Identification of candidate genes for reproductive traits in cattle using a functional interaction network approach

Francisco Alejandro Paredes-Sánchez [a]

Daniel Trejo-Martínez [b]

Elsa Verónica Herrera-Mayorga [c]

Williams Arellano-Vera [d]

Felipe Rodríguez Almeida [e]

Ana María Sifuentes-Rincón [d*]

[a] Universidad Autónoma de Tamaulipas. IA-UAMM. Mante, México.

[b] Instituto Politécnico Nacional. UPIIZ-, Zacatecas, México.

[c] Universidad Autónoma de Tamaulipas IBI. UAMM. Mante, México

[d] Instituto Politécnico Nacional. Centro de Biotecnología Genómica. Laboratorio de Biotecnología Animal. Blvd. Del Maestro esq. Elías Piña. Col. Narciso Mendoza s/n. Cd. Reynosa, Tam. México.

[e] Universidad Autónoma de Chihuahua. Facultad de Zootecnia y Ecología. Chihuahua, México.

*Corresponding author: asifuentes@ipn.mx

Abstract:

Reproduction is a key element in cattle production systems. Systems biology approaches, including those involving gene networks, have been applied to genetic dissection complex phenotypes in cattle. A set of 385 genes associated with reproductive traits in cattle were included in a protein-protein network analysis to identify and prioritize candidate genes

related to phenotypic differences in cattle reproduction. Genes belonging to the ubiquitin family - Ubiquitin C (*Ubc*, Gene ID: 444874) and Ubiquitin B (*Ubb*, Gene ID: 281370) -had the highest probability of being associated with these traits in cattle. Both proteins were identified as important hubs in a protein-protein interaction network, each having 3,775 interactions of 3,856 possible. Resequencing of the *Ubb* gene coding region to evaluate the presence of SNPs in a discovery population identified the G/T (rs110366695) transversion. This causes emergence of a stop codon and a protein truncated by 287 aa. The allelic frequency distributions found in two beef cattle breeds highlight the promise of further research into the effects of protein truncation and the potential of these proteins as molecular markers for semen quality.

Key words: Bovine, Molecular markers, Semen quality, Ubiquitylation.

The identification of genes encoding complex traits has traditionally been achieved by genome-wide scanning and the candidate gene approach, but these methods do not constitute a reliable strategy for the systematic exploration of a genetic network that causes phenotypic variation in complex traits[1]. Protein networks provide a systems-level overview of genetic organization and enable the functional modules underlying complex traits to be dissected, which facilitates the prediction of novel candidate genes for a trait[2]. In cattle, some approaches related to interaction networks have been utilized to identify candidate genes related to phenotypic differences such as marbling[3], genes involved in estrus (behavior) in dairy cattle[4], and single nucleotide polymorphisms (SNPs) associated with growth traits in Mexican Charolais cows[5].

Reproduction is an essential element of livestock production, and fertility traits are of particularly significant economic importance; it is a very complex process that involves numerous consecutive events, including gametogenesis, fertilization, and early embryo development, that must be accomplished in a well-orchestrated manner to achieve a successful pregnancy[6].

An improved understanding of the mechanisms that control fertility traits at the organ, cellular, and molecular levels could aid the development of strategies to improve and/or monitor fertility[4]. The objective of this work was to conduct a search guided by a functional interaction network to identify key genes controlling reproductive traits in cattle and explore

genetic variation in those identified genes with potential to be associated with reproductive traits.

A literature review was conducted, and the Genie software (http://cbdm-01.zdv.uni-mainz.de/~jfontain/cms/?page_id=6) was used to perform PubMed-based text mining of genes that had been previously associated with bovine reproductive traits (reference genes). To identify and prioritize candidate genes for the functional network, the interactions of the reference genes were extracted, and the degree of association with reproduction (DAR) was calculated for each of the genes in the subnet as follows:

$$DAR = \Sigma_{j \in refgenes} W_{ij} \cdot \Sigma_{j \in refgenes} P_{ij}$$

Where W_{ij} is the weight of the linkage connecting protein i and reference protein j, and P_{ij} is the number of links connecting protein i and reference protein j (excluding itself). Thus, the probability that each of these proteins is associated with reproduction was evaluated based on their interactions with genes whose biological function had already been associated with this trait[5]. To select the candidate genes associated with phenotypic variations in reproductive traits, the DAR score was used to calculate the positive predictive value (PPV), which indicates the likelihood that a gene is associated with reproduction, so the selection criterion was the highest value of PPV obtained in this analysis, i.e., 0.3[5,7].

