

演習課題1 目的のタンパク質のアミノ酸配列・塩基配列情報を入手する

もちいる web site: **Entrez** <http://www.ncbi.nlm.nih.gov/entrez/>

- 1) Entrez にアクセスする
- 2) Protein の選択 (上の黒いバーの中)
- 3) 検索ウィンドウに [axonemal dynein] ([]の中のみ)と入力して、GO をクリックする
- 4) "History"をクリックして、何件のデータがヒットされたかを調べよ。
- 5) 検索語を"axonemal dynein"というように""で囲むと、この連語を含むデータのみが検索される。何件ヒットしたかを調べよ。
- 6) 3) の検索結果にあって、5)にないものには、どのようなものがあるか? 例をあげよ。
- 7) さらに、検索の範囲を狭めることを行う。"Preview/Index"をクリックして、"Add Term(s)"メニューの中から、"Organism"を選択し、検索語として"human"を入力せよ。検索を狭めるには、その後、左下の"AND"ボタンをクリックしたのち、"Preview"ボタンを押す。これで、上の5)の検索結果の中でヒトのデータが検索される。何件あったか?
- 8) これまでのヒットを、GenBank でアノテーション(注釈付け)された遺伝子配列データベースである RefSeq に限定することができる。"Limits"をクリックし、"Limits to"の右下側にある選択メニューの中で、[Only from:]の項で "RefSeq"を選択する。その後、検索語記入枠の横の"Go"ボタンをクリックする。ヒット件数は何件となったか?。
- 9) 今、興味のある遺伝子は、上のヒット件の中で、セン毛運動の不全の見られる Kartagener Syndrome (カータゲナー症候群)に関係したものとする。上の"human"で絞り込んだ方法にならって、この原因となる遺伝子の記述されたデータベースを探し出せ。何件あったか。また、NP で始まるコード番号で表記されているが、その番号は何か? NP 識別子は、断片化された情報を再編し、遺伝情報のギャップをなくすなど種々の編集作業(キューレーション)を経たデータで、高い精度のあることを意味する。
- 10) NP****のコード番号をクリックすると、目的のタンパク質のアミノ酸配列が得られる。このとき、右上の"Link"をクリックし、"Nucleotide"を選択すると、そのタンパク質に対応した mRNA とゲノム配列が得られる。目的の遺伝子は、第何番目の染色体上にあるか?
- 11) 上と同じ方法で、ヒトの繊毛やベン毛にみられる疾患に関係した遺伝子で、ダイニンではないものはどのようなものがあるか調べよ。その中の一例で、タンパク質の種類、染色体上の位置、発見者、参照論文名をあげよ。
- 12) 現在、あなたが興味を持っているヒトの疾患や病気の名前を1つあげよ。
- 13) この疾患や病気に関わるタンパク質を、上と同様の方法で検索して、データベースの中から RefSeq を持つ関連タンパク質を1つ見出せ。そのタンパク分子名、コード番号、ヒトの染色体上の位置について述べよ。
- 14) 上記タンパク質について、マウス、および、ショウジョウバエのゲノムの中から関連するものを見つけ出し、そのタンパク分子名、コード番号を述べよ。