

## 生物医学分野の仮説データベース構築に向けて

高松 邦彦  
村上 勝彦  
上田 國寛

過去10年、様々な種のゲノムがシーケンスされてきた。これらの巨大なデータは、研究スピードを加速させ、また生物医学分野の論文出版数を増大させてきた。研究者は、新しいアイデアを考え、また論文を書くために多くの論文を読む。論文には、2種類の情報が含まれている。1つは実験結果、すなわち「事実」である。もう1つは、仮説である。既に、事実を扱ったデータベースは数多く存在するが、生物医学分野の論文全文で、仮説を扱ったデータベースは存在しない。ここで我々は、生物医学分野の仮説データベースの構築のための初解析を報告する。我々は、5つの論文から、コンピュータ抽出用のルールを手で構築した。次にこのルールで5つの論文から84の仮説を含む文をコンピュータを用いて抽出した。84文のうち、87%にあたる文が、結論と考察に含まれていた。将来、我々は仮説データベースを構築し、web上で公開する予定である。

キーワード：仮説データベース

第32回 日本分子生物学会ポスター発表済