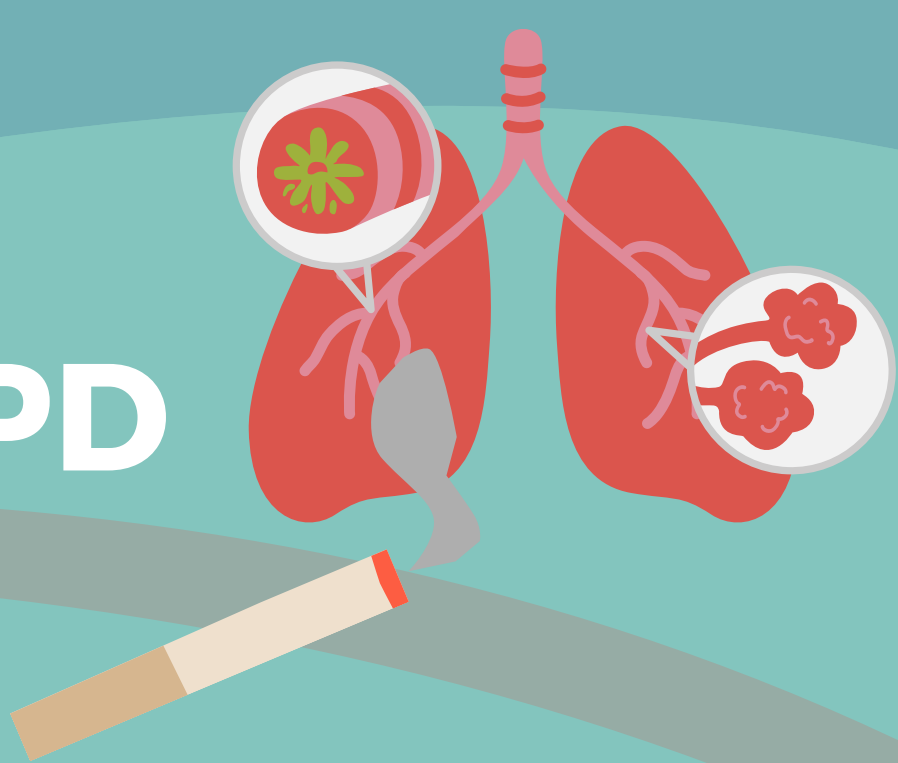


How do gene expression profiles differ in COPD patients by disease severity and smoking exposure?

Annelieke Bos, Onoufrios Fragkioudakis, Robin van Nes, Rosa Rutte, Monique ter Schure, Terri Thahmaz



INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a prevalent disease, causing about 3.5 million deaths worldwide in 2021 (Elflein, 2025).

The main demographic affected is smokers, as 20% of smoking individuals will eventually develop the disease (Terzikhan et al., 2016).

The disease is characterized by chronic inflammation that leads to **airway narrowing, limited airflow**, and, in advanced stages, **tissue destruction** and eventual **collapse of the respiratory system and death** (Agarwal, Raja, & Brown, 2023).

