

# TransMEP: Transfer Learning for Mutation Effect Prediction

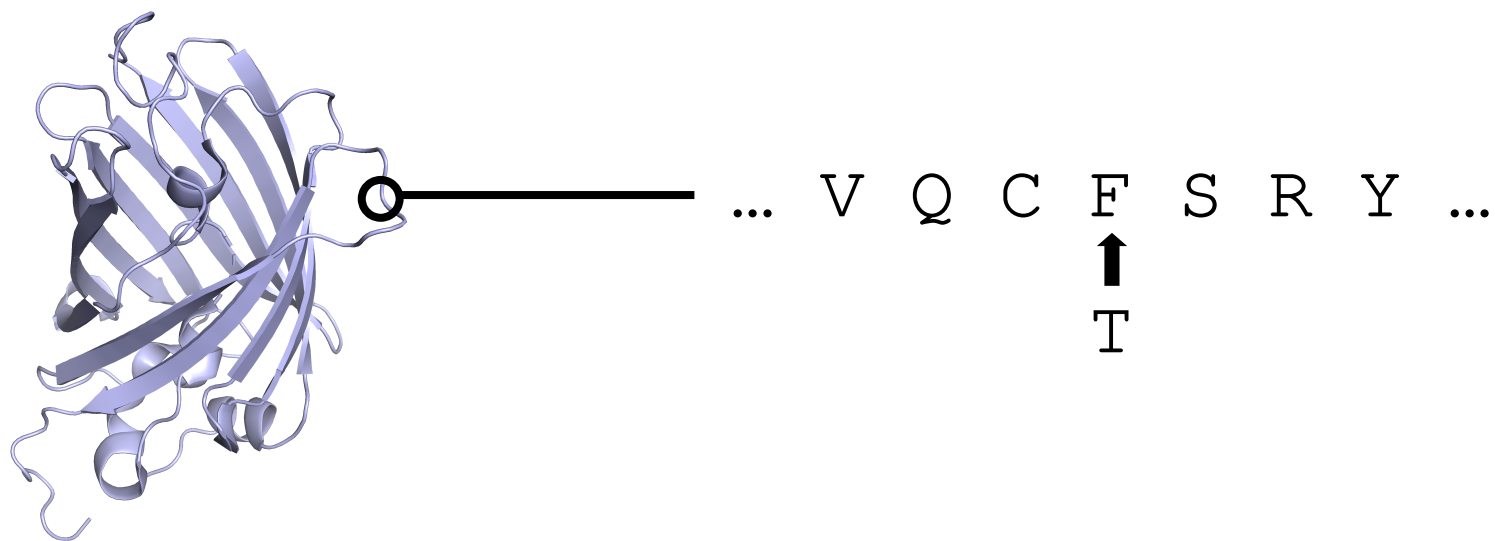
## An Easy-to-Use Pipeline

Tilman Hoffbauer<sup>1</sup>, Birgit Strodel<sup>1,2</sup>

<sup>1</sup>Computational Biochemistry Group, Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Germany

<sup>2</sup>Computational Biochemistry Group, Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, Germany

