

単純グラフ同型な新規薬物骨格デザインのための分子進化プログラミング Molecular evolutionary programming for structure design based on simple graph isomorphic scaffold

丸野 裕史
Hiroshi Maruno

高橋 由雅
Yoshimasa Takahashi

豊橋技術科学大学 工学部 情報・知能工学系
Department of Computer Science and Engineering, Toyohashi University of Technology

In this paper, we propose an approach to finding new molecular skeletons which are expected to have a desirable property. The method is based on evolutionary programming. In this approach a seed structure (an NTG in simple graph expression) is submitted to generate an initial population. Individuals of the initial population are generated along with evolutionary operations, including mutation and recombination. Each chemical structure of the individuals is coded with an atomic connection table. Computational trial was carried out with a seed structure and a fitness function on structural similarity to examine whether the method can make it evolve into a target structure. The result suggested that the present method allows us to generate possible candidate structures of our interest. Basic idea of the method and its implementation are presented with illustrative examples.

1. 背景と目的

新薬開発における候補構造の探索は、 10^{60} 個以上の解候補が存在する膨大な空間を対象としており、開発コスト縮小のため様々な探索手法が検討されている。先行研究において吉丸らは、薬物構造の側鎖に着目し、進化的プログラミングを用いた新規薬物構造創出へ向けた可能性を示した[吉丸 08]。

一方、図 1 に示すように、NTG(NTG:Non-Terminal vertex Graph) [大野 02]の形状の似た薬物は、発現する活性も似ることが知られている。そこで本研究では、薬物の基本的骨格構造に着目し、この NTG の単純骨格を保持したうえで進化的プログラミングを適用し、新薬開発の基盤となりうる NTG をデザインする方法について検討を行う。最終的には任意の活性に着目し、注目する活性発現が期待される骨格構造の提案を目指す。

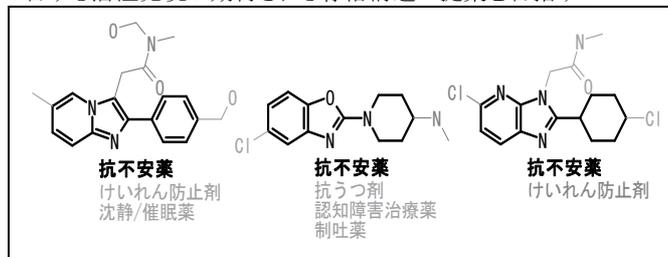


図 1 NTG の形状と活性の関係

2. 方法

進化的プログラミング (EP:Evolutionary Programming) は、自然界における生物の進化を模した探索手法の一つである。生物が突然変異・交叉・適者生存を繰り返すことで周囲の環境により適した形態へと進化してきた過程に倣い、例えば新規薬物構造のように明確な解が提示されない問題に対して、条件に適合したより望ましい解を得るために用いられる。本研究では、進化対象となる個体を NTG とし、その結合行列を染色体として用いた。単純骨格グラフからの NTG 分子進化プログラミングの概要を図 2 に示す。

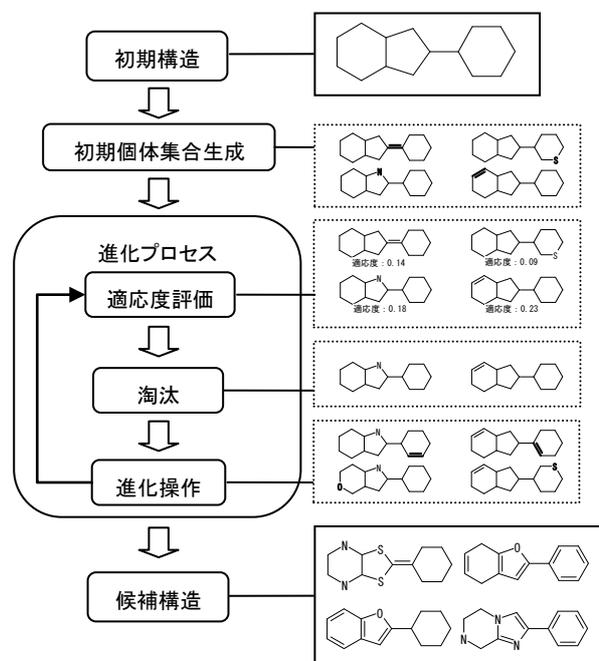


図 2 NTG 分子進化プログラミングの概要

2.1 初期構造の選択

特定の活性が報告されている市販の薬物構造の基本的骨格を初期構造として用いることで、その活性が発現する可能性の高い新規薬物骨格をデザインできることが期待される。また、市販の薬物には用いられていない NTG を初期構造とすることで、未発掘の候補構造空間の探索を実現することができる。

2.2 初期個体集合の生成

本研究では、初期構造として炭素と単結合で構成された基本骨格を与え、この構造の頂点(原子種)や辺(結合種)に対し進化操作を行っていく。進化プロセスへ入力する初期個体集合は、初期構造の一部をランダムに変異させることによって生成した。

2.3 適応度評価

適応度評価関数を用い、進化操作により生成された各個体について適応度を算出する。本研究では単純骨格グラフからの目標構造の進化による再現実験に際し、適応度評価関数には構造類似度を用いた。類似度評価には薬物構造の特徴を数値化することが必要であり、本研究では Topological Fragment Spectra(TFS)法を用いた[Takahashi 98]。