From candidate genes, the Ubiquitin B (*Ubb*) gene was selected as a target. The genetic variation in the *Ubb* gene was investigated using eleven DNA samples from four different cattle breeds (3 Holstein, 2 Charolais, 3 Brahman and 3 Angus). Primers UBB-F 5'-GAGAGATTTGTGAGAGATCTTGACG-3' and UBB-R 5'-CCATTTTAACCTGTTGAGTACCCA-3' were designed to cover and resequence the bovine *Ubb* gene (GenBank accession number AC_000176.1). The resulting PCR fragments were purified using Exo-SAP-it (Thermo Fisher Scientific, Waltham, MA, USA), and bidirectional sequencing was achieved using the BigDye® Terminator procedure and an ABI PRISM 3100 Genetic Analyzer DNA sequencer (Applied Biosystems, Foster City, CA, USA). Sequences were aligned with ClustalX 2.0.8[8].

The presence of SNPs in the resulting sequences was determined by visual inspection of the sequence chromatograms, and the SNPs were defined according to their presence in the screening population associated with the three expected genotypes.

Amplification-created restriction sites coupled to PCR (PCR-ACRS) were designed to genotype the nonsynonymous SNP rs110366695 identified in the previous sequencing screening, and following PCR, the fragments were digested using 2.5 U of *Hinf* I enzyme

and analyzed on a 2.5% agarose gel. The following digestion patterns were observed: 210+132+130+18 bp (allele G) and 210+155+150 bp (allele T).

A population of sixty-seven young Angus and Charolais bulls were genotyped with PCR-ACRS. The allelic and genotypic frequencies were calculated for each breed, and deviations from Hardy-Weinberg equilibrium were tested by running GENEPOP statistical package version 4.2[9].

A set of 385 reference genes associated with reproductive traits in cattle, through SNPs, expression profiles or their biological function, were identified. According to the PPV, the genes that presented a DAR≥11 had a greater than 33% probability of being associated with reproductive traits in cattle, and those meeting this criterion belonged to the ubiquitin family: Ubiquitin C (*Ubc*; Gene ID: 444874) and B (*Ubb;* Gene ID: 281370).

The importance of these proteins in the topology of the interaction network was determined according to the number of interactions; in this case, *Ubb* and *Ubc* have 3,775 interactions of 3,856 possible, so they are very important hubs. According to BiNGO, a Biological Network Gene Ontology tool (https://www.psb.ugent.be/cbd/papers/BiNGO), in the subnet that forms *Ubb* and *Ubc,* the annotation of Gene Ontology 51094, "positive regulation of developmental process", is overrepresented with a p-value of 4.8 E-09. This result makes sense and relates to reproduction in cattle, as this term of biological process refers to any process that activates or increases the rate or extent of development and whose specific outcome is the progression of an organism over time from an initial condition (e.g., a zygote, or a young adult) to a later condition (e.g., a multicellular animal or an aged adult). Figure 1 depicts the *Ubb* and *Ubc* interaction with 23 reference gene modules, i.e., genes previously associated with reproduction in cattle[8-33].

