Figure 1: Flow diagram of the analytical procedure followed to perform this analysis. The dataset consists in the expression data and the mutational status of the patients. The genes are ranked according to their discriminatory power measured by their Fisher's ratio and absolute fold change values. Their predictive accuracy is established through leave-one-out-cross validation (LOOCV) and different hold-out analysis. Finally, the Correlation Networks and the Pathway Analysis using Gene Analytics are established and serve to find the biological processes impacted by the NOP16 mutation and how the most discriminatory genes are related.

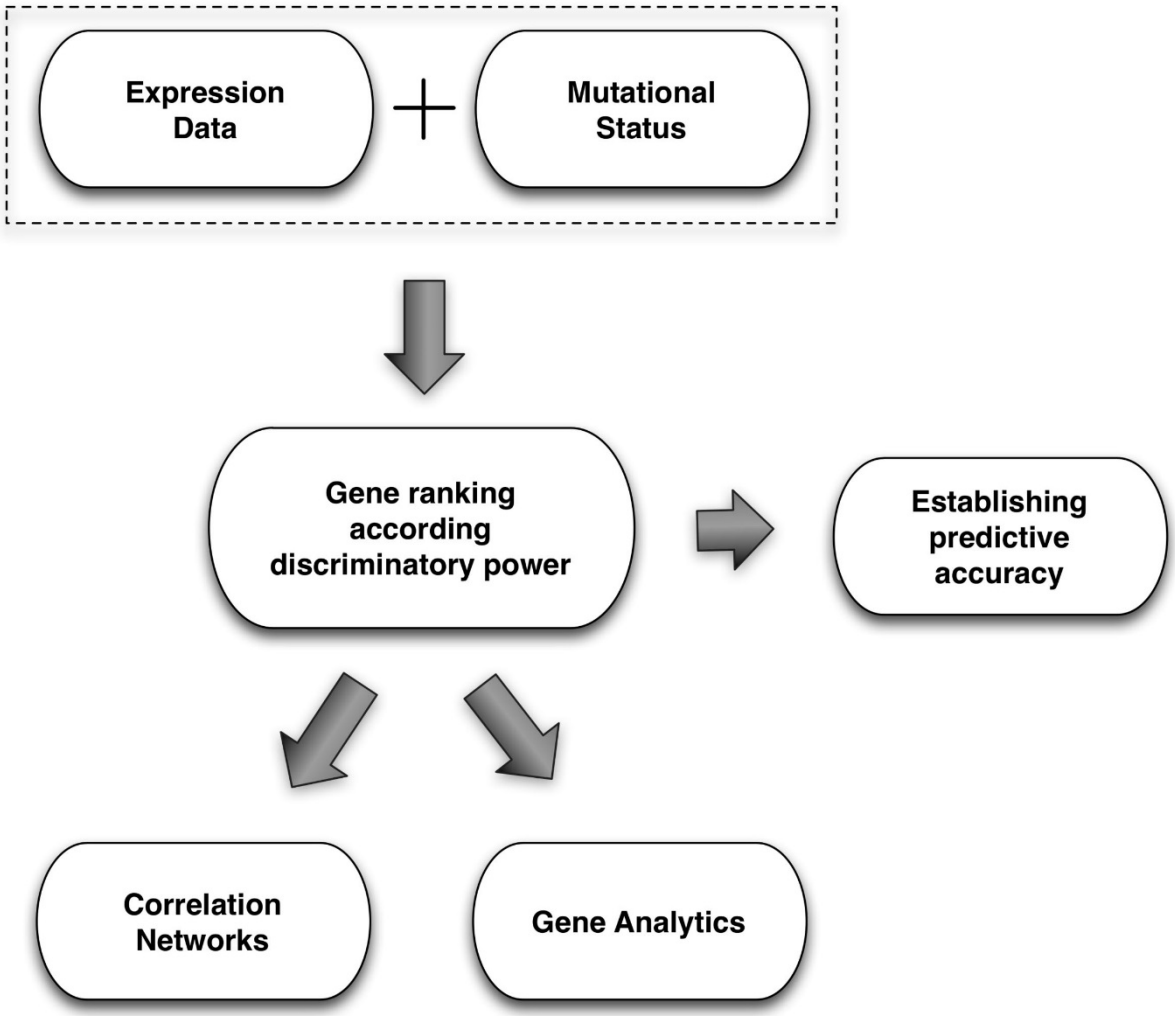
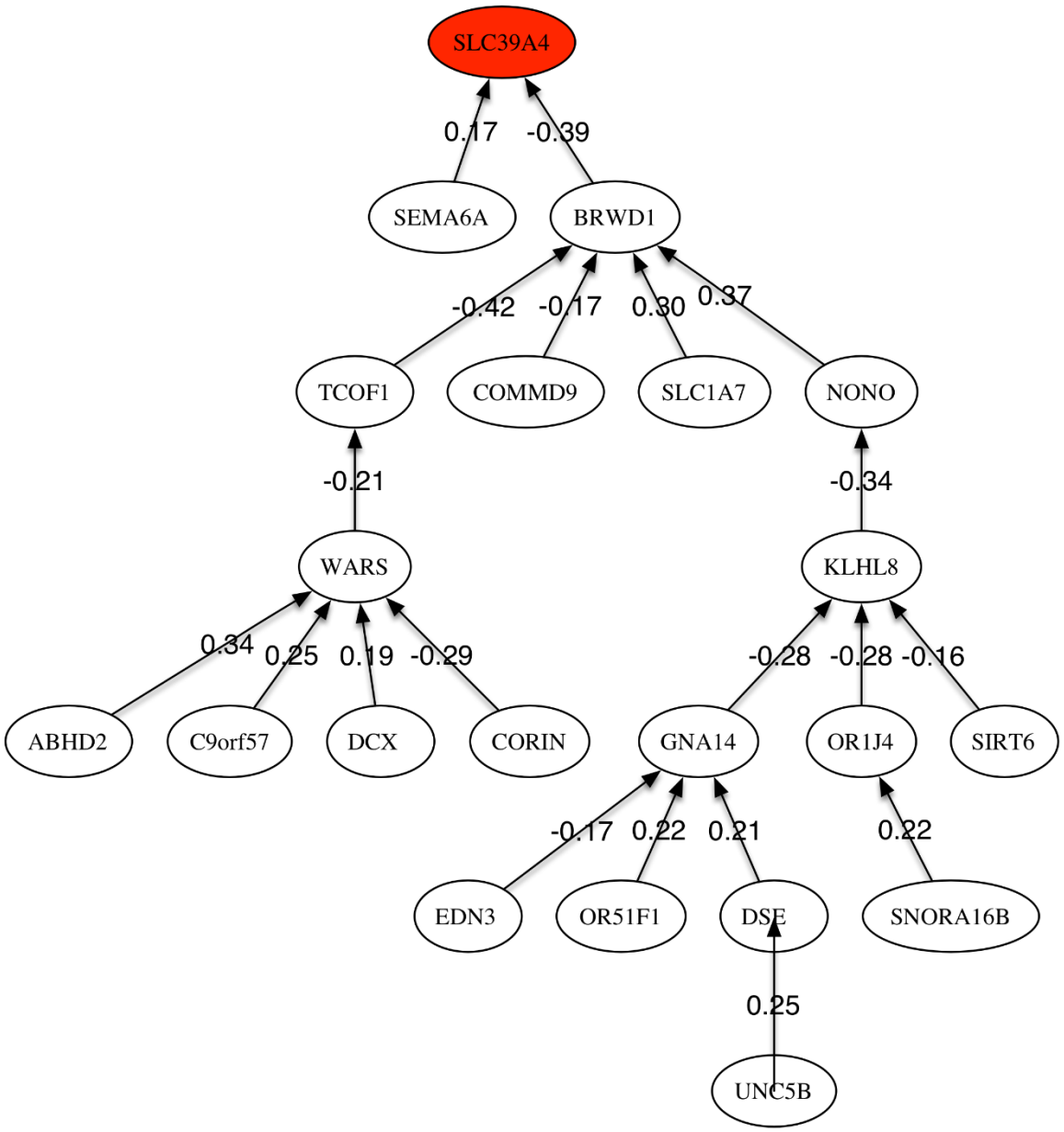


Figure 2: Correlation network of the most discriminatory genes (Fisher's ratio) for the NOP16 mutational status prediction using the Pearson correlation coefficient. The tree is restricted to the most discriminatory genes for

the sake of visual clarity. The tree begins by the header gene (*SCLC39A* in this case) and the branches show the nodes (genes) that control expression and have the highest absolute Pearson correlation coefficient.