2.4 淘汰

適応度評価で算出した値をもとに、進化操作により生成された 100 個体から、トーナメント法を用いて 20 個体を次世代へ生き残る個体として選出する。

2.5 進化操作

淘汰で生き残った個体に対し、あらかじめ設定した確率に応じて交叉および変異いずれかの操作を施し、次世代の個体集合(100 個体)を生成する。

進化操作における変異は、構造の原子種または結合種を変化させることで実現する。なお、現在原子種変化において扱う原子は、NTG 骨格内に一般的にみられる炭素、窒素、酸素、硫黄の 4 種のみを対象としている。交叉は、対象となる 2 つの構造をランダムに選択し、任意の 2 点で組み替えることで実現している。

3. 計算機実験

3.1 実験概要

実装した EP アルゴリズムの有効性を検証するため、特定の薬物の骨格構造を目標構造とし、その構造との類似度を適応度に設定したうえで、単純グラフ骨格から目標構造の再現を試みた。100 個体/世代で 100 世代目までの進化実験を 100 セット行い、各世代の適応度の平均値、最良値、そして最悪値の 100 試行平均を求めた。

3.2 目標構造と適応度評価

本実験で目標と設定した構造は図 3 の通りである。この構造は、ある程度頂点数が多いという基準のもと無作為に選出した。これは、頂点数の少ない構造だと数世代で容易に目標構造へ辿りついてしまい、世代毎の適応度の推移を観測するには不向きなためである。

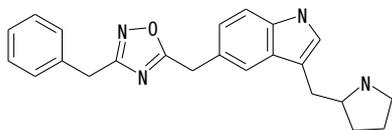


図 3 目標構造

各個体の適応度を算出するために必要な評価関数には、被評価側構造の TFS ベクトルと、目標構造の TFS ベクトルとの類似度を用いた。この計算のため、オブジェクト間の類似度の尺度として広く使われている、Tanimoto 係数を用いた。算出のための式は次の通り。

$$S(x, y) = \frac{\sum x_i y_i}{\sum x_i^2 + \sum y_i^2 - \sum x_i y_i} \quad (0 \leq S \leq 1)$$

ここで、x は被評価側構造の TFS パターンベクトルを、y は目標構造の TFS パターンベクトルを表す。

3.3 実験結果

実験結果として、最終的に生き残った候補構造の一例および、世代毎の適応度の推移を図 4 に示す。100 世代後に生き残った構造の中で、目標構造と同一なものを 100 試行中 2 回再現することが出来た。初期構造の変化可能部位が頂点 28, 辺 32 箇所であることを考えると、この結果は良好なものと言える。

目標構造を復元できなかった試行において、最終的に生き残った構造の適応度は、最高 0.987(図 4(a)), 最低 0.928(図 4(b))であり、平均は 0.952 となった。この結果から、適応度(目標構造との類似度)を高める方向で進化を進めていくという狙い通りに、プログラムが動作することを確認できた。

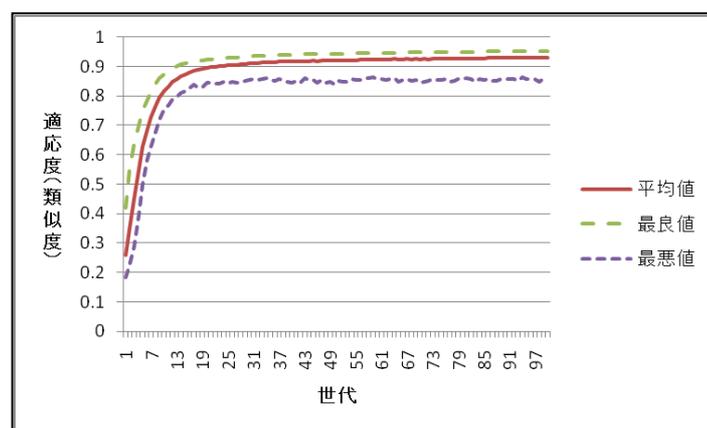
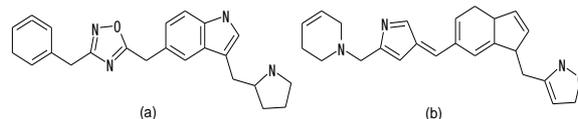


図4 実験結果

4. まとめ

本論文では、進化的プログラミングを新薬開発支援に応用する方法を提案し、計算機実験によりその方法を実装したシステムの妥当性を示した。今後の予定として、まず化学的な知識をプログラムに組込むことで、例えば図 4(a) のような合成が現実的ではない構造の除去を含めた進化の効率化を目指す。また、現在対象としているのは側鎖のない構造のみだが、その有無が活性発現に大きな関わりを持つことが知られるカルボニル基等、特定の置換基を考慮した構造も取り扱えるよう、検討していく予定である。また、上記内容を拡張した後、薬物構造からの活性予測モデルを用い、望んだ活性を持つ可能性の高い薬物骨格のデザインへ向けた計算機実験についても合わせて進めていきたい。

参考文献

- [吉丸 2008] 吉丸和孝, 高橋由雅: 進化的プログラミングを用いた創薬候補化学構造の創出, 人工知能学会全国大会論文集, 人工知能学会, 2008.
- [大野 2002] 大野貴生, 高橋由雅: NonTerminal Vertex Graph(NTG)を利用した薬物の構造特徴解析, 第 30 回構造活性相関シンポジウム講演要旨集, 日本薬学会構造活性相関部会, 2002.
- [Takahashi 1998] Y. Takahashi, H. Ohoka, and Y. Ishiyama: Structural Similarity Analysis Based on Topological Fragment Spectra, JAI Press, Greenwich, 1998.