The ubiquitin protein (Ub) is widespread in all eukaryotic cells, it has a conserved structure which has been interpreted as an indication of its important role in cell metabolism. Through the process of ubiquitination, Ub lead protein degradation and regulate a different biological events including cell cycle progression, membrane-receptor endocytosis, antigen occurrence in the immune system, and even retroviral infection[34]. Ubiquitination is achieved through the covalent binding of 76-AA, 8.5 kDa ubiquitin to the ε-amino group on the Lys-residues of the substrate via the C-terminal AA residuum of ubiquitin (G76). This process requires ATP hydrolysis and a set of ubiquitin-conjugating factors including ubiquitin-activating (UBA) and conjugating (*Ubc*) enzymes[11].
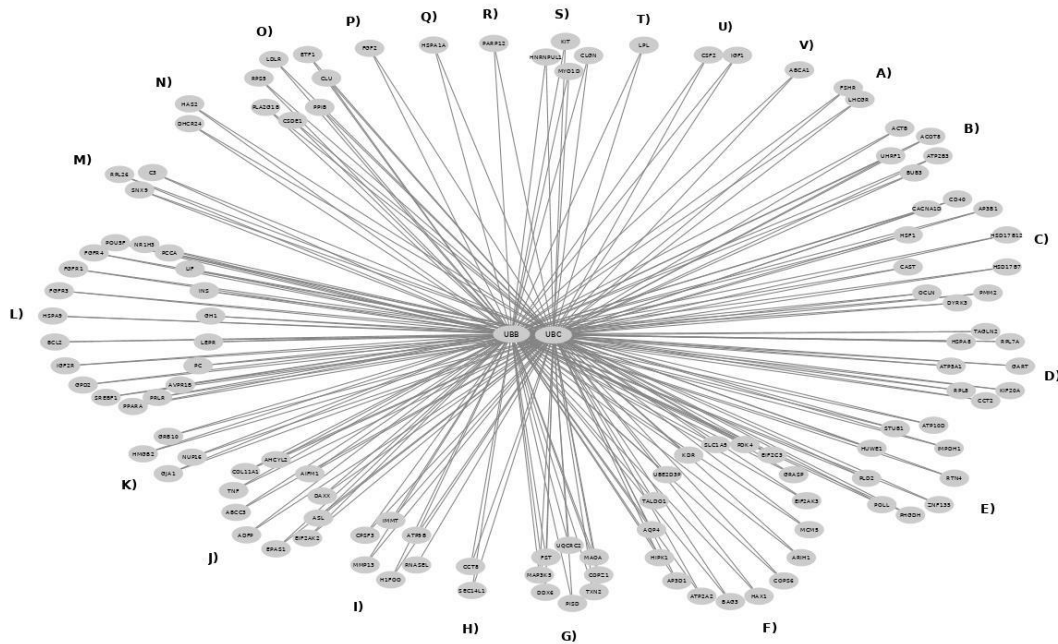
Among multiple functions of Ub system, those involved in developmental and reproductive processes are relevant. In the former case, there are studies in different models including developmental transitions in *Dictyostelium discoideum* and development specificity in *C. elegans*[35,36]. In chicks it has been involved in embryogenesis, and also in the human

myogenesis and brain development[34]. At reproductive processes level, human Ub has been reported as the main protein in seminal plasma and the ubiquitination system has been implicated with fertility problems in humans and other species including cattle[37,38]. It has been reported that a high proportion of ubiquitinated spermatozoa in the ejaculates of different species is related to infertility[38].

In cattle, increased ubiquitin levels have been associated with both increased levels of damage to sperm DNA and with reduced fertility[39]. A negative correlation between sperm ubiquitin and sperm count, swirl and % normal morphology has been reported in bulls and evidence that increased ubiquitin levels in bull sperm are predictive of both poor semen quality and fertility has been also found [39]. This evidence has allowed the use of ubiquitinated spermatozoa as an useful tool to identify fertility problem[40,41].

Even though there is some evidence that the biological mechanisms through ubiquitination systems affect the different species fertility, the process of spermatozoon ubiquitin tagging and the role of this process in sperm biology remains unclarified.

Searching for additional evidence to consider these genes as candidate genes, molecular characterization was achieved on the *Ubb* gene. According to the NCBI database, the length of the gene is 1898 bp, and it contains one exon at position 841 to 1758. In this database, 19 and 15 SNPs have been reported in the coding and noncoding sequences, respectively, and the amplified 1328-bp fragment enabled the identification of 5 SNPs in the study population, 3 (rs109592218, rs110007734 and rs110366695) in the coding region and 2 (rs720990890 and rs439271103) in the noncoding region. The transversion rs110366695 (G/T) located at exon 1 is particularly interesting because it causes a nonsynonymous functional change, and the GAG codon that produces glutamic acid (Glu, E) changes to UAG, which is a stop codon, thus predicting a truncated protein that is 287 amino acids shorter than the nonmutated protein. Figure 1 shows the allelic frequencies of SNP rs110366695 evaluated in the Angus and Charolais beef cattle breeds; allele G showed the highest frequencies (0.542 and 0.750, respectively). Significant departures ($P<0.001$) from Hardy-Weinberg equilibrium were identified for the Angus breed, and lower-than-expected numbers of heterozygotes were found for these loci.

