



Chemotargets CLARITYによる化合物プロファイル

Chemotargets CLARITY (以下、CLARITY) は、医薬品候補化合物の薬理的に重要な側面であるターゲット、安全性、代謝物について適切にプロファイルを予測するための計算プラットフォームを提供します。CLARITYには、T-link (Molecule-Target)、S-link (Molecule-Safety)、M-link (Molecule-Metabolism) という3種類のリンク情報が組み込まれています。CLARITYは高度にモジュール化された計算アルゴリズムを統合することで、化合物と3種類のデータ空間、つまりターゲット、安全性、代謝物の関係を同定します。ここでは、CLARITYにおける化合物のプロファイル予測手法について紹介します。

■予測手法

■ T-link: Molecule - Target

一般的に各ターゲットタンパク質に対して生物活性を持つ化合物は、分子記述子により、そのリガンドとしての特徴を表現することができます。CLARITYのMolecule-Targetモジュール(T-Link)は、以下のSAS、SAR、SIM、SEA、MLM、XPIの6種類の独立した計算手法を統合してターゲットの予測を行います。

1. Simplest Active Subgraph similarity (SAS)

SASは、活性を持つための必要最小限のファーマコフォアグラフとして定義されます。SASにより、従来は同様の活性があるが類似化合物とみなされなかったものでも、適切に類似性を評価することができます。

2. SAS-based QSAR model (SAR)

SASベースのQSARモデルは、高速かつ簡単に3D構造に依存しない活性予測手法を提供します。

3. Molecular Similarity (SIM)

それぞれ独立した化合物の指標であるPHRAG (Pharmacophoric fragments)、FPD (Feature-pair distribution)、SHED (Shannon entropy descriptors)の3種類の2D記述子を用いて類似性を評価します。

4. Similarity Ensemble Approach (SEA)

Keizerらにより開発されたSimilarity Ensemble Approachは、元々リガンド間の類似性によりタンパク質の関連性を同定する方法ですが、化合物に応用してリガンド-ターゲット相互作用の予測を行います。

5. Machine Learning Methods (MLM)

FPD分子記述子をベースに作られた1000を超える高品質なニューラルネットワーク、サポートベクターマシン、ランダムフォレストの各モデルのコンセンサス・スコアです。

6. Cross-pharmacology indices (XPI)

2種類のターゲット間における類似リガンドの数を元にして評価します。パブリックドメインにおける数千の低分子薬理データを元に頑健な予測を行います。

■ S-link: Molecule - Safety

CLARITYの安全性モジュール(S-Link)は、各安全性項目に関連した医薬品の豊富なデータを統計解析することにより、有害事象を同定します。異なるタイプの有害事象を評価するために4種類の異なる計算手法が適用されます。

1. Ligand Based (LB)

各エンドポイントにアノテーション付けされた化合物を参照して類似性を評価します。

2. Fragment Based (FB)

事前に定義されたフラグメントと生成したRECAPフラグメントを加えたデータセットを参照して、安全性に関わる分子フラグメントを評価します。

3. Target Based (TB)

既知のターゲットタンパク質またはエンドポイントにアノテーション付けされたタンパク質に対する活性予測に基づいて安全性を評価します。

4. Pathway Based (PB)

パスウェイから安全性を評価します。化合物のターゲットプロファイルは、薬理的プロファイルに含まれるあらゆるターゲットの全てのパスウェイをアノテーション付けることにより、パスウェイ・プロファイルに翻訳されます。

■ M-link: Molecule - Metabolism

フラグメントベースのQSARモデルにより化合物に対して全ての代謝物が予測されます。M-link QSARモデルは、フェーズIとIIに含まれる生体異物について実験に基づく生体内変換のデータセットをトレーニング・セットとして構築されています。

■ユーザー・インターフェース

画面に入力化合物、代謝物とその安全性情報、ターゲットタンパク質と安全性のエンドポイントが見やすく配置されます(図1)。また、それぞれの予測や検索結果の根拠となったデータソースへアクセスするリンクが作られており、即座に関連情報を参照することができます。



図1. CLARITYの予測結果の表示例