

奈良先端大蛋白質機能予測学人材養成ユニット

計画の目標・概要

1. 目標

奈良先端科学技術大学院大学におけるバイオインフォマティクス分野の人材養成の一環として、修士・博士課程学生の教育を行い、蛋白質機能予測学分野の技術者・研究者を養成する。

目標：平成17年度までに修士学位取得者計20名程度を、博士学位取得者計5名程度を、各年度、民間企業、国公立研究機関に供給する

2. 内容

情報科学研究科情報生命科学専攻に蛋白質機能予測学講座（ユニット）を設置する。

客員教授として郷信広博士（原研・量子生命情報解析グループリーダー、兼任）と土居洋文博士（セレスター・レキシコ・サイエンシズ株式会社社長、兼任）を、客員助教授として Gautam Basu 博士、川端猛博士、土井晃一博士を招聘する。

蛋白質機能予測学講座、さらに、情報生命科学専攻における、人材養成の補助と専門研究を推進するために、9名の博士研究者を雇用し、より高度なバイオインフォマティクス研究者として育成する。

蛋白質機能予測学講座に招聘する研究者は、情報生命科学専攻における講義・演習に参画すると共に、蛋白質の構造と機能の情報学的解析を始めとする専門的教育と研究を推進する。

諸外国の現状等

1. 現状

世界各地で、ゲノム配列決定のプロジェクト、大規模に立体構造の決定を目指すプロジェクト（構造ゲノム科学）が進行しており、それらのデータを維持・管理する拠点が作られ、多くのバイオインフォマティクス研究がなされている。

蛋白質の立体構造や機能予測に関して、Structural Genomix 社などのベンチャー企業が構造決定を推進し、一方で Molecular Simulation 社などではバイオインフォマティクス手法による蛋白質構造機能予測を進めている。

2. 我が国の状況

いくつかの研究機関において世界的に優れたバイオインフォマティクス研究が行われているが、組織的な人材養成の体制整備が遅れており、国全体としてはバイオインフォマティクス研究推進のパワーが不足していることが問題となっている。

計画進展・成果がもたらす利点

1. 世界の水準との関係

奈良先端大学院情報生命科学専攻の教育体制が充実し、我国のバイオインフォマティクス研究の将来を担う技術者・研究者の組織的養成が進み、我国のバイオインフォマティクス研究パワーが世界的な水準に充実するために貢献できる。

本申請で強化しようとする領域である、立体構造情報から各遺伝子産物の生物学的機能への理解を目指す計算構造生物学や生物情報学は、世界的にもこれからの発展が期待されている分野であり、世界をリードする技術者・研究者の養成が期待される。

2. 波及効果

蛋白質の立体構造や機能の予測はゲノム創薬などを通じて産業界への寄与が大きい。また遺伝子多型情報にともなった蛋白質立体構造変化を調べることで、テーラーメイド医療にも貢献する。

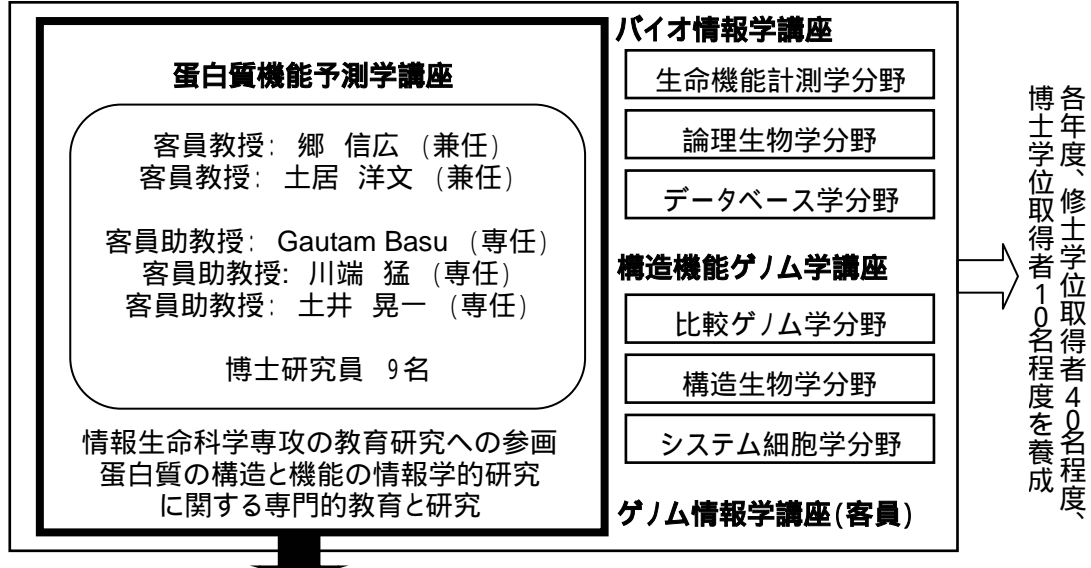
奈良先端科学技術大学院大学は関西文化学術研究都市の研究拠点であり、地域との連携により、関西地区における企業、大学等におけるバイオインフォマティクス分野の研究の活性化に寄与する。

