AlloDriver V3

服务器

- IP: 175.185.16.17
- 用户名：hipeson
- 密码：hipeson@123

数据

- 简介：共有17130条突变数据，已平衡。其中8565为driver（阳性数据），另8565为passenger（阴性突变）。/home/hipeson/driverteam/Data/all_balance_17130.csv（数据在服务器上的位置）。
- all_balance_17130.csv每行数据格式如图：
  - 每一行相继记录了一条driver数据（；前）和一条passenger数据（；后），以；隔开
  - 每一条数据依次记录 PDB id, Chain id, PDB residue index, PDB residue (three letter), Uniprot id, Uniprot residue index, Uniprot residue (one letter), Mutation (driver or passenger)，以， 隔开
  - 所以6cms,A,73,THR,Q06124,73,T,T73I指 蛋白Q06124的三维结构PDB id 6cms在A链的73号位残基Thr可突变为Ile (I)，此突变为driver。
  - 相应的蛋白三维结构已经提前下载至/home/hipeson/driverteam/Data/PDBs。推荐使用biopython进行处理，或对缺失数据重新下载
  - 具体数据划分可参考 https://academic.oup.com/nar/article/51/W1/W129/7132339 进行随

机8:1:1划分

```
6cms,A,73,THR,Q06124,73,T,T73I;6cms,A,491,PRO,Q06124,491,P,P491F
5mcv,B,273,ARG,P04637,273,R,R273L;5mcv,A,229,CYS,P04637,229,C,C229F
2pkl,A,877,THR,P10275,878,T,T878A;2pkl,A,801,PRO,P10275,802,P,P802L
421p,A,58,THR,P01112,58,T,T58I;421p,A,33,ASP,P01112,33,D,D33Y
2j1y,C,126,TYR,P04637,126,Y,Y126N;2j1y,B,248,ARG,P04637,248,R,R248H
1ycs,A,179,HIS,P04637,179,H,H179D;1ycs,A,160,MET,P04637,160,M,M160T
7dhz,A,273,ARG,P04637,273,R,R273C;7dhz,A,174,ARG,P04637,174,R,R174H
4nmm,A,12,CYS,P01116,12,G,G12A;4nmm,A,147,LYS,P01116,147,K,K147T
5ecg,A,179,HIS,P04637,179,H,H179P;5ecg,B,133,MET,P04637,133,M,M133V
4ark,A,115,LEU,Q02750,115,L,L115P;4ark,A,189,ARG,Q02750,189,R,R189K
3rs2,A,13,GLY,P01112,13,G,G13D;3rs2,A,64,TYR,P01112,64,Y,Y64H
2bim,B,197,VAL,P04637,197,V,V197M;2bim,B,97,VAL,P04637,97,V,V97D
5yu9,A,724,GLY,P00533,724,G,G724S;5yu9,A,790,MET,P00533,790,T,T790A
4ifj,A,415,ARG,Q14145,415,R,R415H;4ifj,A,334,TYR,Q14145,334,Y,Y334H
1xkk,A,858,LEU,P00533,858,L,L858R;1xkk,A,800,ASP,P00533,800,D,D800N
3ii5,A,593,ASP,P15056,594,D,D594A;3ii5,A,462,ILE,P15056,463,I,I463S
```

模型

- 目标：**使用surface革新driver prediction蛋白表征**
- 参考 https://github.com/gcorso/DiffDock/blob/dev/models/all_atom_score_model.py 中build_cross_conv_graph中 ATOM to RECEPTOR交互建图进行RECEPTOR to SURFACE交互建图
- **也可完全参考 https://academic.oup.com/bioinformatics/article/40/7/btae413/7697100 ，与我们目的一致，只是任务不同**