RESULTS: DGE CONTROL VS COPD

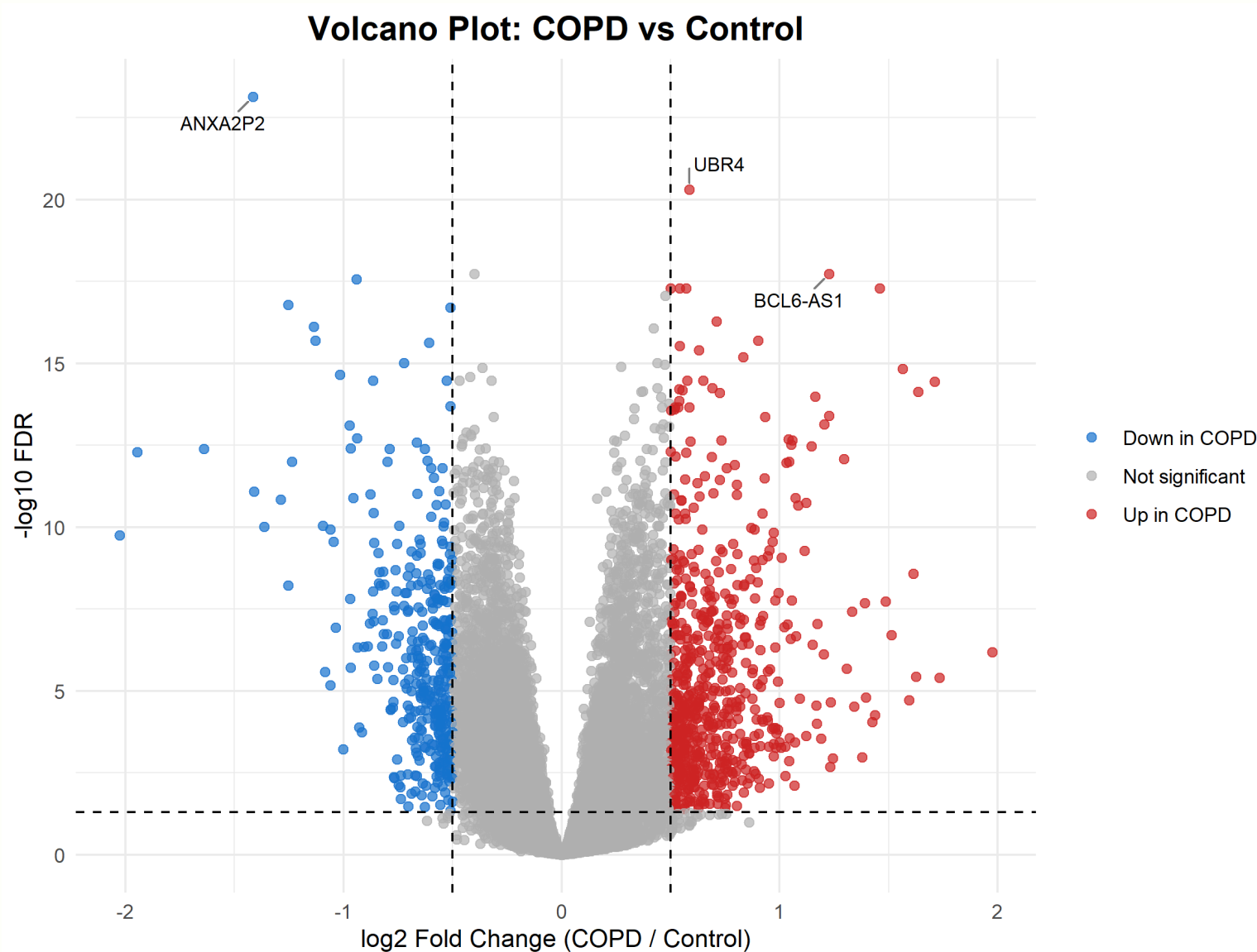


Figure 1. Differential Gene Expression Analysis of COPD Patients Compared to Healthy Controls. A volcano plot was created that displays differentially expressed genes between COPD patients and healthy controls. The Benjamini-Hochberg correction has been used due to multiple testing in addition to the quasi-likelihood F-test (glmQLFTest) that determines significance. Grey points in the plot represent genes without significant differential expression between the two groups. Red points highlight genes that are significantly upregulated in COPD, while blue points indicate genes significantly downregulated in COPD. The top three most statistically significant genes are labeled, which are **ANXA2P2**, **UBR4** and **BCL6-AS1**.