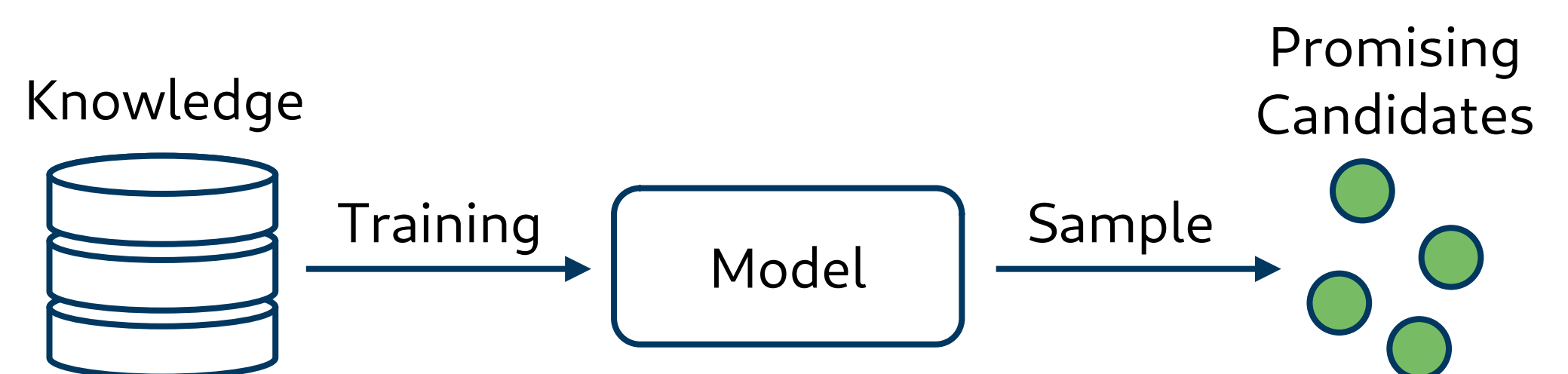
### Protein Engineering demands Mutant Assessment



- Mutations change protein properties, e.g. activity
- Mutant assessment: Quantify mutation effects
- Artificial mutations are used in protein engineering

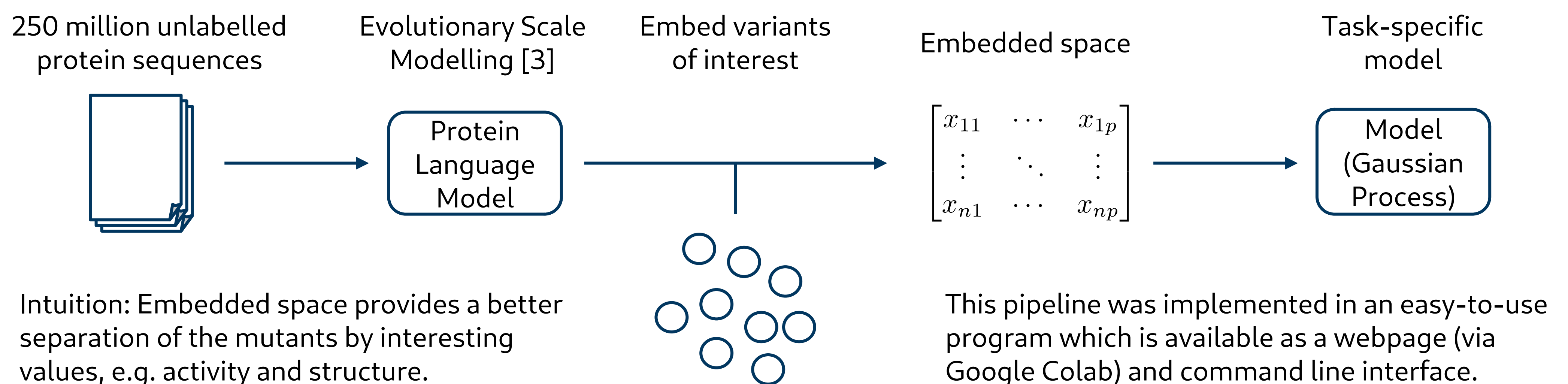
Example: GFP (UniProt ID 5N9O)