Figure 1: *Ubb* and *Ubc* interaction network

A) Heifer conception rate, SNPs related to interval to insemination. **B)** Effects of the Well of the Well (WOW) system and embryo density on developmental rates, genes differentially regulated in embryos cultured in vitro. **C)** Daughter pregnancy rate, heifer conception rate, cow conception rate. **D)** Term survival in embryos of differentially regulated genes, potential of pretransfer endometrial and embryo gene expression patterns **E)** Blocking apoptosis in bovine embryos, gene differentially regulated embryos treated with CSF2. **F)** Immune function and developmental genes expressed in the endometrium, endometrial genes differentially regulated in lactating cows. **G)** Genes differentially regulated in oocytes compared to 8-cell embryos, global activation of the embryonic genome. **H)** Cow conception rate. **I)** *Igf1* acts in thermoprotection on bovine embryos, genes differentially regulated in embryos treated with *Igf1*. **J)** Endometrial genes differentially regulated in pregnant cows and associations with fertility of lactating dairy cows. **K)** Genes differentially regulated in oocytes compared to blastocysts, candidate genes for the characterization of the development. **L)** Liver genes differentially regulated during the transition period, determination of hepatic adaptations occurring from late pregnancy. **M)** Genes differentially regulated in the oviduct of cows at diestrus compared to estrus. **N)** Genes in cumulus cells regulated by the LH surge, cumulus cells in development and fertility of oocytes. **O)** Genes differentially regulated at different stages of oocyte maturation. **P)** Estimated relative conception rate, net merit, and fat yield. **Q)** Calving rate (beef cattle), net merit, fat percent, and productive life. **R)** SNPs related to interval to insemination. **S)** Embryo development on the blastocyst stage. **T)** Differential regulation in cumulus cells from *in vivo* embryos compared to *in vitro* embryos. **U)** Antiapoptotic in embryos improves embryo competence. **V)** Mammary genes differentially regulated during lactation.

To the current knowledge, there have been no previous molecular studies aimed at evaluating the effects of *Ubb* genetic variation on semen quality despite the demonstrated physiological importance of the *Ubb* gene. The obtained results support the *Ubb* gene as a strong candidate gene with genetic variations to be tested for association with reproductive traits.

Unfortunately, in Mexico, phenotyping for reproductive traits is not a common practice, and additional efforts must be made to prepare a wide database that allows confirmation of its genetic influence on these traits, particularly of the transversion rs110366695 (G/T).

An analysis-based protein-protein interaction network has been previously validated as a useful tool for identifying causal genes associated with economic traits in bovines and other species. The obtained results provide information about the potential of *Ubb* and *Ubc* as candidate genes for reproductive traits, particularly semen quality, and justify further research aimed at exploring both the effects of protein truncation and its potential as a molecular marker.

Acknowledgments

Literature cited:

1. Jiang Z, Ott TL. Reproductive genomics in domestic animals. Iowa, USA: Wiley-Blackwell; 2010.

2. Lee T, Hwang S, Kim CY, Shim H, Kim H, Ronald P, *et al*. WheatNet: A genome-scale functional network for hexaploid bread wheat, *Triticum aestivum*. Mol Plant 2017;(8):1133-1136.

3. Lim D, Kim NK, Park HS, Lee SH, Cho YM, Oh SJ, *et al*. Identification of candidate genes related to bovine marbling using protein-protein interaction networks. Int J Biol Sci 2011;(7):992-1002.

4. Hulsegge I, Woelders H, Smits M, Schokker D, Jiang L, Sørensen P. Prioritization of candidate genes for cattle reproductive traits, based on protein-protein interactions, gene expression, and text-mining. Physiol Genomics 2013;(10):400-406.

5. Paredes-Sánchez FA, Sifuentes-Rincón AM, Segura CA, García PCA, Parra BGM, Ambriz MP. Associations of SNPs located at candidate genes to bovine growth traits, prioritized with an interaction networks construction approach. BMC Genet 2015;(91):1-12.

6.- Han Y, Peñagaricano F. Unravelling the genomic architecture of bull fertility in Holstein cattle. BMC Genet 2016;(1):143.