RESULTS: DGE MILD VS MODERATE

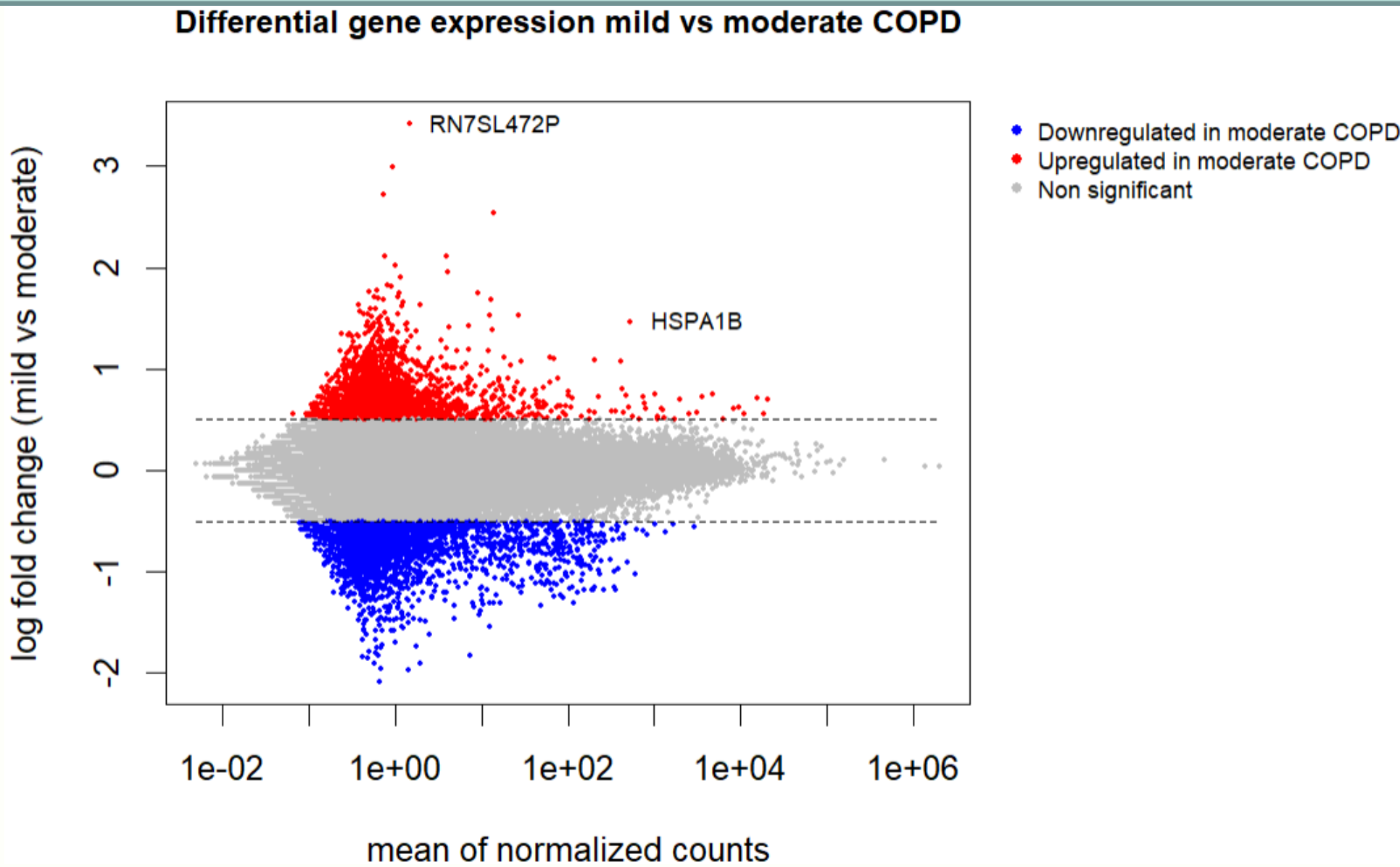


Figure 2. Differential gene expression in mild vs moderate COPD. In R, an MA plot was made, which shows the differential gene expression between mild and moderate COPD. Normalization was done with DESeq. The X-axis shows the average gene expression of the samples on a logarithmic scale. The Y-axis shows the log fold change (moderate COPD/ mild COPD). A non-significant area was added between ± 0.5 log fold change. Each dot represents a gene. The non-significant genes are grey, upregulated in moderate COPD are red, and downregulated in moderate COPD are blue. The genes with differential gene expression with a p-value < 0.05 were annotated. These are the upregulated genes **RN7SL472P** and **HSPA1B**.