奈良先端大蛋白質機能予測学人材養成ユニット

代表者名：鳥居 宏次（奈良先端科学技術大学院大学学長）

実施機関：奈良先端科学技術大学院大学

奈良先端科学技術大学院大学情報科学研究科情報生命科学専攻



平成17年度までに、修士学位取得者計20名程度、博士学位取得者計5名程度を、民間企業、国公立研究機関に供給する

人材養成計画の趣旨・概要

【人材養成計画の趣旨】

バイオインフォマティクス分野等の新たな研究領域を担う研究者、技術者の養成のためには、主導的な研究教育機関を機動的に整備することが重要である。奈良先端科学技術大学院大学は、先端科学技術分野における高度な基礎および応用研究を推進するとともに、先端科学技術分野の研究開発に携わる人材を組織的に育成することを使命としている。また、学部を持たない大学院大学であるという組織の機動性を生かし、先端科学技術分野間の融合領域の教育研究に積極的に対応していくことも本学の重要な課題である。こうした背景の下、バイオインフォマティクスの教育研究を推進するために、情報科学研究科及びバイオサイエンス研究科の組織を再編し、平成14年度に情報科学研究科内に情報生命科学専攻を設置し、博士前期課程および後期課程の教育を開始した。

情報生命科学専攻は、ゲノム情報科学、ゲノム機能解析、蛋白質構造機能解析の3つを研究教育の柱としている。その組織は大講座制であり、2基幹講座（教育研究領域）、すなわちバイオ情報学領域および構造機能ゲノム学領域、さらに1客員講座、すなわちゲノム情報学講座より構成されている。バイオ情報学領域は、3つの教育研究分野、すなわち、ポストゲノム研究を支える基盤技術である「データベース学」分野、生命体の様々な現象を情報システムとして研究し、システム生物学への展開を図る「論理生命学」分野、ミクロからマクロまでの生命現象の計測手法の開発から医療面への応用を図る「生命機能計測学」分野から構成される。また、構造機能ゲノム学領域は、蛋白質構造機能解析を行う「構造生物学」分野、ゲノム機能解析から細胞システムの解明を目指す「システム細胞学」分野、ゲノム・遺伝子の進化・多様性の原理の解明を進める「比較ゲノム学」分野から構成される。

一方、バイオインフォマティクス分野は現在急速に発展している分野であり、1大学のスタッフのみでは必要な全分野の教育をカバーすることが困難であることも否めない。そのため、客員講座に加えて、学内措置で運営している「民間機関等との連携による大学院教育」（連携講座）、すなわち生体膜情報学講座（産業技術総合研究所との連携）、計算神経科学講座（国際電気通信基礎技術研究所との連携）などを活用し、基幹6講座ではカバーできない研究分野については学内外の関連研究者との連携による教育を推進している。