7. Garza-Brenner E, Sifuentes-Rincón AM, Randel RD, Paredes-Sánchez FA, Parra-Bracamonte GM, et al. Association of SNPs in dopamine and serotonin pathway genes and their interacting genes with temperament traits in Charolais cows. J Appl Genet 2016;(3):1-9.

8. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al*. Clustal W and Clustal X version 2.0. Bioinformatics 2007;(21):2947-2848.

9. Raymond M, Rousset F. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. J Heredity 1995;(3):248-249.

10.- Cory AT, Price CA, Lefebvre R, Palin MF. Identification of single nucleotide polymorphisms in the bovine follicle-stimulating hormone receptor and effects of genotypes on superovulatory response traits. Anim Genet 2013;(2):197-201.

11.- Hoelker M, Rings F, Lund Q, Ghanem N, Phatsara C, Griese J. Effect of the microenvironment and embryo density on developmental characteristics and gene expression profile of bovine preimplantative embryos cultured in vitro. Reproduction 2009;(3):415-425.

12.- Cochran SD, Cole JB, Null DJ, Hansen PJ. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genet 2013;(49):14-49.

13.- Fair T, Carter F, Park S, Evans AC, Lonergan P. Global gene expression analysis during bovine oocyte *in vitro* maturation. Theriogenology 2007;(68):91-97.

14.- Salilew WD, Hölker M, Rings F, Ghanem N, Ulas-Cinar M, Peippo J, *et al*. Bovine pretransfer endometrium and embryo transcriptome fingerprints as predictors of pregnancy success after embryo transfer. Physiol Genomics 2010;(2):201-218.

15.- Loureiro B, Oliveira LJ, Favoreto MG, Hansen PJ. Colony-stimulating factor 2 inhibits induction of apoptosis in the bovine preimplantation embryo. Am J Reprod Immunol 2011;(65):578-588.

16.- Cerri RL, Thompson IM, Kim IH, Ealy AD, Hansen PJ, Staples CR, *et al*. Effects of lactation and pregnancy on gene expression of endometrium of Holstein cows at day 17 of the estrous cycle or pregnancy. J Dairy Sci 2012;(10):5657-575.

17.- Fayad T, Lévesque V, Sirois J, Silversides DW, Lussier JG. Gene expression profiling of differentially expressed genes in granulosa cells of bovine dominant follicles using suppression subtractive hybridization. Biol Reprod 2004;(2):523-533.

18.- Misirlioglu M, Page GP, Sagirkaya H, Kaya A, Parrish JJ, First NL, *et al*. Dynamics of global transcriptome in bovine matured oocytes and preimplantation embryos. Proc Natl Acad Sci 2006;(50):18905-18910.

19.- Cochran SD, Cole JB, Null DJ, Hansen PJ. Single nucleotide polymorphisms in candidate genes associated with fertilizing ability of sperm and subsequent embryonic development in cattle. Biol Reprod 2013;(3):1-7.

20.- Bonilla AQ, Oliveira LJ, Ozawa M, Newsom EM, Lucy MC, Hansen PJ. Developmental changes in thermoprotective actions of insulin-like growth factor-1 on the preimplantation bovine embryo. Mol Cell Endocrinol 2011;(1-2):170-179.

21.- Mamo S, Sargent CA, Affara NA, Tesfaye D, El-Halawany N, Wimmers K, Gilles M, Schellander K, Ponsuksili S. Transcript profiles of some developmentally important genes detected in bovine oocytes and in vitro-produced blastocysts using RNA amplification and cDNA microarrays. Reprod Domest Anim 2006;(6):527-534.

22.- Graber M, Kohler S, Kaufmann T, Doherr MG, Bruckmaier RM, van Dorland HA. A field study on characteristics and diversity of gene expression in the liver of dairy cows during the transition period. J Dairy Sci 2010;(1):5200-5215.

23.- Bauersachs S, Rehfeld S, Ulbrich SE, Mallok S, Prelle K, Wenigerkind H, *et al*. Monitoring gene expression changes in bovine oviduct epithelial cells during the oestrous cycle. J Mol Endocrinol 2004;(2):449-466.