RESULTS: IMPACT OF SMOKING EXPOSURE ON COPD SEVERITY AND GENE EXPRESSION

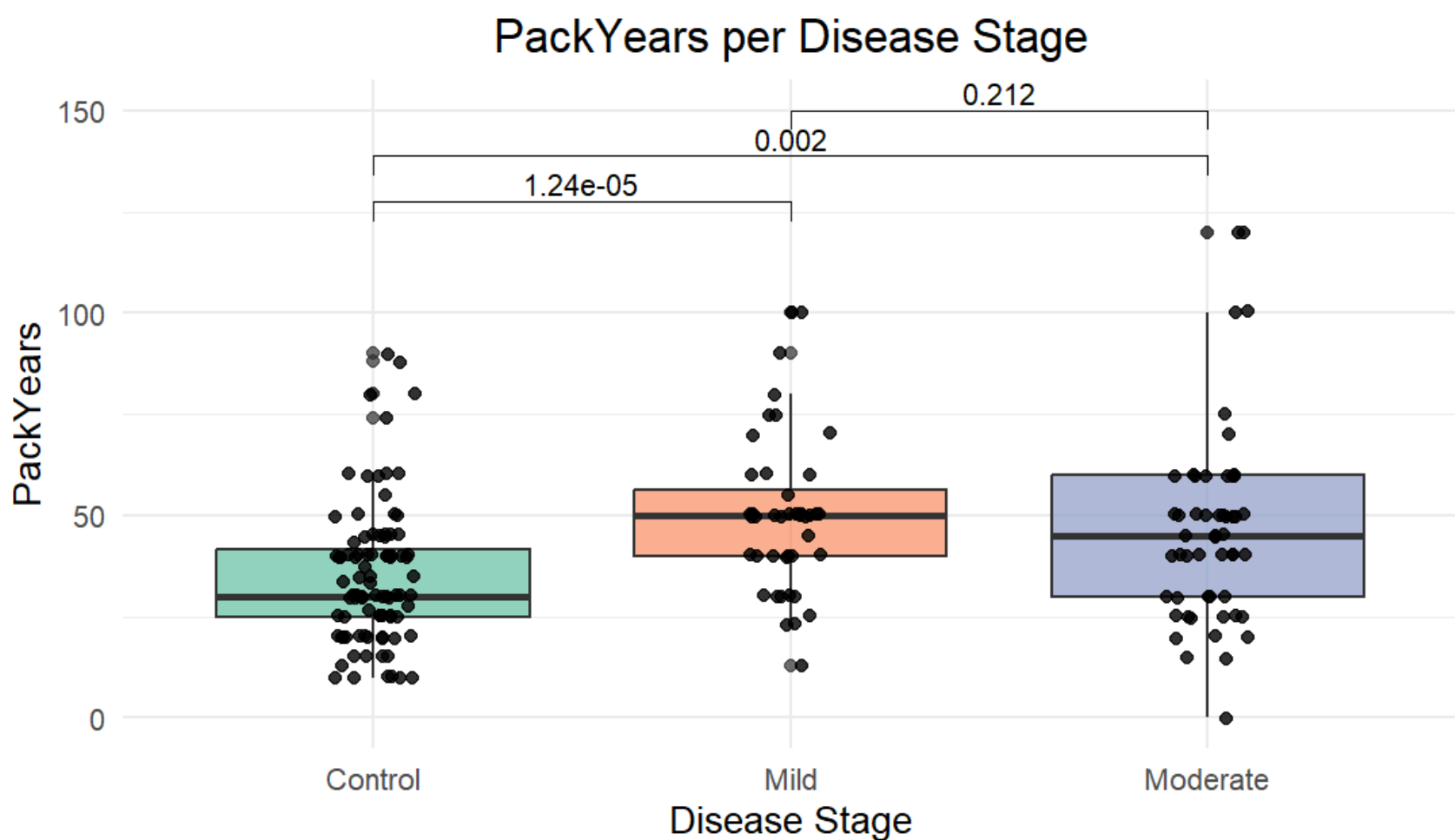


Figure 3. Distribution of smoking exposure (PackYears) across COPD disease stages (stage 0-2). The x-axis represents the stage of disease. The y-axis represents the PackYears. Black dots represent individual samples. Pairwise comparison between groups was performed using the Wilcoxon rank-sum test. Additionally, adjusted p-values (Benjamini-Hochberg correction with a threshold of $p < 0.05$) are shown. Significant differences in PackYears were observed between the control group and both Stage 1 and Stage 2 COPD, whereas no significant difference was detected between Stage 1 and Stage 2.

