

生体情報工学演習 第2-5回

北海道大学情報エレクトロニクス学科

生体情報コース3年

学生番号: 02180144 茂木貴紀

作成日: 2020/12/19

光受容体「オプシン」に関する解析

これまでの学習を生かし, オプシンについて解析を行う. まず, アミノ酸配列をドメイン検索することで, 保存されている機能的な立体構造について観察する. 次に, このアミノ酸配列に近い, 脊椎動物のアミノ酸配列を組み合わせ分子系統樹を作成し, 色覚の進化について考察する. 最後に, この遺伝子がヒトゲノムのどの位置にあるかを確認, 周囲からどのような影響を与えられているか考察する.

オプシンと視覚の関係性については, 既に資料に掲載されていたため, 省略する.

アミノ酸配列のドメイン検索

まず, NCBIのサイトにアクセスし, ホモ・サピエンスの長波長感受性オプシンlong-wave-sensitive opsin 1 アミノ酸配列を取得した. 次にその概要を示す.

link: https://www.ncbi.nlm.nih.gov/protein/NP_064445.2

(https://www.ncbi.nlm.nih.gov/protein/NP_064445.2).

long-wave-sensitive opsin 1 [Homo sapiens]

NCBI Reference Sequence: NP_064445.2

[Identical Proteins](#) [FASTA](#) [Graphics](#)

Go to:

LOCUS NP_064445 364 aa linear PRI 11-DEC-2020
DEFINITION long-wave-sensitive opsin 1 [Homo sapiens].
ACCESSION NP_064445
VERSION NP_064445.2
DBSOURCE REFSEQ: accession [NM_020061.6](#)
KEYWORDS RefSeq; MANE Select.
SOURCE Homo sapiens (human)
ORGANISM [Homo sapiens](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.
REFERENCE 1 (residues 1 to 364)
AUTHORS Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall R, Cafarelli T, Campos-Laborie FJ, Charleatoux B, Choi D, Cote AG, Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF, Kishore N, Knapp JJ, Kovacs IA, Lemmens I, Mee MW, Mellor JC, Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, Yadav A, ...

MAQQWSLQRLAGRHPQDSYEDSTQSSIFTYTNSTNSTRGPFEGPNYHIAPRWVYHLTSVWMIFVVTASVFT
NGLVLAATMKFKKLRHPLNWILVNLAVDLAETVIASTISIVNQVSGYFVLGHPMCVLEGYTVSLCGITG
LWSLAIISWERWMVCKPFGNVRFDKLAIVGIAFSWIWAAVWTAPPIFGWSRYWPHGLKTSCGPDVFSG
SSYPGVQSYMIVLMVTCIIPLAIIMLCYLQVWLAIKRAVAKQKQKESESTQKAEKEVTRMVVVMIFAYCVC
WGPYTFACFAAANPGYAFHPLMAALPAYFAKSATIYNPVIYVFMNRQFRNCILQLFGKKVDDGSELSSA
SKTEVSSVSSVSPA

次に, <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>
(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) にアクセスし, 保存ドメインを調べた.

結果としては, 以下のような画面が表示され, ドメインは7回膜貫通型Gタンパク質共役受容体の1つであることが判明した. これはも最も大きいタンパク質ファミリーの1つである.

Conserved domains on [lcl|seqsig_MAQQW_8d165b14ba29c3f96bb7c95f54e4a2d3]

View

Concise Results

Local query sequence

Protein Classification

7tmA_LWS_opsin domain-containing protein (domain architecture ID 11606785)
7tmA_LWS_opsin domain-containing protein

Graphical summary

☐ Zoom to residue level show extra options »

Query seq.

150200250300350364

TM helix 1

TM helix 2

TM helix 3

TM helix 4

TM helix 5

TM helix 6

TM helix 7

putative ligand binding site

Specific hits

7tmA_LWS_opsin

Superfamilies

7tm_GPCRs superfamily

Search for similar domain architectures

Refine search

List of domain hits

| Name | Accession | Description | Interval | E-value |
|----------------|-----------|---|----------|---------|
| 7tmA_LWS_opsin | cd15081 | long wave-sensitive opsins, member of the class A family of seven-transmembrane G ... | 42-333 | 0e+00 |