### Shortcut: Machine Learning (ML)

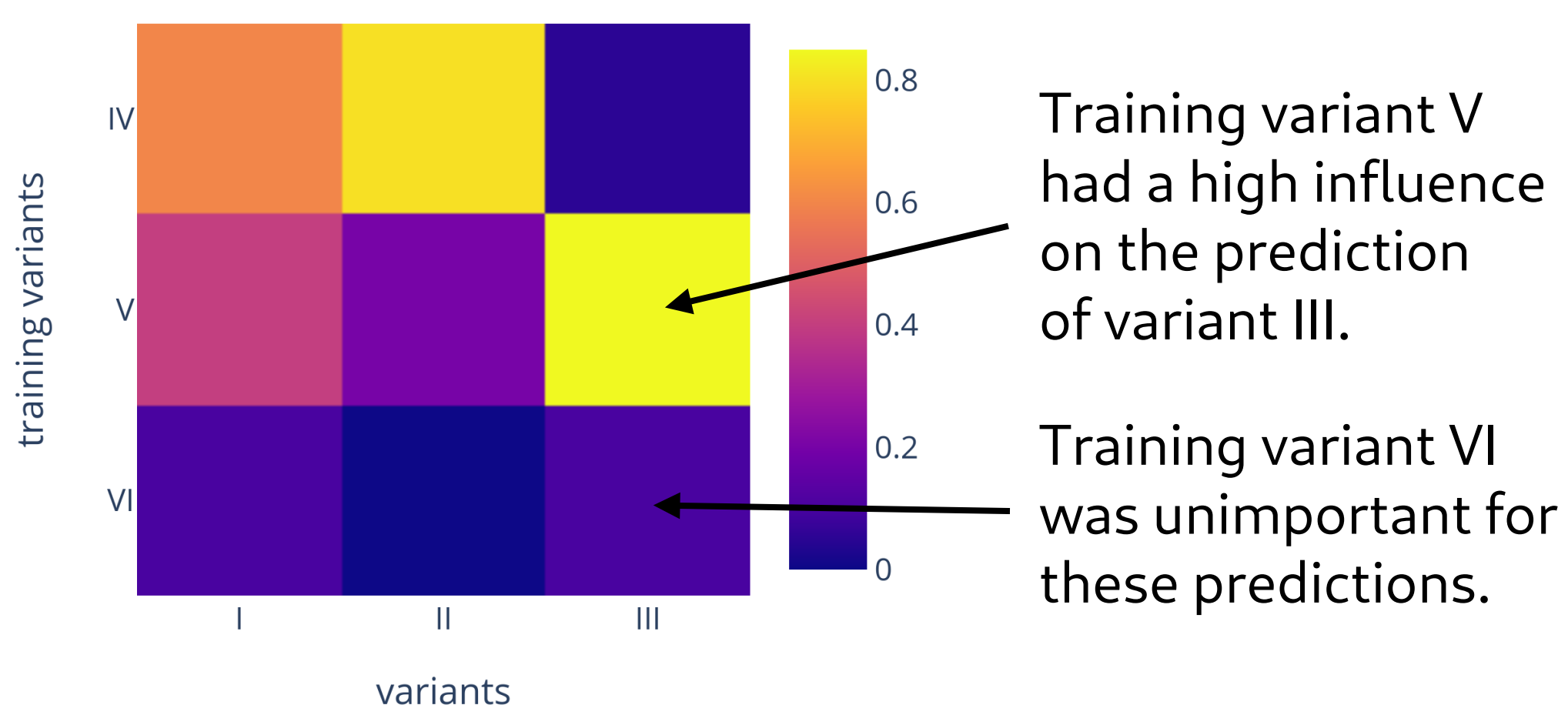


- ML is cheaper & faster than experimental assessment
- Promising candidates require experimental validation
- But: Naive ML is not accurate enough [1, 2]

