# Analyzing Retrotransposons in RNA Data: Can I Hand It Over to AI or Do I Still Have to Work?

Ostašov Pavel1,2, Klieber Robin1,3, Dekojová Tereza1,3, Holubová Monika1,3,4

1 Charles University, Faculty of Medicine in Pilsem, Laboratory of Tumour Biology and Immunotherapy

2 Charles University, Faculty of Medicine in Pilsem, Department of Histology and Embryology

3 University Hospital Pilsen, Department of Hematology and Oncology

*4Consortium for iNKT Research and Therapy*

**Background**  
With the growing popularity of AI in the form of LLMs for coding, we tested whether such models could simplify complex bioinformatics tasks sufficiently for non-experts to generate not only basic analysis code but also scripts for less common tasks such as retrotransposable element (TE) analysis.

**Aims**  
We evaluated whether LLMs could set up the environment and perform TE analysis with a single prompt: link samples to FASTQ files, align reads, quantify gene and TE family expression, and compare TE expression between two groups of patients, where TE expression could be altered by a disease status and affect immune response.

**Methods**  
Cloud-based models (GPT-o3, GPT-4o, Sonnet, Gemini) and local models (Qwen3, Gemma3, DeepSeek-r1) were given a one-shot TE analysis task. The results were evaluated for workflow design, code quality, and whether a non-expert could identify and correct errors. Analyzed data were represented by RNA-seq of invariant NKT cells from two groups of myeloma patients with partial or complete remission.

**Results**  
All models produced errors, including missing dependencies, broken data links, and flawed logic. GPT-o3 even tried to implement its own method for differential expression analysis. Cloud models generated compact scripts; local models provided more stepwise instructions. RNA-seq alignment and basic differential expression were mostly correct, but TE analysis and environment setup often failed. No model created a fully functional workflow. Reasoning ability and model type did not consistently predict better performance.

**Conclusion**  
LLMs can assist in generating analysis code but cannot yet replace experts, especially in niche or emerging fields where limited training data is available. Dividing the task into smaller subtasks would improve accuracy but still risk guiding users into dead ends.

**Funding**  
Supported by the Ministry of Health of the Czech Republic in cooperation with the Czech Health Research Council under project No. NW24-03-00079.