

産業技術総合研究所 生命情報科学技術者養成コースにおける 計算機実習の紹介

Introduction to computer practice CBRC bioinformatics engineer course in National Institute of Advanced Industrial Science and Technology (AIST)

野口 保

Tamotsu Noguchi

薬学教育研究センター

E-Mail: noguchit@my-pharm.ac.jp

1. 生命情報科学技術者養成コース

2005年度から2009年度までの期間、文部科学省科学技術振興調整費 新興分野人材養成ユニット(企業等の研究者、技術者の再教育プログラム)の補助により実施された産業技術総合研究所(産総研)生命情報科学技術者養成コースは、2010年12月に公表された「平成22年度研究評価部会ライフサイエンス系人材養成評価作業部会評価結果報告書」¹⁾で、総合評価Sの高評価を得た。本技術者養成コースは、産総研 臨海副都心センターに設立された生命情報科学研究センターにおいて2005年度からスタートし、2009年度からは後継の生命情報工学研究センターにおいて引き継がれた。

1.1. 生命情報科学技術者養成コースの目的

生命情報科学は、バイオインフォマティクスとも称される1990年頃から始まった生物学と情報科学の境界領域の学問分野で、それ以前は系統的に教育する機関はなく、現在もその数は少ない。また、ヒトゲノム解析プロジェクトや構造ゲノミクスプロジェクトなどの国際連携プロジェクトや次世代シーケンサーや遺伝子発現実験技術などの発展により、大量の情報(文献情報、実験データなど)が蓄積されたため、バイオ分野の研究において、バイオインフォマティクス技術の必要性が認識されるようになってきた。

このような状況において、バイオインフォマティクスの専門教育を受けた社会人はまだ少なく、社会人と

して活躍中の人材に学習機会を提供する場がなく、その必要性が増していた。また、企業内ではバイオインフォマティクス技術者の密度が低く、孤立しがちなため、情報交換をする人材ネットワークの構築も望まれていた。さらに、生物学と情報科学はともに進展が早く、バイオインフォマティクス分野も同様のため、知識や技術の陳腐化が起こり得る。そのため、それらのフォローアップも重要となっている。

このように、社会人になってからも実践的かつ先進的な教育・再教育を施す目的で、産総研 生命情報科学研究センターにおいて生命情報科学技術者養成コースが実施されることになった。

1.2. 生命情報科学技術者養成コースの内容²⁾

生命情報技術者養成コースでは、バイオインフォマティクスの初心者を中心にカリキュラムを作成したが、経験者を対象にしたコースも含め、表1に示した4つのコースを用意した。

	対象	養成方式	到達スキル
バイオインフォマティクスリーダ養成・再教育コース	現役B技術者 PJ立案者	末所してOJT セミナー聴講	最先端の知識 問題解決能力
バイオインフォマティクス速習コースI (e-ラーニング 2009年～)	未経験者 基礎知識不足	夜間20コマ 理論・応用を講義	B技術者 認定試験合格 ^{*)}
バイオインフォマティクス速習コースII	未経験者 基礎知識不足	昼間30コマ 講義・計算機実習	上記+計算機による 解析力
創薬インフォマティクス技術者養成コース	創薬に関する 背景知識あり	短期集中30コマ 講義・計算機実習	創薬ソフトを実務レベルで操作

*)日本バイオインフォマティクス学会認定

表1 生命情報科学技術者養成コースの各コース

対象を社会人としたので、バイオインフォマティクス速習コースⅠでは、企業の就業時間後を想定して夜間(18時30分以降)に行ったり、創薬インフォマティクス技術者養成コースでは、夏休み期間に短期集中で行ったりと、社会人でも受講しやすい工夫をした。

講義風景と実習風景を図1と図2に示す。講義・実習会場は、産総研 臨海副都心センター別館8階のコラボレーションコーナーで、講義中やセミナー中(図1)はモニターが収納される机にパソコンが常備されており、実習時(図2)にはモニターを取りだし、受講者が各自パソコンを操作しながら実習できる環境を構築されている。また、テレビ会議システムを用いて講義を他の会場に発信するシステムも構築し、バイオインフォマティクス速習コースⅠの講義のみ、産総研つくばセンターで受講希望者があった場合、そのシステムを用いて発信した。