### Variant Prediction using Transfer Learning Pipeline

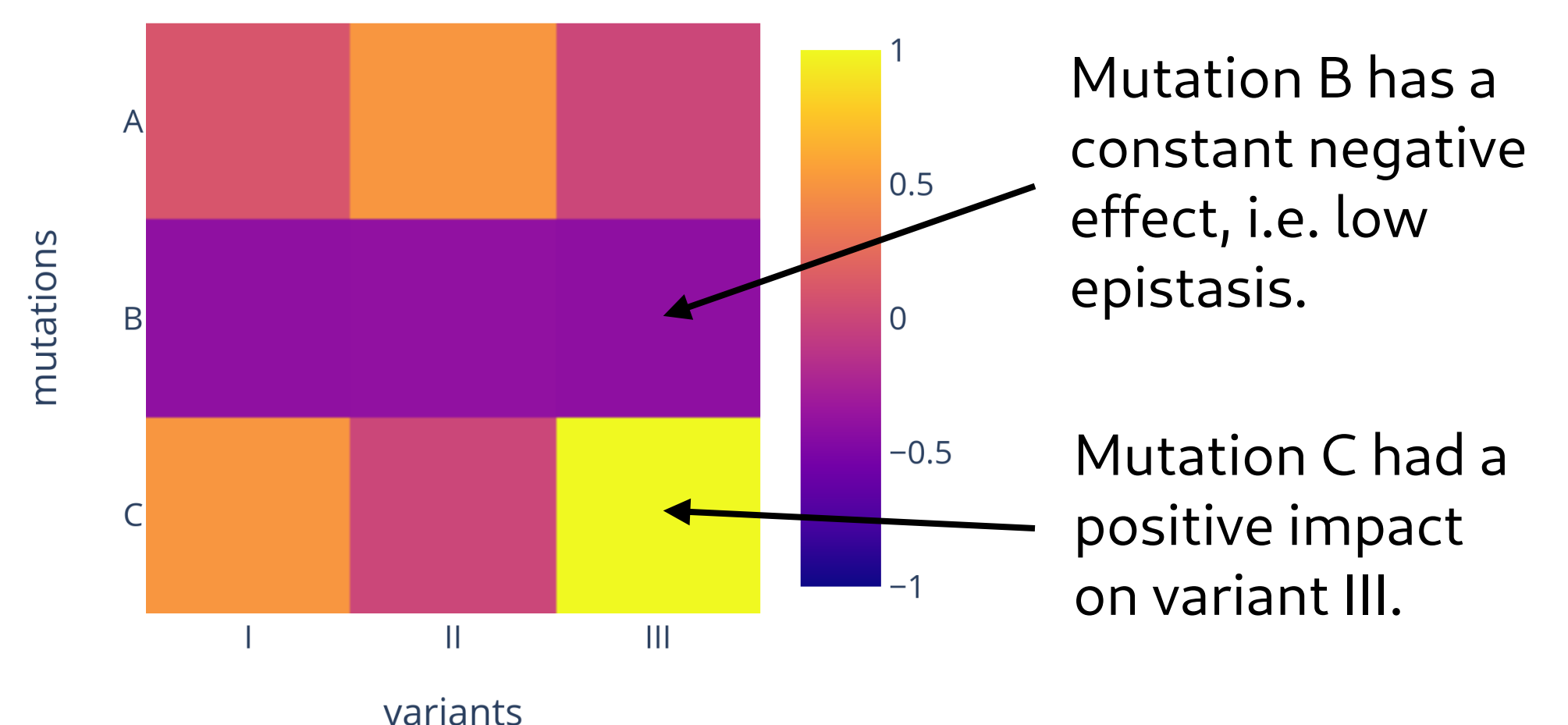


### Attributing to Training Samples



→ Can be inferred from model weights

### Attributing to Mutations



→ Computation with integrated gradients [4] via backpropagation

### TransMEP: An Easy-to-Use Pipeline

- Improved mutation effect prediction using transfer learning
- Simple usage, no installation required
- Allows attribution of predictions to individual mutations
- Detailed reports for analysis

Publication will follow soon!

E-Mail: [tilman.hoffbauer@rwth-aachen.de](mailto:tilman.hoffbauer@rwth-aachen.de)

We thank Anna Jäckering for fruitful discussions and RWTH Aachen University for HPC resources.

[1] Xu et al., JCI, 2020 [2] Hoffbauer, Bachelor Thesis at RWTH Aachen University, 2021

[3] Rives et al., PNAS, 2021 [4] Sundararajan et al., PMLR, 2017