HOW TO SELECT A SHORTLIST OF GENETIC MARKERS FOR MALE INFERTILITY?

In reproductive genetics, the application of high throughput techniques such as Whole Exome Sequencing has extended the number of pathogenic gene variants associated with male infertility, opening the debate about the utility of the clinical implementation of diagnostic gene panels. But, which genes should we include in these panels to obtain a high diagnostic yield and high specificity and sensitivity?

So far to date, the most common strategy used for gene selection has been literature search in the MEDLINE-Pubmed database. Nevertheless, one of the limitations of such a strategy is the difficulty in classifying and in selecting the genetic variants according to their degree of association with male infertility, that is, the difficulty of ordering them from the most to the least predictive.

This article published in *Cytogenomics and Genome Research* introduces an interesting alternative. Based on regression model analysis, the authors have developed a ranking of genes according to their potential as fertility biomarkers. The analysis takes as a starting point genes associated with male infertility in knockout mice (n=251) and genes selected from studies in infertile men (n=52). Both lists are further analyzed considering sequence gene conservation, testis transcription level, and their connectivity in a protein-protein interaction network. Using this strategy, the authors identify 9 genes with the highest predictive values suggesting their potential as candidates for becoming a part of a gene panel. Besides, additional genes with high values are also suggested, compiling a final list of 22 markers. Interestingly, the predictive value of the selected genes was validated using quantitative proteomics in spermatozoa from fertile and infertile patients, and differential expression was observed in 23% of the genes reinforcing their potential as infertility markers.

Although additional work should be done to validate the list provided in this article, some characteristics of this selection method (use of public database information, quantitative interpretation of the results, and the possibility to add new predictive variables) make it an interesting and novel approach.