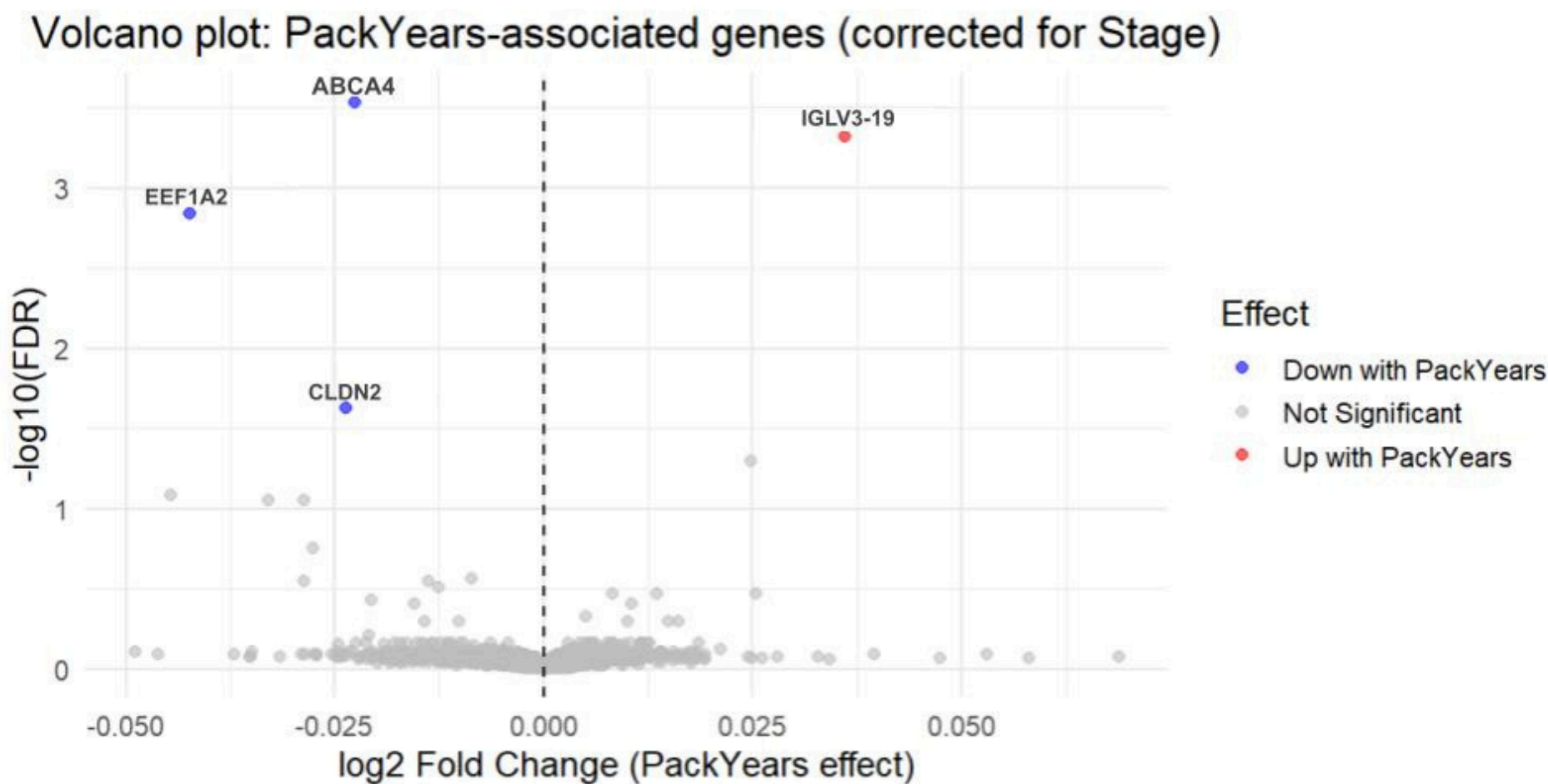


Figure 4: PackYears-associated genes, corrected for COPD disease stage. The x-axis represents the estimated $\log_2$ fold change in gene expression per unit increase in PackYears. The y-axis represents the $-\log_{10}$ of the adjusted p-value (FDR). Red dots indicate significantly upregulated genes, and blue dots indicate significantly downregulated genes (FDR < 0.05). Grey dots represent non-significant genes. Vertical dashed line denotes no change in expression ($\log_2FC = 0$).

- Gene **IGLV3-19** shows upregulation. Genes: **ABCA4**, **EEF1A2**, and **CLDN2**, show downregulation.

FUNCTIONAL OVERVIEW OF MAJOR GENES

UBR4 encodes a protein that is an E3 ubiquitin-protein ligase. It attaches to retinoblastoma-associated proteins and calcium-bound calmodulin (Dryer et al., 2025). It is also significantly expressed in COPD in the findings of Kim et al. (2025).

