

相互作用マシンラーニング法による化合物スクリーニングシステム

相互作用マシンラーニング法 — Chemical Genomics-Based Virtual Screening (CGBVS) —

■ 高速・高精度のin silicoスクリーニングを実現

機械学習を応用した独自の的方法論により、高速で精度の高い in silico 化合物スクリーニングを実現します。

■ アッセイデータの有効活用が可能

ライブラリー評価実験で蓄積されたアッセイデータ等を学習データとして活用する事で、予測モデルをカスタマイズしたり、新たに構築したりする事ができます。

■ マルチターゲット予測に対応

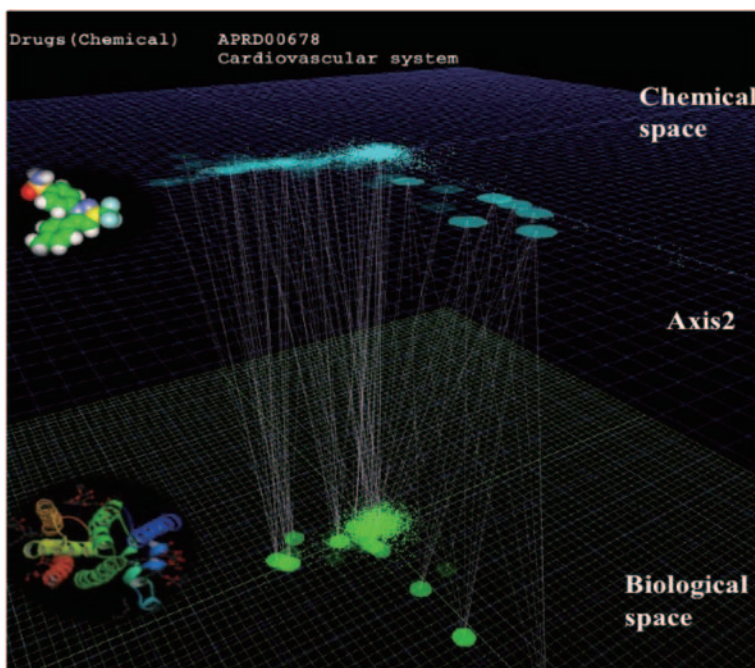
複数の標的タンパク質に対するスコアリングが可能です。標的選択性を考量した化合物スクリーニングや、オフターゲットの予測などに応用できます。

■ 豊富な利用実績

当社の受託計算サービスにおいて、GPCR やキナーゼ、イオンチャネルなど様々なターゲットを対象とした計算実績があります。

■ スーパーコンピュータ「京」による大規模計算

スーパーコンピュータ「京」の産業利用プロジェクト（平成 24 年度、課題番号：hp120048）において、Czeek S を用いて 189 億通りのタンパク質 - 化合物相互作用の大規模予測計算を実施しました。



膨大な数の既知化合物と標的タンパク質の相互作用データをもとに機械学習によるモデリングを行い、目的の化合物について相互作用を予測します。



CGBVS は京都大学の奥野恭史教授により開発された in silico スクリーニング手法（特許第 5448447）です。
株式会社京都コンステラ・テクノロジーズは京都大学より独占実施権を付与されて CGBVS の製品・サービスを提供しています。

システムライセンス

CzeekS では、CGBVS の予測モデル構築（機械学習）及び予測計算（スコアリング）の機能をコマンドラインの操作で実行することができます。複数の計算 node を使用する事により、計算処理の速度アップが可能です。

node 数	1 年間ライセンス価格 (税抜)	3 年間ライセンス価格 (税抜)
1 node	¥500,000	¥1,350,000
5 node	¥750,000	¥2,100,000
10 node	¥950,000	¥2,700,000
無制限	¥1,900,000	¥5,500,000

予測モデル

当社保有の化合物 - タンパク質相互作用データを用いて構築した下記の予測モデルを提供しています。

モデル名	標的タンパク質数	学習データ数 (相互作用数)	モデル内容 (標的タンパク質の種類)	価格 (¥: 税別)
GPCR	242	158,405	ClassA α , ClassA β , ClassA δ , ClassA γ , ClassB, ClassC	500,000
Kinase	412	154,079	AGC, CAMK, CMGC, STE, TK, TKL, others	500,000
Ion Channel	218	142,282	Voltage-gated, Ligand-gated, others	500,000
Transporter	148	112,659	Electrochemical, ATPase, ATP-binding cassette, others	400,000
Nuclear Receptor	41	139,051	NR1, NR2, NR3, NR4, NR5	400,000
Protease	229	141,691	Endopeptidase, Exopeptidase	500,000

※ 表は ver.2 の内容を記載しています。最新情報については HP をご確認ください。 <http://www.k-ct.jp/service/czeeks.html#plist>

化合物記述子計算ソフト

CGBVS の計算では化合物およびタンパク質の構造をそれぞれの記述子に変換して利用します (右図)。

CzeekS ご利用の際には、化合物の記述子化に必要な計算ソフト DRAGON (Kode srl 提供) のライセンスが別途必要となります。CzeekS ユーザー向けに DRAGON の特別ライセンスがご利用頂けますので、詳細をお問合せ下さい。



動作環境

CPU	マルチコア CPU (Intel, AMD)
メモリ	16GB
ディスク	100GB 以上の空き容量
OS	Linux (CentOS, Fedora Core)

リファレンス

- Analysis of multiple compound-protein interactions reveals novel bioactive molecules. *Mol. Syst. Biol.* 7, 472, 2011
- Systems biology and systems chemistry: new directions for drug discovery. *Chem. Biol.* 19(1), 23-8, 2012
- Unifying Bioinformatics and Chemoinformatics for Drug Design. *Systems and Computational Biology – Bioinformatics and Computational Modeling*, 99-120, 2011

CGBVS 計算フロー