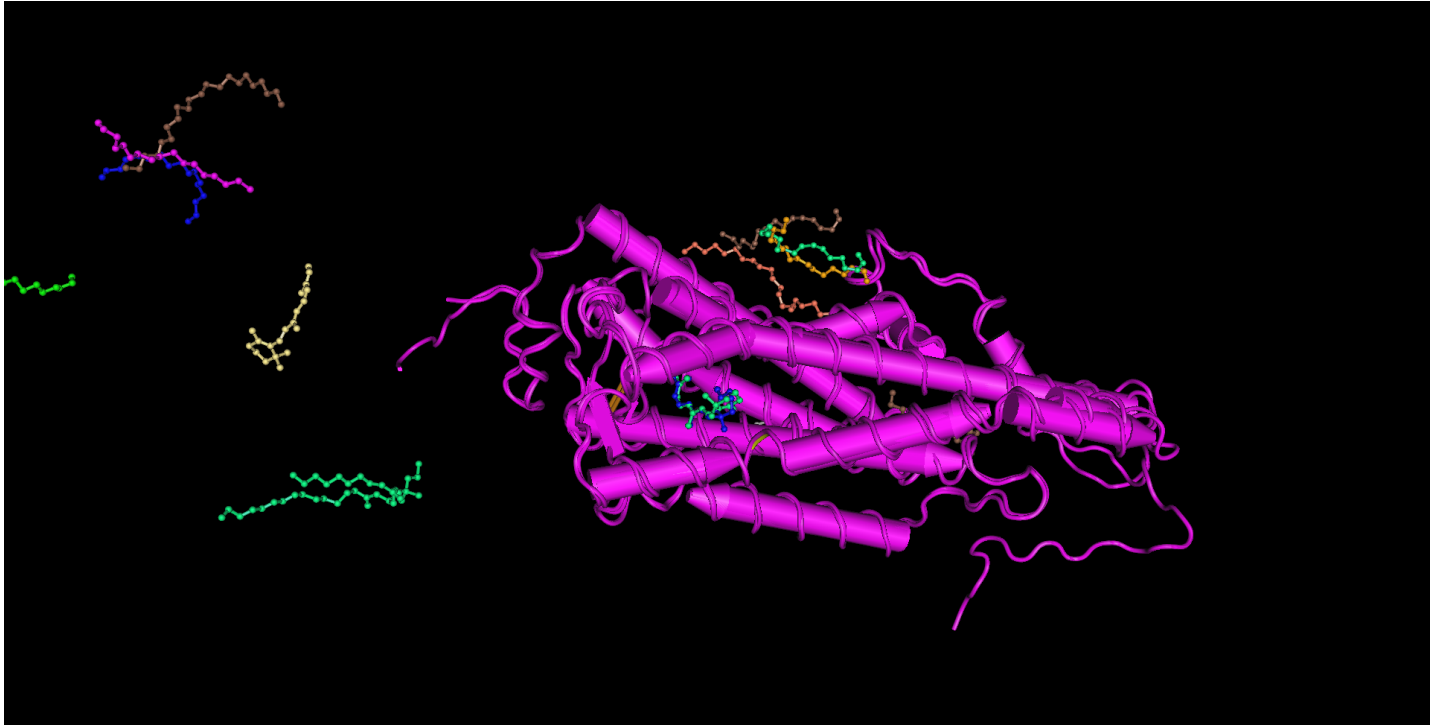
References:

https://hackmd.io/b73n8372QgSjFO65v3ahwg?view

2/8

また, <https://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>

(<https://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>) から, Cn3Dと呼ばれるタンパク質の立体構造を確認できるソフトウェアをインストールし, 立体構造について確認した. 以下に得られた構造を示す.



立体構造を観察すると, ファミリーの中心部にオプシンとは別の物質と考えられる構造が確認できる. この物質はレチノールと呼ばれる物質である. 正確には, オプシンではなくレチノールが光を受容して, これがシス型からトランス型へと変化することで, オプシンの構造が変化し, Gタンパク質を活性化する. 活性化したGタンパク質によってシグナル伝達が起こり, 光が脳で検知される, という仕組みになっている. また, オプシンとレチノールの複合体をロドプシンと呼ぶことも多い.

この立体構造ではレチノールが膜タンパク質内に閉じ込められているが, 光の受容に伴って構造が変化すると, 化学結合が外れたり, 外に飛び出ることも考えられる.

