

PubMed検索 あのキーワードのヒット数は?

ラムのネタに困ったので、PubMedで本領域に関係ありそうなキーワードを年ごとに検索してヒット数を調べてみた。本題に入る前に、そもそもPubMedの母集団の数はどのくらいなのだろう。詳細検索の「Limits」欄で、ある年だけに絞り、検索窓は空欄で調べたのが図1だ。2007年で75万件ほどで、この10年でだいたい1.7倍くらい増えている。古くなるほど登録されていない文献が増えるのであろうが、ずいぶん論文数は増

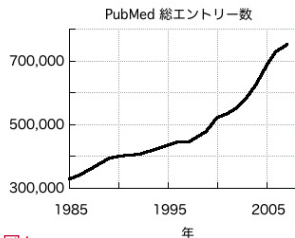


図1

えているのである(ちなみにReview数も同時にわかる。総エントリーの1割以上は総説であり、ずいぶん多く感じる)。

さて本題の「タンパク質の社会」関連。思い付くままにキーワードを選んで調べて、chaperone, Hspあたりがちょうど見やすいようなグラフになるものを集めたのが図2である(他にもいろいろ調べたが最近でも100以下のものは入れていない)。Chaperoninは「chaperonin or GroEL or hsp60」というように無理がない範囲で同義語をOR検索した。これを見ると、例えば、Hspはシャペロンより早くから増え始めているな、とか、シャペロン御三家のHsp70, Hsp90, シャペロニンではHsp70の論文がいつも多い、というようなことが読み取れる。「シャペロン」は94年ごろから増え始めており、Ellisによる1987年のシャペロンの再定義から少しラグがあるのもわかる。興味深いと思ったのは、「protein folding」と「シャペロン」のカーブが奇妙なほど似通っていることである。フォールディングの研究はシャペ

ロンの概念出現のだいぶ前から行われているはずであるが、シャペロンが登場したことでフォールディングの研究そのものも活性化されたと言えるかもしれない。

気になるのは2007年がどれも下がっていること。2007年はまだ未登録の文献が多数あるのかもしれないが、対照として選んだ「プロテオム」は順調に増えているし、総エントリー数も減っているわけではない。分野は成熟を迎えたのであろうか……。今後の追跡調査が必要である。

以上、PubMedの仕組みに詳しいわけでもないのですが何かカン違いしているかもしれないが、案外おもしろいものである。余談であるがヒット数がどんどん下がっていくキーワードを見つけることが案外難しかった。いいのがあれば教えてください。

(田口 英樹)

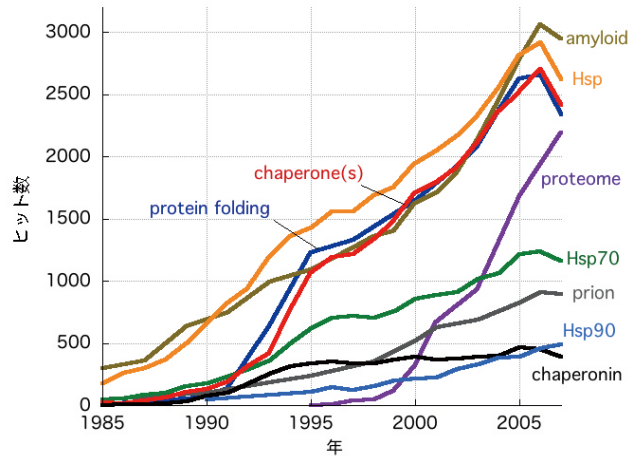


図2