



諏訪牧子

(ゲノム情報科学チーム、
チーム長)

「私が最も気に入っている研究」という題で、いつもつこく話している内容があります。それは私の研究ではないのですが、個々の銀河を宇宙空間にプロットしたものです。宇宙でたくさんの泡が噴出しその表面に無数の銀河“粒”が張り付いた構造が見えるのです。銀河とは何？という問いに対し、個別に観察するのは決して見えない法則性が総体として見て初めて判ったのです。

これは、ゲノム情報科学チームの研究に非常に近いものです。私たちは、全ゲノムレベルでの解析から、遺伝子やタンパク質の機能発現メカニズムを理解して機能予測につなげることを目指していますが、特定のタンパク質ファミリーについて完備したセットを解析することで、機能発現メカニズムの法則性が解明できると考えています。私たちが現在精力的に集めているのが、7本の膜貫通ヘリックスを有する受容体（7TMR）です。極めて多様なリガンドと結合して、シグナル伝達を行なうメカニズムに興味深いうえ、現在ある薬の半数近くが7TMRをターゲットとしていることもあり、網羅的に集める意義は非常に大きいのです。

先日、ヒトゲノム配列から7TMRを網羅的に発見、収納したデータベース“SEVENS”（図1、<http://sevens.cbrc.jp>）を公開しました。ここから各遺伝子候補について、相同な配列、ゲノム上での遺伝子構造、モチーフ、ヘリックス領域等が視覚的に観察できます。

ヒトゲノムからの遺伝子発見には、隠れマルコフモデル（HMM）を基とする遺伝子発見プログラム（GeneDecoder、浅井氏（数理モデル・知識表現チーム）作）に7TMR特有のコーディングポテンシャルを織り込んで行いました。

SEVENS: 7本膜貫通ヘリックス型受容体の 網羅的コレクション

さらに、7TMR候補を抽出する際は、既知の配列に対する相同性検索だけではなく、モチーフ、ドメイン帰属、膜貫通ヘリックス予測を行い、遠縁のホモログ発見までを視野に入れています。このときに既知のデータで評価した各ツールの閾値を組み合わせ、多少感度を犠牲にしても選択性を100%にした場合から、逆に感度を100%にした場合までの間、様々な信頼度のデータセットを提供した点が特徴的で、少なくとも827個から2,014個の候補が見つかりました。11番染色体にシングルエクソンの嗅覚受容体が非常に多く、偽遺伝子も含めて所狭しと並んでいたり（図1）、ケモカイン受容体が3番染色体に多い、あるいは最大長の7TMRは2000残基以上など、次々に興味あることが見えてきました。

私は、今これら“銀河集団”を観察している最中であり、7TMRとは何？という問題に取り組んでいます。



図1 SEVENS データベース