



# 生物学はコンピュータに向かってする時代に？

## ーバイオビッグデータの衝撃ー

高木 利久（生物科学専攻 教授）

白衣を着て日夜実験台に向かっていて、という生物学者に対するイメージは、そろそろ修正されなければならない時代になってきたようだ。コンピュータに向かってきていることも多いし、まったく実験せずにコンピュータだけを使って研究している人も増えつつある。

このような背景には、膨大なデータの出現がある。その最たるものがゲノムデータである。次世代シーケンサーと呼ばれるゲノム決定装置はペタ（10の15乗）オーダーのデータをいとも簡単に生み出してしまふ。データ駆動型科学という言い方があるように、研究者はそれをコンピュータで解析して仮説を立てたり検証したりするようになってきた。逆に、そうやって効率的に研究しないと競争に勝てないようになってきた。コンピュータに向かう時間がいきおい増える訳である。ビッグデータ時代に生命研究も突入したのである。

このシーケンサーの性能の伸びはコンピュータのそれをはるかに凌駕しており、ゲノム情報解析技術の進歩とあいまって、10年ほど前までは15年の歳月と数千億円のお金がかかったのに、いまでは数日で10万円ほどの費用で個人個人のゲノムを決めることも可能になってきた。これにより個人個人の体質（病気のかかりやすさや薬の効きやすさなど）を診断することも可能になってきた。さらに、ゲノムデータ、生化学データ、画像データ、臨床情報、生活習慣情報などを何千人何万人と解析することにより、病気発症のメカニズムを解き明かすこともできるようになってきた。



データの標準化と統合化のための合宿形式の国際ワークショップ

これは疾患研究や遺伝子診断の例であるが、基礎生物学、農学、薬学などのさまざまな研究分野に、さらにはバイオビジネスにも、バイオビッグデータ時代が到来しつつあり、生命研究のスタイルを根底から変えようとしている。

ビッグデータの意味するところは、膨大なデータから統計処理や機械学習と呼ばれる手法を使って規則性を見いだすことにあるが、生命研究のデータは、他のサイエンスやビジネスのそれとは異なり、多様性、曖昧性、文脈依存性、冗長性、多義性などの特徴があるため、計測したデータをそのままコンピュータで解析すればいいというわけに行かない。

筆者の研究室では、科学技術振興機構バイオサイエンスデータベースセンター、情報・システム研究機構ライフサイエンス統合データベースセンター、遺伝学研究所DDBJ（DNA Data Bank of Japan）センターと連携して、膨大なデータと知識の整理統合とそのためのコ

ンピュータ技術の研究開発を行っている。また、世界各国のデータベース開発者と協力してフォーマットや語彙の国際標準化に取り組んでいる。

バイオビッグデータの解析には、生物学や医学の知識とコンピュータの技術の両方をもった人材が不可欠であるが、我が国ではそのような人材が圧倒的に不足している。筆者の属する生物情報科学科ではそのような人材の教育に精力的に取り組んでいるがまったく需要に追いついてないのが現状である。意欲のある若い方がこの分野に飛び込んで、これからの新たな生物学を切り開いてくることを期待したい。



## 星の最期の大爆発の瞬間を探して

諸隈 智貴 (天文学教育研究センター 助教)

宇宙の年齢(約138億年)を1年とすると、人の寿命(約80年)はたったの0.2秒でしかない。そんなまばたきほどの短い間であっても、宇宙は決して静的ではなく、ダイナミックな姿を見せる。その一つが超新星爆発と呼ばれる、(一部の)星がその一生の最期に迎える大爆発である。

2011年のノーベル物理学賞(遠方超新星爆発の観測による宇宙の加速膨張の発見)の対象となった観測が行われた1990年代以降、超新星の観測は精力的に進み、現在では年間数百個もの発見報告がなされる。しかし、そのほとんど全ては爆発から数日経過したものである。というのも、超新星爆発が稀な現象であること、我々が「どの星がこれから爆発するか」を予言できない(厳密にいうとニュートリノを観測できれば事前にわかるがニュートリノの検出自体も難しい)ことが原因である。ではどうすればよいか。「できるだけたくさんの星をまとめて一度に観測する」のである。それには何

が必要か。「一度に空の広い領域を観測することのできる望遠鏡とカメラ」である。

私たちの研究グループでは、世界有数の広視野観測能力を持つ中年の(昨年40周年を迎えた)木曾観測所シュミット望遠鏡と、2012年に完成した2歳の優秀な息子KWFC(Kiso Wide Field Camera)を使って、超新星爆発の瞬間を捉えるプロジェクト木曾超新星探査(Kiso Supernova Survey: KISS)を2012年4月から行っている。爆発の瞬間の観測は非常に難しいため、1年に1個見つけられるかどうかの野心的なプロジェクトであるが、発見できればその科学的価値は非常に高い。観測では、6400万画素のデータが約5分おきに生成され、それをパソコンで処理する。主に3人の研究者と2人の大学院生・大学生が交替で担当し、この膨大なデータをチェックし、超新星の候補を見つけ出す。この珍しい現象を見落とさないよう、



■ KWFCで取得したアンドロメダ銀河の広視野画像

アマチュア天文家の方々にもボランティアでご自宅からご協力いただいている。また、年間100晩もの超新星探査観測を安定して行うためには、観測所員による望遠鏡・カメラの日々のメンテナンスも欠かせない。この三位一体の取り組みは、NHK長野放送局によるドキュメンタリー番組、さらにはドラマ「木曾オリエオン」の題材としてとりあげいただくこともできた。あとは爆発の瞬間が観測できる幸運を待つのみである。

KWFCの生み出す、1日あたり約100ギガバイトにも及ぶデータは世界的に見ても膨大なものであるが、天文学では、より大量のデータ、いわゆるビッグデータをリアルタイムに処理することが求められる時代を迎えている。私たちも、KWFCの後継機として、CMOS(Complementary Metal Oxide Semiconductor)センサを用いた超広視野カメラ(新しい娘の名前を木曾町にゆかりのある巴御前にちなんでTomo-e Gozenと命名)の開発を始め、次は、人間のまばたきくらいの時間で宇宙の変動を捉える観測を行おうとしている。



■ 観測所員の日々のメンテナンスにより、1974年から立派に動き続けている望遠鏡