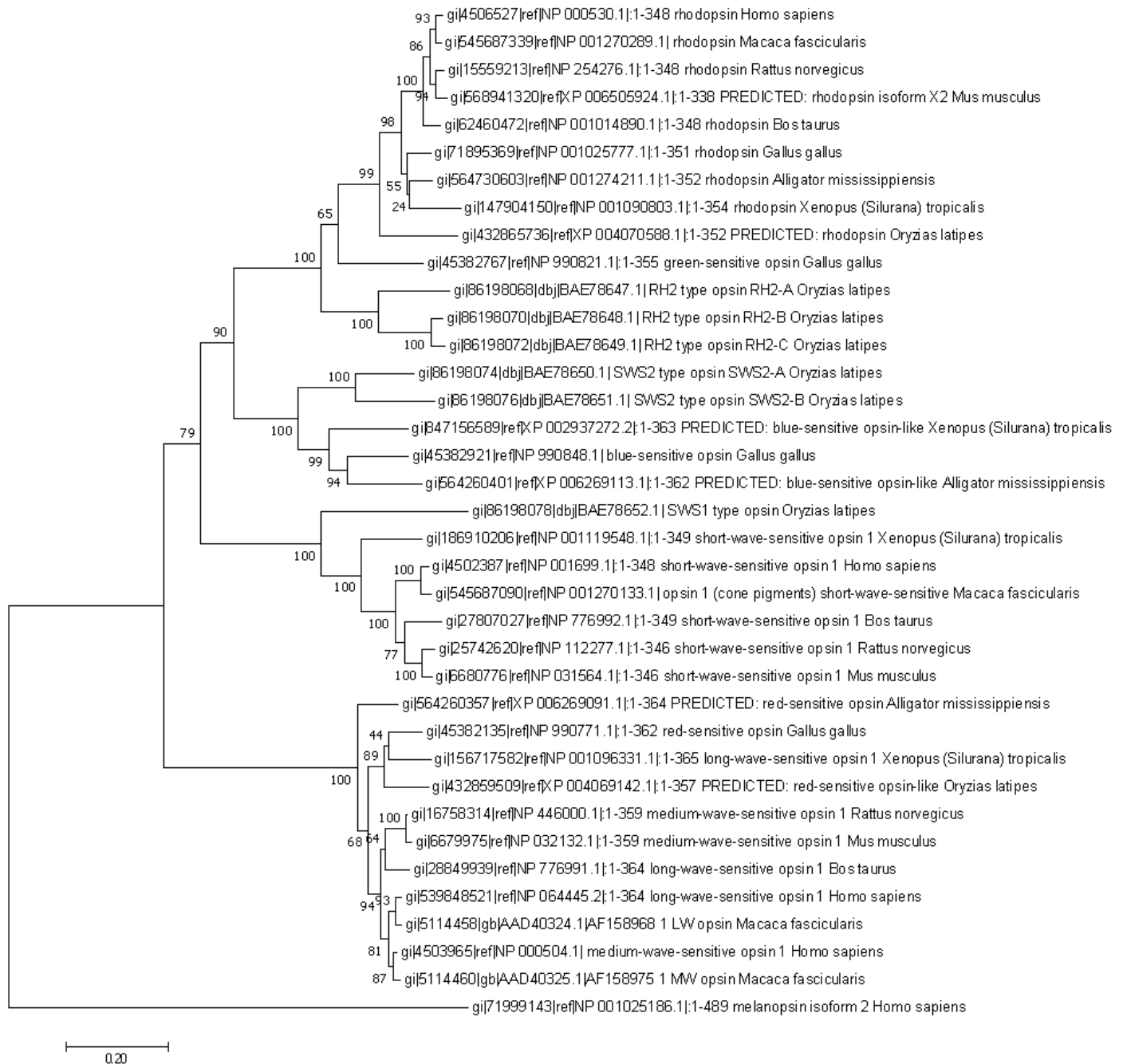
分子系統樹の作成

次に, 予め与えられた脊椎動物のオプシン配列から分子系統樹を作成した. NJ法と最尤法を用いて作成し, bootstrapを100回行った.

bootstrapとは, 得られた樹形がどの程度妥当なものであるかを知るために用いられる手法であり, 系統樹作成以外の場面でも, 様々な信頼区間を得るために用いられる. ここでは, 復元抽出によってサイトをリサンプリングして擬似アラインメントを作成し, それぞれにおいて系統樹を作成し, 元々作成されたクレードが出現する確率を求めることで, 妥当性を判断する. 配列の長さが長くなるにつれ, 擬似アラインメントの数は膨大になり, そのことによって計算量も膨大になる.