24.- Assidi M, Dieleman SJ, Sirard MA. Cumulus cell gene expression following the LH surge in bovine preovulatory follicles: potential early markers of oocyte competence. Reproduction 2010;(6):835-852.

25.- Salhab M, Tosca L, Cabau C, Papillier P, Perreau C, Dupont J, *et al*. Kinetics of gene expression and signaling in bovine cumulus cells throughout IVM in different mediums in relation to oocyte developmental competence, cumulus apoptosis and progesterone secretion. Theriogenology 2011;(1):90-104.

26.- Khatib H, Monson RL, Huang W, Khatib R, Schutzkus V, Khateeb H, et al. Short communication: Validation of in vitro fertility genes in a Holstein bull population. J Dairy Sci 2010;(93):2244-2249.

27.- Rosenkrans Jr C, Banks A, Reiter S, Looper M. Calving traits of crossbred Brahman cows are associated with Heat Shock Protein 70 genetic polymorphisms. Anim Reprod Sci 2010;(3-4):178-182.

28.- Pimentel EC, Bauersachs S, Tietze M, Simianer H, Tetens J, Thaller G, *et al*. Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNPs within candidate genes derived by expression profiling. Anim Genet 2011;(3):251-262.

29.- Gad A, Besenfelder U, Rings F, Ghanem N, Salilew-Wondim D, Hossain MM, et al. Effect of reproductive tract environment following controlled ovarian hyperstimulation treatment on embryo development and global transcriptome profile of blastocysts: implications for animal breeding and human assisted reproduction. Hum Reprod 2011;(7):1693-1707.

30.- Tesfaye D, Worku D, Rings F, Phatsara C, Tholen E, Schellander K, *et al*. Identification and expression profiling of microRNAs during bovine oocyte maturation using heterologous approach. Mol Reprod Dev 2009;(7):665-677.

31.- Jousan FD, Hansen PJ. Insulin-like growth factor-I promotes resistance of bovine preimplantation embryos to heat shock through actions independent of its anti-apoptotic actions requiring PI3K signaling. Mol Reprod Dev 2007;(2):189-196.

32.- Mani O, Körner M, Sorensen MT, Sejrsen K, Wotzkow C, Ontsouka CE, *et al*. Expression, localization, and functional model of cholesterol transporters in lactating and nonlactating mammary tissues of murine, bovine, and human origin. Am J Physiol Regul Integr Comp Physiol 2010;(2):642-654.

33.- Fortes MR, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, *et al*. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. J Anim Sci. 2011;(6):1669-1683.

34.- Bebington C, Doherty FJ, Fleming SD. The possible biological and reproductive functions of ubiquitin. Hum Reprod Update 2001;(1):102-111.

35.- Clark A, Nomura A, Mohanty S, Firtel RA. A ubiquitin-conjuganting enzyme is essential for developmental transitions in *Dictiostelium*. Mol Biol Cell 1997;(8):1989-2002.

36.-Zhen M, Schein JE, Baille DL, Peter E, Candido M. An essential ubiquitin conjugating enzyme with tissue and developmental specificity in the nematode *C. elegans.* EMBO J 1996;(15):3229-3237

37.-Muratori M, Marchiani S, Forti G, Baldi E. Sperm ubiquitination positively correlates to normal morphology in human semen. Hum Reprod 2005;(20):1035-1043.

38.- Sutovsky P, Geary T, Baska KM, Manandhar G, Feng D, Lovercamp KW, Sutovsky M. Ubiquitin as an objective marker of semen quality and fertility in bulls. Proc Nebraska Appl Reprod Strat in Beef Cattle 2004;185-199.

39.- Rodríguez-Lozano I, Ávalos-Rodríguez A, Castillo-Juárez H, Borderas-Tordesillas F, Roa-Vidal JJ, Rosales-Torres AM. Percentage of ubiquinated spermatozoa does not correlate with fertilizing capacity of thawed bovine semen. Reprod Dom Anim 2013;(1):27-31.

40.- Sutovsky P, Terrada Y, Schatten G. Ubiquitin-based sperm assay for the diagnosis of male factor infertility. Hum Reprod 2001;(2):250–258.

41.- Sutovsky P, Hauser R, Sutovsky M. Increased levels of sperm ubiquitin correlate with semen quality in men from an andrology laboratory clinic population. Hum Reprod 2004;(3):628–638.