図1 講義風景



図2 実習風景

また、バイオインフォマティクス速習コースⅠの講義は、2008年度の講義を基に図3のようなeラーニ

ングをウェブシステムから受講できるシステムを構築し、2009年度から受講できるように公開した。

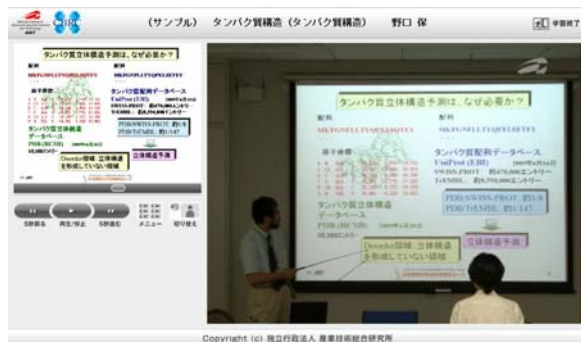


図3 eラーニング

1.3. 生命情報科学技術者養成コースの実績

生命情報技術者養成コースの修了者数は、5年間で合計446名でした。各コース別では、バイオインフォマティクス速習コースⅠ:246名(内eラーニング:88名)、バイオインフォマティクス速習コースⅡ:81名、創薬インフォマティクス技術者養成コース:103名、リーダー養成・再教育コース:16名でした。毎年ほぼ定員を満たし、受講者を選抜する年もあったことから、eラーニングでの開講を行うことになった。そのeラーニングは、最終年度の1年間であったが、88名の修了者を出し、修了者の大幅増加に貢献し、修了生を図4に示すように全国規模に広げる効果があった。

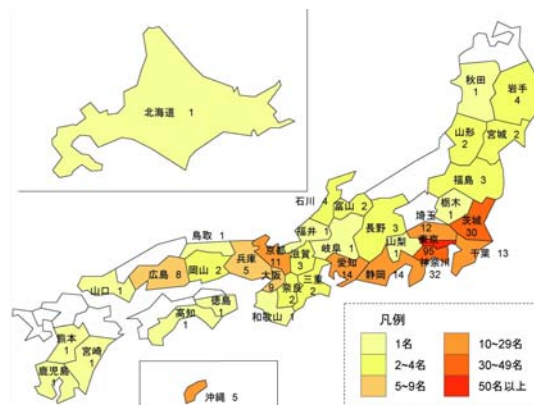


図4 eラーニング修了者分布

また、創薬インフォマティクス技術者養成コースは、毎年受講生を選抜するのに苦勞するほどの需要があり、多くの製薬メーカーからの修了者を輩出した。

また、本技術者養成コース修了者間の人材ネットワークの構築も進み拡がり続けている。

一方、応募者のプロフィールを調査したところ、バイオインフォマティクス速習コースは、全般に公的研究機関や大学関係者が約4割と高いのに対し、創薬インフォマティクス技術者養成コースは、製造業（医薬品、化学、食品）関係者が約5割を占めていた。バイオインフォマティクス速習コースは、社会人にバイオインフォマティクスの学習機会を与え、創薬インフォマティクス技術者養成コースは、企業の現場での再教育と人材ネットワークの構築と言う、本コースの目的を果たせた内容であったと考えている。

2. 計算機実習例：タンパク質構造解析

本計算機実習は、バイオインフォマティクス速習コースⅡの中の1コマで行ったタンパク質構造解析～タンパク質立体構造予測実習～で、ヒトプリオンタンパク質を用いたタンパク質立体構造予測実習である。本実習の流れを図5に示す。実習に用いたソフトウェアは、大きく2グループに分類できる。1つは、ウェブで利用できるもので、Disorder 予測 (POODLE)³⁾、膜貫通部位予測 (SOSUI)⁴⁾、ドメイン予測 (DomCut)⁵⁾、二次構造予測 (PSI-RRED)⁶⁾、鋳型構造検索 (PSI-BLAST)⁷⁾、構造評価 (Verify3D)⁸⁾がある。もう1つは、ソフトウェアを手元の計算機にインストールして利用するもので、構造構築 (MODELLER)⁹⁾がそれに当たる。一般に、計算時間がかかるソフトウェアは、このように手元で行うものが多い。

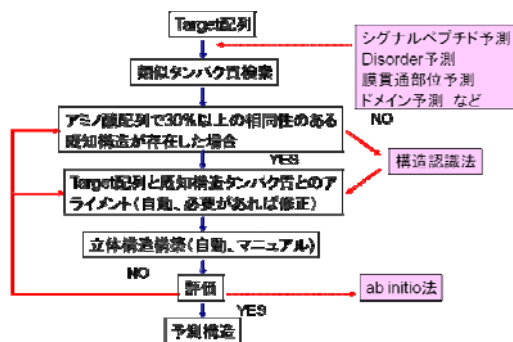


図5 実習の流れ

本実習では、ウェブで利用するソフトウェアは、図

2にあるような受講者各自に割り当てられたパソコンを用いて行い、MODELLER による構造構築だけ実習用に用意した Linux サーバー (Xeon 3.0GHz 16 CPU) で行った。