BCL6-AS1 is classified as a lncRNA, which is an immune regulator for T follicular cells differentiation (Dryer et al., 2025). BCL6 is regulated during hypoxia-induced pulmonary hypertension, where it increases T follicular cell activity, disrupting immunological balance (Li. et al., 2022).

HSPA1B is a heat shock protein, which are usually elevated in patients with COPD (Cappello et al., 2015).

EEF1A2 is an eukaryotic translation elongation factor found to be associated with smoke-induced COPD (Gao, 2023)

CLDN2 is a major integral membrane protein localized exclusively at tight junctions. Claudins are expressed in an organ-specific manner and regulate tissue-specific physiologic properties of tight junctions (Elkouby-Naor, 2025).

IGLV3-19 enables antigen-binding activity and is predicted to be involved in the immune response. Active smoking induced an adaptive immune response with relatively high levels of (class-switched) memory B-cells (Dryer et al., 2025).

CONCLUSION

Our DGE analysis revealed significantly expressed genes. Some of the observed genes could directly affect COPD by influencing the lung tissues, air flow, or inflammation. Others, however, could be used as biomarkers or therapeutic targets. Smoking exposure correlates with both COPD onset and distinct gene expression changes associated with smoke-induced COPD and inflammation. Future studies should assess whether these molecular changes are reversible and if they can predict disease progression. Additionally, they should aim to reveal the exact mechanisms by which these genes affect COPD manifestation.

References

Agarwal, A. K., Raja, A., & Brown, B. D. (2023). Chronic obstructive pulmonary disease. In StatPearls. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK559281/>

Dyer, S. C., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Barrera-Enriquez, V. P., Becker, A., Bennett, R., Beracochea, M., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Campbell, L. I., Martinez, M. C., Charkhchi, M., ... Yates, A. D. (2024). Ensembl 2025. Nucleic Acids Research, 53(D1). <https://doi.org/10.1093/nar/gkac1071>

Elflein, J. (2025, January 21). Chronic obstructive pulmonary disease (COPD) worldwide - Statistics & facts. Statista. <https://www.statista.com/topics/33124/chronic-obstructive-pulmonary-disease-copd-worldwide/>

Kim, W. J., Lim, J. H., Lee, J. S., Lee, S., Kim, J. H., & Oh, Y. (2015). Comprehensive analysis of transcriptome sequencing data in the lung tissues of COPD subjects. International Journal of Genomics, 2015, 1-9. <https://doi.org/10.1155/2015/206937>

Terzikhan, N., Verhamme, K. M., Hofman, A., Stricker, B. H., Brusselle, G. G., & Lahousse, L. (2016). Prevalence and incidence of COPD in smokers and non-smokers: The Rotterdam Study. European Journal of Epidemiology, 31(8), 785-792. <https://doi.org/10.1007/s10654-016-0132-z>

Li, C., Liu, P., Yao, H., Zhu, H., Zhang, S., Meng, F., Li, S., Li, G., Peng, Y., Gu, J., Zhu, L., Jiang, Y., & Dai, A. (2022). Regulatory B cells protect against chronic hypoxia-induced pulmonary hypertension by modulating the Tfh/Tfr immune balance. Immunology, 168(4), 580-596. <https://doi.org/10.1111/imm.13589>

Cappello, E., Macario, A. J., & Di Stefano, A. (2015). Hsp27 and Hsp70 in chronic obstructive pulmonary disease: certainties vs doubts. Cell stress & chaperones, 20(5), 721-723. <https://doi.org/10.1007/s12192-015-0617-9>

Elkouby-Naor, L. (2010). Functions of Claudin tight junction proteins and their roles in the paracellular barrier. Tissue Barriers. Advance online publication. <https://doi.org/10.1016/j.tibs.2010.07.005>

Gao, J., Liu, H., Wang, X., Wang, L., Gu, J., Wang, Y., Yang, Z., Liu, Y., Yang, J., Cai, Z., Shu, Y., & Min, L. (2023). Associative analysis of multi-omics data indicates that acetylation modification is widely involved in cigarette smoke-induced chronic obstructive pulmonary disease. Frontiers in Medicine, 9, 1030644. <https://doi.org/10.3389/fmed.2022.1030644>