Fernández-Martínez et al., J Mol Genet Med 2017, 11:4
DOI: 10.4172/1747-0862.1000295

Figure 3: Correlation networks of the most discriminatory genes (Fold-change) for the *NOP16* mutational status prediction using the Pearson correlation coefficient. The tree is restricted to the most discriminatory genes for the sake of visual clarity.

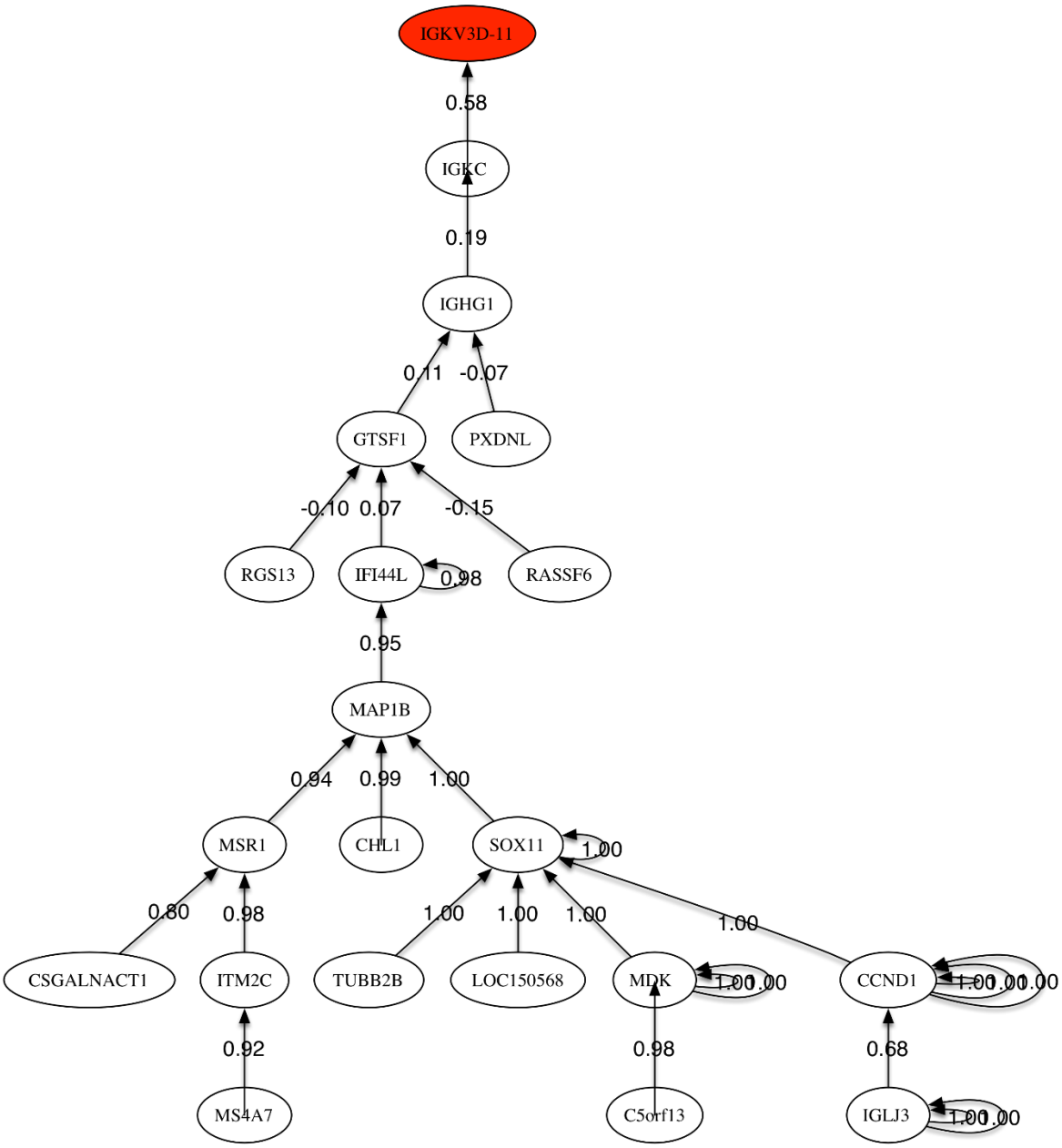


Figure 4: Hold-Out stability analysis. Cumulative probability functions for the predictive accuracy of the small-scale genetic signature of the *NOP16* mutation. The minimum scale signature is very stable since the accuracy varies from 97.5% to 100%.