こうした本大学院大学におけるバイオインフォマティクス教育研究体制の整備の一貫として、ゲノム創薬などの産業応用を考える上でも重要である蛋白質機能予測に関する教育研究の充実を図るために、科学技術振興調整費による人材養成ユニット設置の申請を行い、採択されたものが、「蛋白質機能予測学」講座である。本講座はその設置主旨から、情報生命科学専攻の教育研究体制の迅速な立ち上げへの寄与が第一の責務である。また、蛋白質機能予測学に関する専門的な教育を講座において行い、修士及び博士を養成することが第二の責務である。大学院学生の教育において研究室での研究活動が大きな比重を占めるため、質の高い教育を行うためには、高レベルのアクティビティを有する研究室を整える必要がある。そのためには、教授・助教授の人材だけではなく、現場での研究を遂行しつつ、大学院学生の指導するための博士研究員も雇用することとした。一方で、特に国内ではバイオインフォマティクスの博士の取得者の数が少ないということから、関連分野で学位を取得している研究者、あるいは海外の研究者を積極的に博士研究員として雇用し、彼らを本ユニットにおいて再教育することで、より短期でバイオインフォマティクスの独立した研究者に育てあげることも推進している。

【人材養成計画の概要】

(1) 情報生命科学専攻の教育体制の整備に寄与すると共に、情報生命科学専攻における教育活動に参加し、バイオインフォマティクス分野の研究者・技術者として、修士及び博士の学位取得者を養成する。

(2) 教育・養成する大学院学生の目標人数として、計画当初においてはユニットとして「各年度、修士課程学生6名程度、博士課程2名程度を選考し、学位取得のための専門教育を行う」と設定した。しかし、その数値目標は制度上可能な上限が設定されており、現実には初年度の学生の受け入れが事実上不可能であったこともあり、平成15年度の計画からは以下のような形で目標人数を設定した。

a) 情報生命科学専攻全体として、各年度修士学位取得者40名程度、博士学位取得者10名程度を養成する。専攻設置より現在2年目であるので、3年目までに、修士学位取得者40名、5年目までに、修士学位取得者120名、博士学位取得者20名が数値目標となる。

b) 蛋白質機能予測学講座として専門的教育を行い、5年目までに修士学位取得者計20名程度、博士学位取得者計5名程度を民間企業、国公立機関に供給することを目標とする。この目標からは、3年目までには、修士学位取得者6名の養成が数値目標となる。

(3) 蛋白質機能予測学講座さらに情報生命科学専攻全体における人材養成の補助と専門研究を推進するために、博士研究員を雇用する。

成果の概要

【人材養成計画の進捗状況】

平成 13 年度は、10 月に郷信広博士、土居洋文博士を非常勤の客員教授として、川端猛博士、土井晃一博士を常勤の客員助教授として招聘し、人材養成ユニットの研究環境の整備を行うと共に、情報生命科学専攻の授業カリキュラム整備等の新専攻の学生受入準備に参画し、今後の研究・教育活動のための準備的な活動を行った。3 月には Gautam Basu 博士を常勤の客員助教授として招聘した。平成 14 年度は、情報生命科学専攻が正式に設立され、47 名の前期課程学生と 8 名の後期課程学生が配属され、情報生命科学の人材養成を目指した新しいカリキュラムによる授業も開始された。人材養成ユニットで雇用されたスタッフによる蛋白質機能予測学講座も本専攻の研究室として積極的に教育・研究に参加した。また、専攻配属学生のうち、3 名の前期課程及び 1 名の後期課程の学生が蛋白質機能予測学講座に配属され、このうち 1 名はこの年度に博士前期課程を短期修了した。平成 15 年度は、4 月に 30 名の前期課程学生と 14 名の後期課程学生が新たに情報生命科学専攻に入学・進学し、その中から蛋白質機能予測学講座にも 3 名の前期課程学生が配属された。また、前年度における専攻での教育システムの評価を行い、一層の教育体制の充実を図った。10 月にも 2 名程度、専攻に配属が行われる見込みである。なお、15 年度末には、45 名が前期課程を修了見込みであり、その進路は就職・復職など 29 名、後期課程進学 16 名の予定である。人材養成ユニットは、こうしたバイオインフォマティクス分野の研究者・技術者の組織的な養成という、我国で最初の取り組みに大きく貢献している。また、修士の学位を持つものを社会に輩出するという点でも、情報生命科学専攻全体としては目標を上回ることができ、また短期修了制度により即効的な人材供給も実現した。人材養成ユニットとしては、かならずしも当初の数値目標を達成できていないが、今後、目標の達成に迫ることができる体制が整備された。

一方、この間 1 2 名の博士研究員を雇用し、教育の補助を行いつつ、バイオインフォマティクスの研究に従事させることにより、独立した研究者として養成することを進めた。この点は当初計画には明示されていないが、バイオインフォマティクス分野での短期的な人材養成という意味で、多大な意義があったと評価できる。

