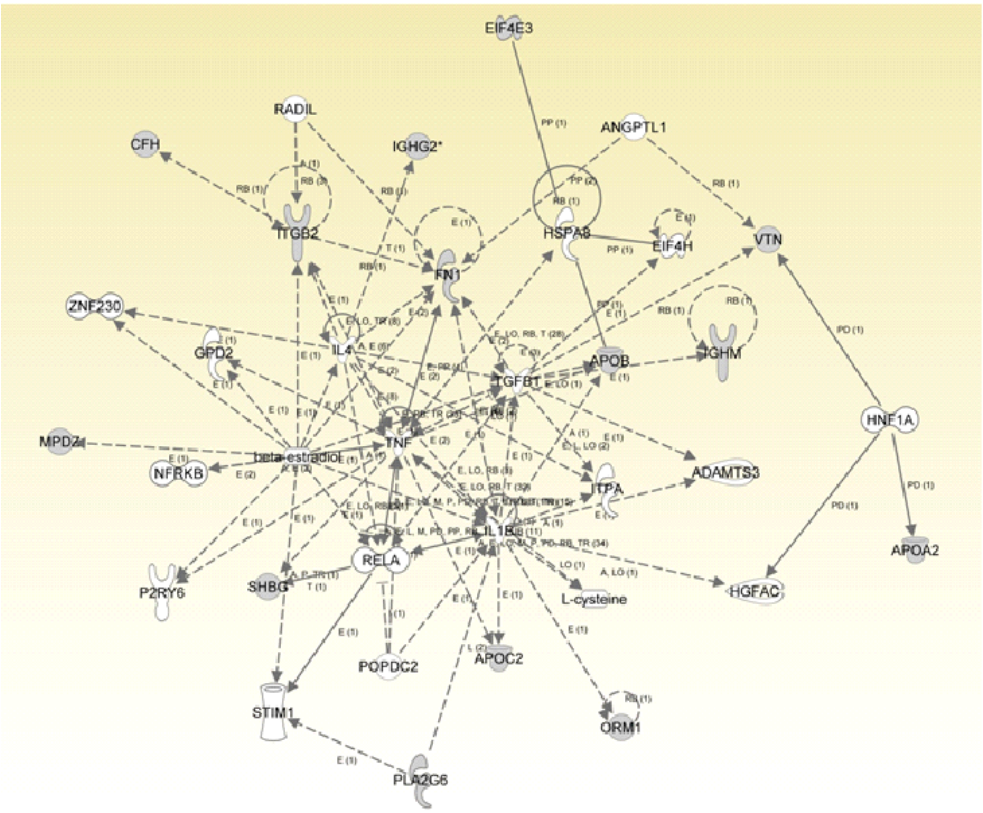

Figure 4

Pathway analysis of network signaling of 32 proteins represented in the ingenuity pathway analysis database, included 14 PPs identified in this study (grey shading). Only thirty-two (32) of the 100 proteins were mapped in IPA database and were implicated in the pathway analysis of multiple signaling networks including Cell-To-Cell Signaling and Interaction, Cell Death and Survival, Cellular Movement, and Immune Cell Trafficking (The image was generated using ingenuity pathway analysis program (IPA Version 49:309) http://qiagen.force.com).