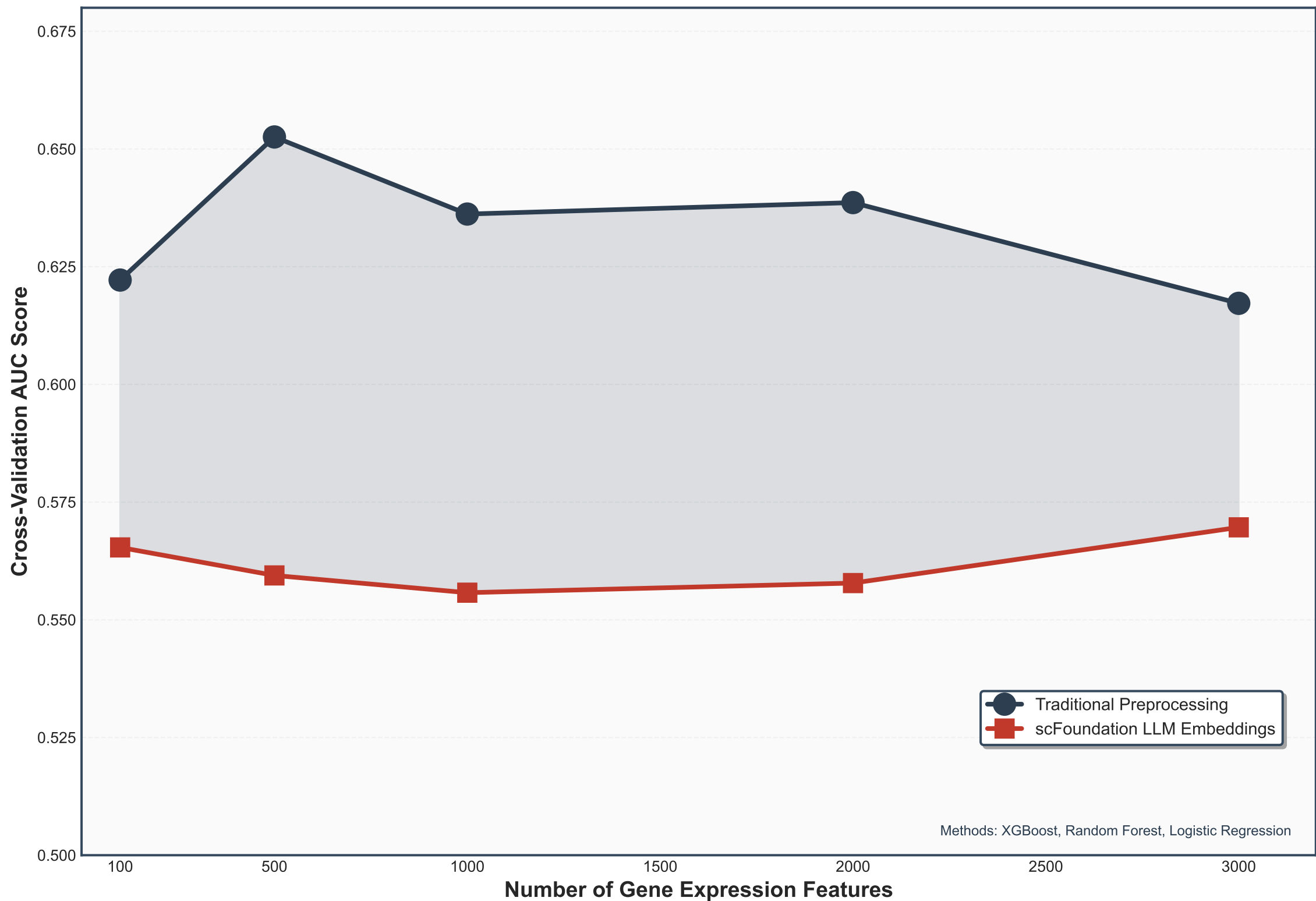


## Gene Expression Feature Engineering: Traditional Preprocessing vs scFoundation LLM



**Figure 7. Comparison of traditional gene expression preprocessing versus scFoundation large language model embeddings. Across all feature counts tested, traditional preprocessing consistently outperforms LLM embeddings, achieving up to 8.3% higher AUC. This suggests that task-specific feature engineering remains superior to general-purpose embeddings.**