As was mentioned in the main text of the manuscript, the target proteins are divided into four groups including nuclear receptors, GPCRs, ion channels and enzymes. All of the operations of this study were implemented separately for each group. For example, the datasets and implemented source codes for the nuclear receptors group include the following:

- **Nuclear receptor file:** The Known drug-target interactions proposed by Yamanishi et al.
- **Nucleardescriptors file:** The drug compound descriptors extracted by PaDEL-Descriptor software.
- **Nuclearfeatures file:** The target protein features extracted by protr R package.
- **NuclearFASTA folder:** The sequence of all target proteins in nuclear receptors group (in FASTA format).
- **Nuclearreceptordrugs folder:** The mol file of all drug compounds so that have known interaction with at least one target protein in the nuclear receptors group.

Supplementary folder: This folder contains a table of evaluation results and new predicted interactions in each target protein group that are referred to in the main text of the manuscript.

Nuclear code folder: This folder contains three subfolders and an R file named CLUTER. The CLUSTER file contains source code for k-medoid and Kmeans clustering. Three subfolders contain source codes related to each of the three negative samples selection methods mentioned in the main text of the article. Subfolders in this folder are including RNIDTP, SELF_BLM, and Random which contain source codes for each of the three algorithms mentioned in the manuscript. For example RNIDTP contains three source codes including:

- FS_NEG.R: The source codes for feature selection and reliable negative sample selection according to the RNIDTP algorithm (FS_NEG file)
- SVM_CV.R: The source codes for 10-fold cross-validation by using the SVM classifier.
- RF_CV.R: The source codes for 10-fold cross validation by using RF classifier.

The source codes for the other three groups of target proteins can be generalized exactly the same as the source codes presented in the nuclear receptor group according to the available dataset in every group.