NJ法による作成

NJ法とは、近隣結合法ともいい、距離行列法の1つである。この手法では、配列間の距離を調べあげ、それに基づいて段階的に系統樹を作成する方法である。アミノ酸のモデルは、JTTモデルと呼ばれるサイトごとの進化速度が一定のものをを用いた。以下にその結果を示す。



最尤法による作成

最尤法とは、名前の通り、最も尤度が高い系統樹を採用する方法である。ある系統樹とある確率遷移行列が与えられた時、その系統樹が存在する確率（尤度）が求まる。尤度が高い系統樹をひたすら探していく手法である。モデルにはWAGモデルを用い、ある割合のサイトの変異速度がガンマ分布に従うようにした。その結果を示す。



考察

まず, 最尤法でbootstrapを適用すると, 系統樹作成に膨大な時間がかかった. 実際, 最尤法はそれぞれの系統樹に対して尤度を求めていくため, 計算量は圧倒的に大きくなり, bootstrapも組み合わせると膨大になってしまう.

それに対して, 近隣結合法はそれぞれの進化距離が得られれば短時間で得ることができる. ただ, 距離行列ではアミノ酸同士のギャップを考慮していないので, アラインメントしたのちに, ギャップが多い配列によって悪影響が及ぶ可能性がある. 実際, 今回の配列は (感覚的なものではあるが) ギャップが多い印象を覚えたため, この方法で得られた系統樹ではなく最尤法で作成した系統樹で視覚の進化について考察するべきであると考えた.

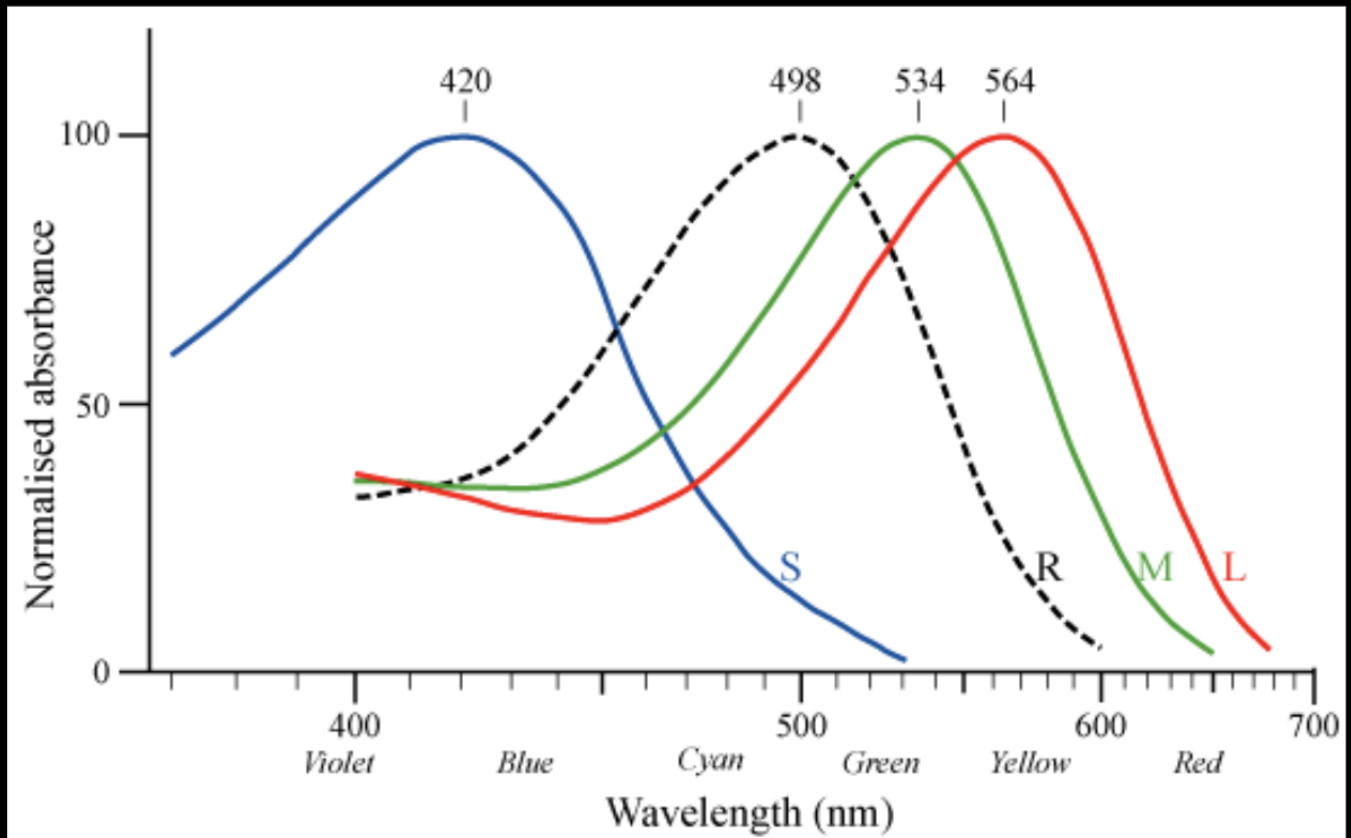
系統樹から、多くのオプシンはメラノプシンと呼ばれる物質とは異なる遺伝子を継いでいる考えることができる。メラノプシンは短波長の光を吸収する光受容体で、体内時計の役割を担っているとされる視交叉上核と呼ばれる脳の領域に直接情報伝達を行う。この物質は、1998年に Ignacio Provencio氏に新しいオプシンとして発見されたものであり、比較的最近に発見されている。実際、メラノプシン遺伝子は10番染色体、青オプシンは7番染色体、緑・赤オプシンは10番染色体に存在することからも、これらが異なるものであるという推定は容易に可能である。