【目標に対する達成度】

(1) 情報生命科学専攻における教育活動への貢献

平成 14 年度から情報生命科学専攻が正式に設置され、情報科学と生物科学の融合領域の教育を目指す新しいカリキュラムによる授業が開始された。人材養成ユニットの教官も多くの講義・演習を担当しており、専攻の教育活動に大きく貢献している。

(2) 情報生命科学専攻及び人材養成ユニットでの大学院学生の養成

以下に現在の情報生命科学専攻の学生総数を示す。学生の出身は情報系とバイオ系の双方からであり、人材養成ユニットに所属された学生はうち博士前期 1 年が 3 人、博士前期 2 年が 4 人、博士後期 2 年が一人である。

	情報系	バイオ系	人材養成ユニット(内数)	合計
博士後期 2 年(14 年度配属)	6	2	1	8
博士後期 1 年(15 年度配属)	10	4	0	14
博士前期 2 年(14 年度配属)	29	18	4(内 1 名短期修了)	47
博士前期 1 年(15 年度配属)	19	13	3	32

(博士前期 1 年については、平成 15 年度 10 月入学者における見込み数を含む)

これらを、目標との関連で整理すると以下ようになる。

養成する人材のレベル	情報生命科学専攻全体の実績(目標)	うち蛋白質機能予測学講座実績(目標)
・博士前期課程	46 名(3 年目までに 40 名)	4 名(3 年目までに 6 名程度)
・博士後期課程	0 名(3 年目までは 0 名)	0 名(3 年目までは 0 名)

(実績は 15 年度までに当該課程を終了または終了見込みの者を記載)

情報生命学専攻全体としては3年目までの目標を達成しているが、人材養成ユニットの蛋白質機能予測学講座については、博士前期課程は目標6人に対し、4人の修士学生の養成にとどまっている。今後目標達成のために一層努力したい。

(3) 博士研究員のバイオインフォマティクス研究者としての養成

質の高い大学院教育を行うためには、高レベルのアクティビティを有する研究室を整える必要があり、そのためには、教授・助教授の人材だけではなく、現場での研究を遂行し、大学院生の指導を行う博士研究員も雇用することとした。しかし、実際には、バイオインフォマティクスの博士取得者の数が特に国内では少なく、適当な人材を集めるのが困難な状況であった。そこで、関連分野で学位を取得している研究者、あるいは海外の研究者を積極的に博士研究員として雇用し、彼らを本ユニットにおいて再教育することで、バイオインフォマティクスの独立した研究者に育てあげ、もって教育・研究に当たらせることにした。様々な関連分野の学位取得者を博士研究員としてのべ12人雇用し、7人は現在も在職中である。残りの5人は本ユニットを転出し、その多くは本ユニットでの経験を生かし、バイオインフォマティクス分野の研究者として学内外で活躍している。

【養成された人材の概要】

情報生命科学専攻によって教育された大学院学生の人材の概要

平成15年度末には、情報生命科学専攻の第1期学生45名が前期課程を修了見込みであり、その進路は就職・復職など29名、後期課程進学16名の予定である。その就職内定先には以下のようなものがある。

ソニー	NEC 2名	日立製作所	三洋電機	NTT 西日本
リコー	トヨタテクノサービス	デンソー	日立ハイテクノロジーズ	
日立システムアンドサービス	日立公共システムインジニアリング	NEC ソフトウェア	キャノンシステムソリューション	エスエムジー
ハネッコーポレーション	東京化学同人	野村不動産	ミツカングループ	シスメックス
タカラバイオ	カバールメント	インテックウェブ&ゲム	ホルデイスクファーマ	アボットジャパン

従来の情報科学研究科における主要な就職先である電気系・通信系・機械系メーカーに加えて、ソフトウェア系、出版系、バイオ系など多彩な企業に人材を輩出していることが分かる。また、バイオインフォマティクス分野に事業展開を行っている、あるいは計画中と思われる企業も見られることから、バイオインフォマティクス分野に関わる高度専門技術者を即効的に供給するという情報生命科学専攻および人材養成ユニットの当初目標は、順調に達成しつつあると考えられる。

