

# Exploring Genes with the Most Disease Associations

Nidhi Bendre  
Northeastern University

**What are the five genes with the highest number of disease associations, and on which chromosomes are these genes located?**

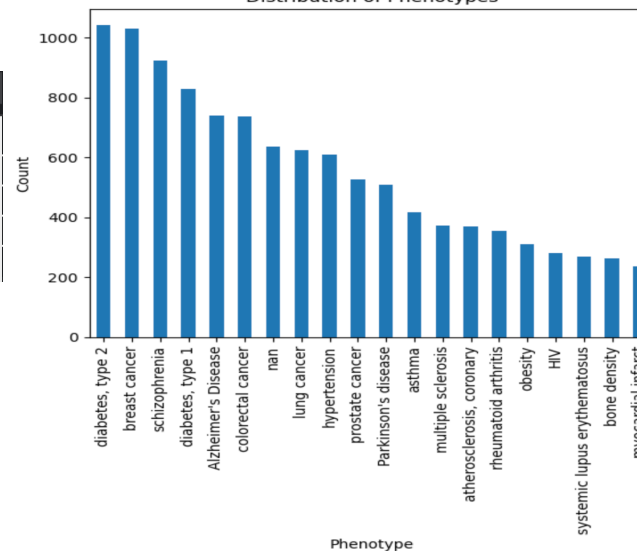
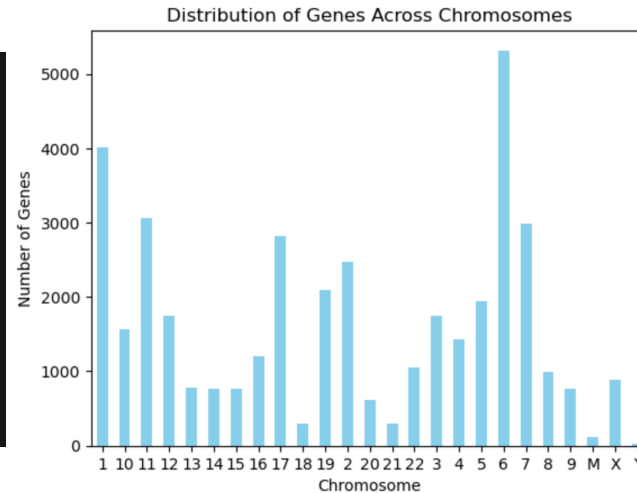
SQL query:

```
SELECT chromosome,  
       gene,  
       COUNT(DISTINCT phenotype) AS number_of_diseases_per_gene  
FROM gad  
GROUP BY chromosome, gene  
ORDER BY number_of_diseases_per_gene DESC  
LIMIT 5;
```

Output:

chromosome	gene	number_of_diseases_per_gene
6	TNF	342
17	ACE	292
6	HLA-DRB1	288
7	NOS3	280
6	HLA-DQB1	267

Visualizations



Results and conclusions

The SQL output shows that the top 5 genes that contribute the most to diseases are TNF, ACE, HLA-DRB1, NOS3, and HLA-DQB1. These genes may play crucial roles in multiple diseases, making them potential targets for further research and drug development.

Based on the distribution of genes across chromosomes, chromosome 6 contains the highest number of genes associated with diseases among the top five genes.

From the distribution of phenotypes chart we can see that diabetes type 2, breast cancer, schizophrenia, diabetes type 1, and Alzheimer's Disease are among the most frequently associated phenotypes.

Overall, the presence of TNF as the gene with the highest number of disease associations suggests its central role in various diseases. The high number of genes associated with chromosome 6 indicates its significance in genetic research, possibly housing genes that are involved in a wide range of diseases. The distribution of phenotypes highlights the diversity of diseases in the dataset. Researchers may focus more on phenotypes like diabetes, cancer, and neurodegenerative disorders due to their higher frequency of gene associations.

These insights can guide researchers in exploring potential therapeutic interventions and personalized medicine approaches, but while doing so the biases and lack of genuineness (if any) should be taken into consideration.

Becker et al., 2004. The Genetic Association Database. Nature Genetics 36(5):431-2



Northeastern University