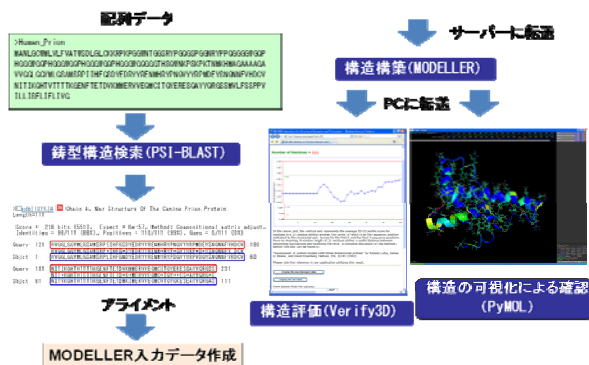


図6 パソコンと Linux サーバーを用いた

実習では図6に示すように、MODELLER の入力データをパソコン上作成し、その入力データファイルを Linux サーバーに転送し、X 端末エミュレータを用いて MODELLER を起動する。MODELLER の計算が終了したら、構造評価を行うために、構築されたタンパク質立体構造予測結果ファイルをパソコン上に転送する。

パソコンに転送したタンパク質立体構造予測結果は、ウェブで利用できる構造評価ソフトウェア (Verify3D) で評価値を計算し、X 線結晶解析で求めた正解構造と同様に計算した評価値と比較して評価する。また、構造表示ソフトウェア (PyMOL) で構造を可視化することにより、X 線結晶解析で求めた正解構造と比較して評価する。

本実習では、タンパク質立体構造予測の流れを理解してもらうことを目的にしているため、配列および構造類似性の高い鋳型が存在しているターゲットタンパク質を用いている。そのため、実習自体は非常に容易にタンパク質立体構造予測が可能である。タンパク質立体構造予測の難易度が高いターゲットの実習は創薬インフォマティクス技術者養成コースに用意されている。

3. まとめ

本技術者養成コースでは、受講者一人に対して

パソコン1台を用意し、実践的な技術習得が可能な実習環境を構築した。また、本技術者養成コースは、2005年度より5か年にわたって実施し、のべ446名の修了者を輩出し、本技術者養成コースの目標を達成し、高い評価を受けた。本技術者養成コースのノウハウを生かし、2010年4月より生命情報科学人材養成コンソーシアム¹⁰⁾が設立され、講習会を継続している。

謝辞

生命情報科学技術者養成コースの初代研究代表者である東京工業大学秋山泰教授および二代研究代表者である産総研生命情報工学研究センター浅井潔研究センター長に感謝致します。また、本技術者養成コースの全ての関係者、特に計算機環境の維持管理および本発表の資料を提供くださった坂井寛子氏、寺田朋子に深く感謝いたします。

参考文献

- 1) 平成22年度研究評価部会ライフサイエンス系人材養成評価作業部会評価結果報告書
http://www.jst.go.jp/shincho/22hyouka/09life_s hin22.pdf
- 2) 平成22年度科学技術振興調整費事後評価対象課題の成果発表会(平成23年3月1日)
http://training.cbrc.jp/page/index.php/doc/seika_poster1.pdf
- 3) Disorder 予測ソフトウェア POODLE
<http://mbs.cbrc.jp/poodle/poodle.html>
- 4) 膜貫通部位予測ソフトウェア SOSUI
http://bp.nuap.nagoya-u.ac.jp/sosui/sosui_submit.html
- 5) ドメイン予測ソフトウェア DomCut
<http://www.bork.embl-heidelberg.de/~suyama/domcut/>
- 6) 二次構造予測ソフトウェア PSI-PRED
<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>
- 7) 鋳型構造予測ソフトウェア PSI-BLAST
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

- 8) 構造評価ソフトウェア Verify3D
http://nihserver.mbi.ucla.edu/Verify_3D/
- 9) 構造構築ソフトウェア MODELLER
<http://salilab.org/modeller/>
- 10) 生命情報科学人材養成コンソーシアムホームページ <http://training.cbrc.jp/>