- メラノプシン: <https://www.ncbi.nlm.nih.gov/gene/94233>
(<https://www.ncbi.nlm.nih.gov/gene/94233>).
- 青オプシン: <https://www.ncbi.nlm.nih.gov/gene/611> (<https://www.ncbi.nlm.nih.gov/gene/611>).
- 緑オプシン: <https://www.ncbi.nlm.nih.gov/gene/2652> (<https://www.ncbi.nlm.nih.gov/gene/2652>).
- 赤オプシン: <https://www.ncbi.nlm.nih.gov/gene/5956> (<https://www.ncbi.nlm.nih.gov/gene/5956>).

また、系統樹から、青オプシンと他の2つのオプシンは別のクレードに位置しており、別々に進化してきたと考えることができる。また、ロドプシンは青オプシンがあるクレードにあり、遺伝子的には青オプシンに近いものであると言える。

起こりうる色覚異常

前章で、それぞれの遺伝子の位置を確認した。それに加えて、まず、それぞれのオプシンがどのようなスペクトルの波長を吸収するか確認したい。Wikipediaから、以下のような図が得られた。



これから、赤と緑の吸収スペクトルが驚くほど一致していることが確認できる。実際にこの2つを吸収するオプシンは同じ染色体に位置していることや、系統樹からも相同性が高いことが確認できるため、例えばこれらの遺伝子間で不等交叉が起こることによって異常が起こることが推定される。実際に調べると、赤と緑の色覚に異常が発生する「赤緑色盲」は、青や黄色の異常「青黄色盲」より遥かに高い割合で発生するようである。この事実からも、遺伝子の突然変異による異常の確率（7番染色体の青オプシン遺伝子が突然変異して異常が見られる）よりも、不等交叉などによって起こる可能性の方が高いこともわかる。

参考URL

- <https://www.nig.ac.jp/color/barrierfree/barrierfree1-5.html>
(<https://www.nig.ac.jp/color/barrierfree/barrierfree1-5.html>).
- 高松 進「分子系統学の基礎」（2005）（URL: http://jppa.or.jp/archive/pdf/59_03_20.pdf）
(http://jppa.or.jp/archive/pdf/59_03_20.pdf).
- <https://ja.wikipedia.org/wiki/ロドプシン>
(<https://ja.wikipedia.org/wiki/%E3%83%AD%E3%83%89%E3%83%97%E3%82%B7%E3%83%B3>).

tags: 生体情報工学演習