Journal of Molecular and Genetic Medicine

Fernández-Martínez et al., J Mol Genet Med 2017, 11:4
DOI: 10.4172/1747-0862.1000295

Open Access

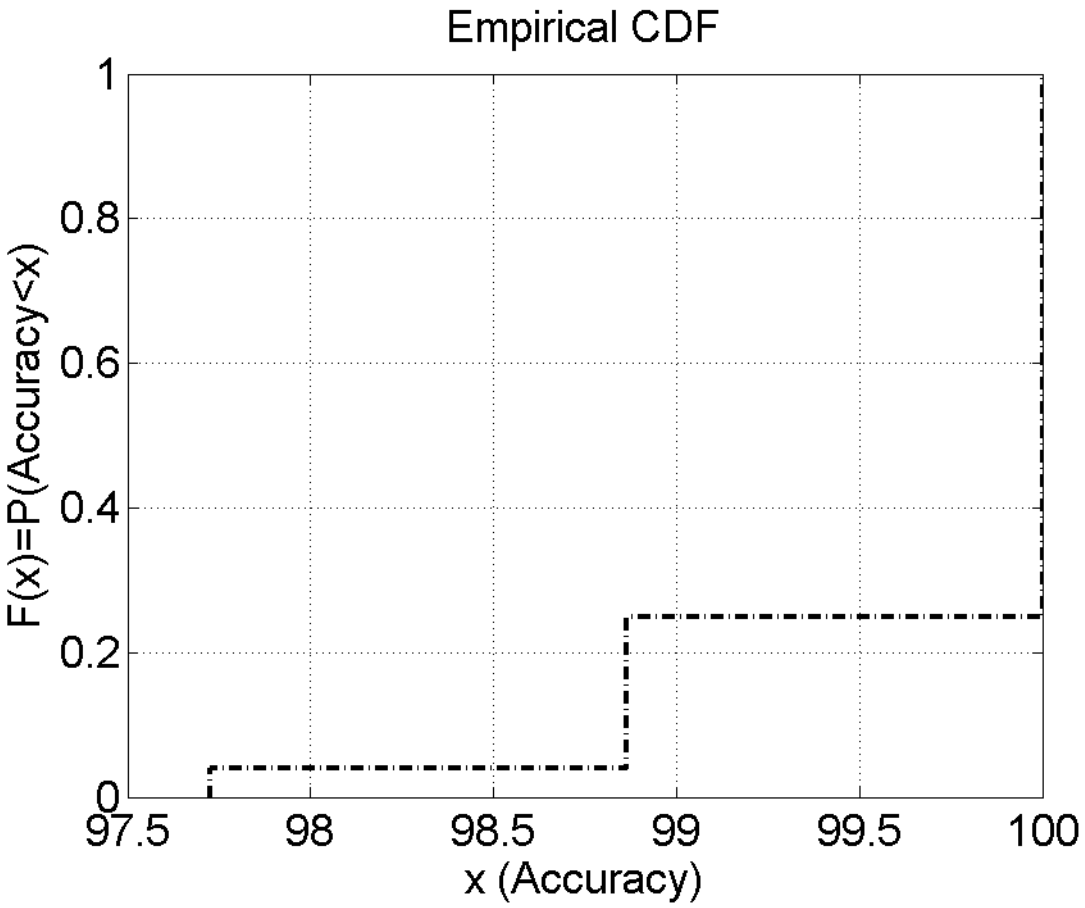
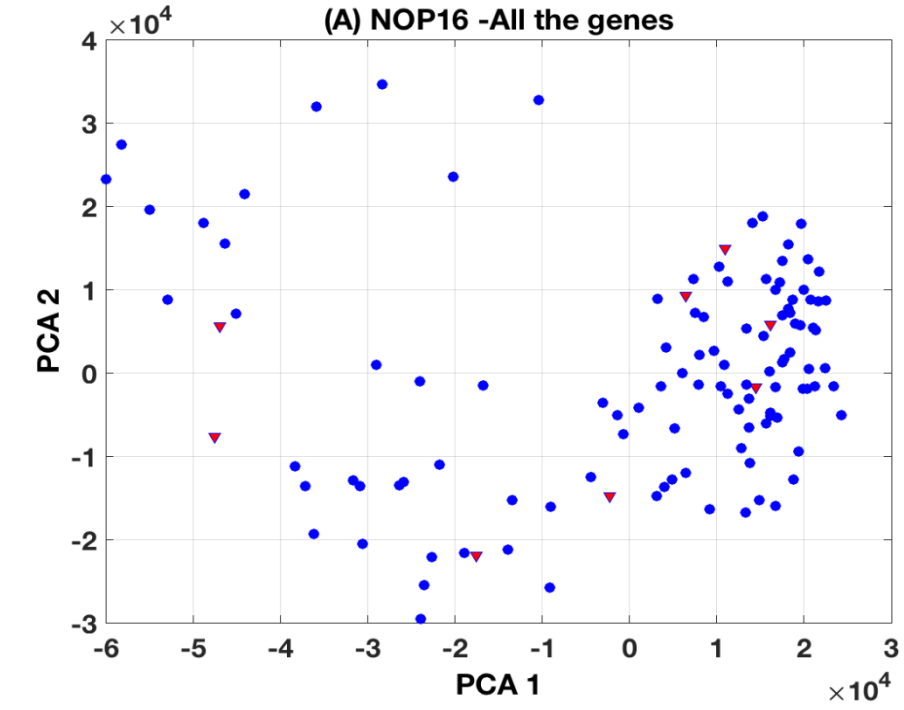


