

生命科学知識の形式的記述

生命現象のメカニズムを理解するためにシグナル伝達パスウェイデータベースが知識基盤として期待されている。ところが、シグナル伝達に関する知見をデータベース化しようとする、どのような知識をどのような形で計算機上に表現するかで頭を悩ませてしまう。というのも"シグナル"が生化学的なレベルの相互作用のみならずプロセスとプロセスの間の関係まで扱っているからである。そしてこのような知識は通常論文中に自然言語や図などの計算機にはなじまない形で表現されている。

シグナル伝達は細胞が外部から受け取った刺激を核に伝達して応答を返す仕組みであるが、さまざまな異質な概念が同列で論じられるため、代謝パスウェイにおける酵素触媒反応のように、標準の記述単位を決めて知識を表現することができない。すべてのステップを書き下すと必要な情報が不足するし、逆に抽象的に記述した知識は背景知識を前提としている場合があって、記述単位を抽象的な表現に統一することもできないためである。

我々はこの問題に対して、階層的で再帰的な表現形式とオントロジーによる意味づけという二本柱で取り組んでいる。我々のアプローチでは、パスウェイ上の各要素をグラフの節点と対応づけ、さらにその内部に下位構

造をもたせることにより、任意の記述単位で生体内プロセスを表現する仕組みになっている(図1)。そして、オントロジーで定義されたさまざまな抽象度の概念によって、階層的な表現の各要素に意味づけがなされている。また、部分構造(パスウェイモチーフ)を明示的にアノテートできることにも注意を払っている。例えば、一つのプロセスを実現する一連の要素を下位構造として括ることによって、サブパスウェイに対する機能のアノテーションを行っている(図2)。

一見バラバラに見えるパスウェイ構造にも、実は生物種を超えて似たような構造がさまざまに再出現することが知られており、これらは進化の過程で生物が獲得していった機能を実現するための部品にあたると思われる。現在、XML データベース上でシグナル伝達データベースの開発に取り組んでおり、パスウェイの検索を実現するだけでなく、このような部分構造に対する検索の提供も試みている。将来的には文献で報告されたパスウェイ構造を比較することで、部分構造をパスウェイモチーフとして整理体系化できるのではないかと期待している。

(本研究は科学技術振興事業団バイオインフォマティクス推進センターから支援をうけて実施している。)

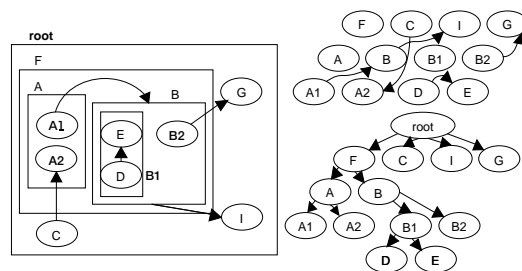


図1 パスウェイデータの階層表現
相互作用関係を定義したグラフと階層関係を定義した木によって定義される。

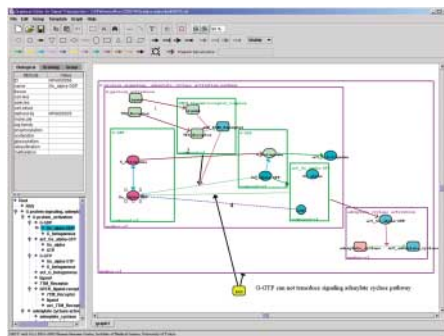


図2 パスウェイエディットツール GEST
われわれの開発したツールでパスウェイを階層的に入力している状態。



ふくだ けんいちろう
福田賢一郎
fukuda-cbrc@aist.go.jp
生命情報科学研究センター

関連情報

- K.Fukuda and T.Takagi : Bioinformatics, Vol.17, Issue 9, 829-837 (2001).
- K.Fukuda and T.Takagi : METMBS'2001, 297-303 (2001).