

# バイオフィォマティクスアプローチによる ヒト膜蛋白質の解析

電気情報工学科 井上泰仁

## 1. はじめに

これまでに、ヒトをはじめとし、マウス、チンパンジー、カタユウレイホヤなどの多くの生物種の遺伝情報が解読されている。それに伴い、膨大な量の塩基配列、および、蛋白質の情報が産出されている。しかし、大部分の蛋白質の機能、および、立体構造については、詳しく調べられていない。

本研究課題では、膜蛋白質の中でも、生体膜内外のシグナル伝達の役割を担っている G 蛋白質共役型受容体 (G-protein-coupled receptor, GPCR) に着目をした。GPCR は真核生物種の中で、巨大なスーパーファミリーを形成しており、生体膜を貫通する 7 個の領域 (膜貫通セグメント領域) と生体膜外で膜貫通セグメント領域をつなぐループ領域から構成され、生体膜の外側にリガンド分子結合部位、生体膜の内側に G タンパク質が結合する部位を有して、細胞外の化学的情報を細胞内に伝達する機能を有する膜蛋白質である (図1)。多くの GPCR は、酵素蛋白質に並んで創薬の候補になっているために、医学、薬学の分野でも非常に重要である。様々な種類の GPCR があり、GPCR に結合するリガンド分子、および、G 蛋白質の種類などの生理学特性、および、配列類似性に基づいて、Class, Subclass, Sub-subclass, Type の階層構造の機能グループに分かれている。しかしながら、まだ、リガンド分子が特定されていない GPCR もある。

ヒトの 5 番染色体長腕 5q14 にある *VLGR1* (Very large G-protein-coupled receptor 1) 遺伝子は、熱性けいれんの責任遺伝子として報告されている<sup>1)</sup>。この遺伝子は、生体膜外からのリガンド分子が結合する N 末端ループ上において、Serine から特定されていないアミノ酸にかわる置換 (S256X) が報告されている。そのため、リガンド分子が膜蛋白質に結合できないため、膜蛋白質としての機能を低下させることにつながっている。また、膜蛋白質が関与していると考えられている疾患も多いが、原因となる蛋白質が特定されていない。そのため、バイオフィォマティクスアプローチによる GPCR の機能、および、立体構造に関する知見を収集、および、解析し、G 蛋白質シグナル伝達系、および、病態の解明を目指す。まず、GPCR の機能、および、構造に重要なループ領域の検出するために、データマイニングアプローチの一つである二分木法を用いて解析を行った<sup>2)</sup>。

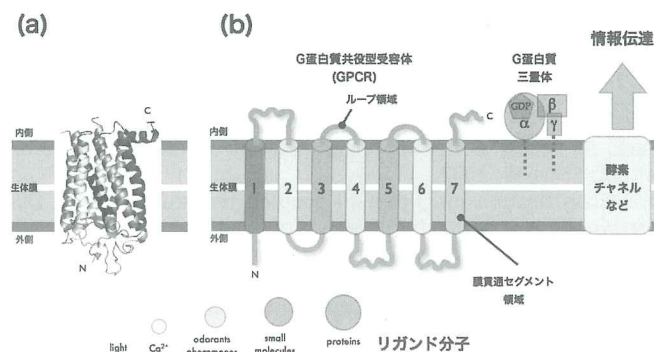


図1 GPCR の立体構造と模式図。(a) 光感受受容体の一つであるロドプシンの立体構造。(b) 模式図。生体膜を貫通している7個の膜貫通セグメント領域とそれらをつないでいるループから構成される。GPCRは生体膜外からのリガンド分子を結合し、膜貫通セグメント領域の構造を変化させ、G蛋白質と結合し、酵素などのセカンドメッセンジャー蛋白質と結合をし、脳へ情報を伝達している。

## 2. データセットと解析方法

### 2. 1 データセット

蛋白質データベース Uni-Prot (Release 56.0)<sup>3)</sup> から, GPCR 統合データベース GPCRDB<sup>4)</sup> に基づいて, 脊椎動物の GPCR を入手した. シグナルペプチド領域を SignalP 2.0<sup>5)</sup>, 膜貫通セグメント領域を HMMTOP 2.0<sup>6)</sup> を用いて予測した. 最終的に, 7 本型膜蛋白質として予測された 1,573 配列の GPCR を利用した (収集した GPCR の詳細な内訳は, Class A, Class B, Class C, Ocular albinis proteins, Frizzled/Smoothed, Vomeronasal receptors, Taste receptors の各機能グループの GPCR は, それぞれ, 1,573, 72, 30, 1, 34, 6, 120配列).

### 2. 2 解析方法

GPCR のループ領域の長さ, および, アミノ酸組成の情報をもとに, 二分木の構築を行う. ジニ係数は, 以下の式で与えられる.

$$GI(t) = 1 - \sum_i [p(i|t)]^2 \quad (1)$$

機能グループ  $i$  でのノード  $t$  における割合  $p(i|t)$  である. ジニ係数が最大になる GPCR のループ領域の長さ, および, アミノ酸組成の最適な値を決定し, 二分木を構築した.

## 3. 結果と議論

脊椎動物の GPCR のループ長, および, アミノ酸組成を二分木法に基づいて解析を行った結果, 結果より, N 末端ループ, 第 2, 3 生体膜外ループ, 第 3 生体膜内ループの長さ, Cysteine, Glutamine, Phenylalanine, Leucine, Tryptohpan のアミノ酸組成に特徴があり, GPCR の各機能グループを特徴づけることができた<sup>2)</sup>. また, ループ領域の長さ, および, アミノ酸組成は, GPCR の機能, および, 立体構造と関連していることが示唆される. 今後は, 膜貫通セグメント領域, および, ループ領域のそれぞれの領域でのアミノ酸組成の特徴を調べ, GPCR の機能, 立体構造, 疾患との関係を解明していきたい.

本研究課題をわかりやすく紹介するためのホームページも構築し, 公開した. また, Learning Management System の一つとして国内外の高等教育機関にて導入されている Moodle (Modular Object-Oriented Dynamic Learning Environment)<sup>7)</sup> を用いて, 生命科学, および, コンピュータによる解析方法などを含めたバイオインフォマティクス自習用ホームページの構築も行っていきたい.

### 参考文献

- 1) Nakayama, J., Fu, Y.-H., Clark, A. M., *et al.*: A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. *Ann. Neurol.* 52: 654-657, 2002.
- 2) Inoue, Y., and Nakagawa, S.: Classification of G-protein-coupled receptors using a Decision tree, *The 2008 Annual Conference of Japanese Society for Bioinformatics, Abstracts*, 2008. (Accepted)
- 3) Bairoch, A., Boeckmann, B., Ferro, S., *et al.*: Swiss-Prot: Juggling between evolution and stability, *Brief. Bioinform.*, 5:39-55, 2004.
- 4) Horn, F., Bettler, E., Oliveira, L., *et al.*, GPCRDB information system for G protein-coupled receptors. *Nucleic Acids Res.*, 31:294-297, 2003.
- 5) Bendtsen, J.D., Nielsen, H., von Heijne, G., *et al.*: Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.*, 340(4):783-795, 2004.
- 6) Tusnady, G.E., and Simon, I.: The HMMTOP transmembrane topology prediction server, *Bioinformatics*, 17(9):849-850, 2002.
- 7) Moodle – A Free, Open Source Course Management System for Online Learning: <http://moodle.org/>.