1. *Sam Boshar, Evan Trop, Bernardo P de Almeida, Liviu Copoiu, Thomas Pierrot, Are genomic language models all you need? Exploring genomic language models on protein downstream tasks,*Bioinformatics*, Volume 40, Issue 9, September 2024, btae529,*[*https://doi.org/10.1093/bioinformatics/btae529*](https://doi.org/10.1093/bioinformatics/btae529)

Paper about the use of genomic large language models for protein synthesis/engineering, which is the same principle as Chat GPT the most famous Large language model. LLM work well because DNA sequences have large groups of unlabled data. Other papers might dive deeper in the actual model itself. For instance:

1. *Outeiral, C., Deane, C.M. Codon language embeddings provide strong signals for use in protein engineering. Nat Mach Intell 6, 170–179 (2024).* [*https://doi.org/10.1038/s42256-024-00791-0*](https://doi.org/10.1038/s42256-024-00791-0)

Paper about the use of such LLM on DNA, but instead of using Amino acids, this model is trained on codons, with the goal to outperform previous models on tasks like species recognition and protein prediction

1. *Machine learning reveals mesenchymal breast carcinoma cell adaptation in response to matrix stiffness* [*https://doi.org/10.1371/journal.pcbi.1009193*](https://doi.org/10.1371/journal.pcbi.1009193)

This study used a machine learning algorithm to predict which factors in the cell and tissue play a role in the development of breast cancer tumors. For instance the tissue/matrix stiffness and cell shape.

1. *Objective quantification of nerves in immunohistochemistry specimens of thyroid cancer utilising deep learning* [*https://doi.org/10.1371/journal.pcbi.1009912*](https://doi.org/10.1371/journal.pcbi.1009912)

This study aims to quantify the amount of nerves in cancer specimens. The current method is time-consuming and expensive. A deep learning model is used to do pixel-by-pixel detection.

1. *In vitro machine learning-based CAR T immunological synapse quality measurements correlate with patient clinical outcomes* [*https://doi.org/10.1371/journal.pcbi.1009883*](https://doi.org/10.1371/journal.pcbi.1009883)

This model predicts the efficiency of different CAR-modified immune cells (used in cancer therapy) in vivo. This machine learning model could help in prediciting the effectiveness of CAR-T cells in patients.