Figure 5: NOP16 mutational status. A) 2D PCA plot using all the genes. B) 2D PCA plot in the small-scale signature. This graph shows that restricting the expression data to the minimum scale signature improves the separability of two classes using unsupervised classification (Principal Component Analysis). PCA1 and PCA2 stand for the two first PCA coordinates of the different samples in the reduced PCA two-dimensional space.



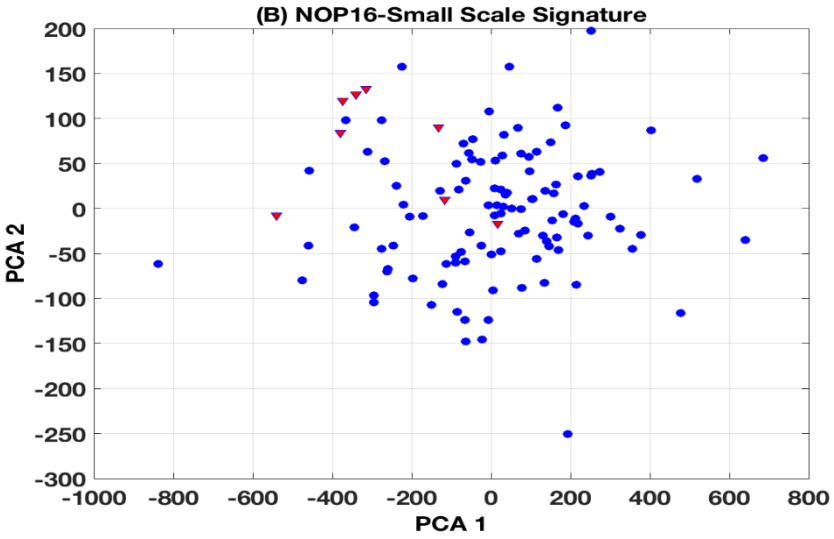


Figure 6: Cyclic diagram showing the intersection between the most discriminatory genes in each mutation: *NOTCH1*, *SF3B1*, *IGHV* and *NOP16*. The four mutations share only 2 genes (*IGHG1* and *RGS13*) within the intersection lists of discriminatory genes of each mutation.

Journal of Molecular and Genetic Medicine

Fernández-Martínez et al., J Mol Genet Med 2017, 11:4
DOI: 10.4172/1747-0862.1000295

Open Access