一方で、情報生命科学専攻および人材養成ユニットは、新興分野における研究・教育を推進していることから、研究志向の強い学生が多いことも特徴である。実際に、現在の博士前期課程2年生45名中16名は博士後期課程への進学が内定あるいは受験準備中であり、これは情報科学研究科全体の進学率(平成15年度は約29%の見込み)よりも高い。このことは、本プログラムの継続にしたがって、バイオインフォマティクス分野の研究者の養成が大いに期待できることを意味している。

人材養成ユニットによって教育された大学院学生の人材の概要

人材養成ユニットの教官は「蛋白質機能予測学講座」という名の講座を構成しており運営上、二人の客員教授を代表として、「郷グループ」と「土居グループ」に分かれている。

郷グループでは、蛋白質立体構造データを中心としたインフォマティクス研究を進めている。「構造ゲノミクス」プロジェクトの進展に伴い、立体構造データの数は現在より大きく増大することが予想され、膨大な立体構造データを生かすための情報学的研究が望まれている。具体的には (A)蛋白質の立体構造分類と機能との相関、(B)蛋白質の立体構造予測、(C)膜蛋白質の構造解析、(D)蛋白質-リガンド相互作用、(E)蛋白質-蛋白質間相互作用の5つの分野に取り組んでいる。土居グループでは、製薬業界で要求の強いゲノム創

薬におけるインフォマティクスの研究を念頭において、ゲノム情報・蛋白質アミノ酸配列や文献情報などをベースとした情報科学的解析による蛋白質機能予測方法の研究開発を進めている。具体的には、土井客員助教授を中心に大腸菌や枯草菌などに関する微生物文献アブストラクトから蛋白質相互作用に関する情報を抽出するためのシステムの開発を開始し、その第一歩として文献から蛋白質名を抽出する機械学習システムの開発を行っている。

本講座は平成14年度から学生の受入を始めた。平成14年度は3人の博士前期課程学生と1名の博士後期課程学生が郷グループに、平成15年度は2名の博士前期課程学生が郷グループに、1名の学生が土居グループに配属された。また、平成15年度に博士前期課程2年の学生が、他研究室から本講座に転配属し、郷グループに配属された。学生の学部での専攻は様々であり、配属された博士前期課程の学生7人のうち、情報系が3人、生物学系が2人、化学系、応用物理系がそれぞれ一人ずつとなっている。研究室内においては、情報系の学生には生物学の知識の取得を、バイオ系の学生にはプログラミング等の情報科学の技術など関連分野の知識を取得するよう指導を行っており、実際に幅広い分野の授業を積極的に履修している学生が多い。また、郷グループにおいては、英語は研究者として必要不可欠なスキルであるとの考えから、研究室の全てのセミナーを英語で行っている。当初、英語での意思疎通に困難を伴う学生が多かったが、2年を経過した現在、その多くは英語でのディスカッションに慣れ、確実な英語力の向上が見られた。

配属された学生は各グループの研究方針にそって、主に蛋白質立体構造データを生かしたインフォマティクスについての研究活動を行っている。平成14年度に入学した4人の学生の修士研究のテーマは、それぞれ、「2次元構造予測等を利用した蛋白質間の弱い相同性を認識する手法の開発」、「表面アミノ酸分布に基づく蛋白質間相互作用部位の予測」、「膜に対する傾きを考慮した膜貫通ヘリックス予測の試み」、「モチーフ情報を用いた蛋白質間弱相同性認識法の改良」であり、いずれもポストゲノム配列時代における重要な課題である。4人の学生の一人は修士課程を1年間で短期修了し、情報生命科学専攻からの最初の博士前期課程の修了生となった。彼は博士前期課程を終えたあと特許庁に戻り、現在「生命工学」という、微生物、酵素、遺伝子工学の特許を扱う部署に配属となり、本学で学んだバイオインフォマティクスの知識を勤務に生かしている。また別の一人は博士後期課程に進学予定であり、本年度生物物理学会年会でポスター発表を行うなど、研究者として自立するための訓練を積極的に受けている。

人材養成ユニットにおいて雇用し再教育を受けた博士研究員の人材の概要

人材養成ユニットでは、バイオインフォマティクス以外の他分野の背景を持つ博士研究員に対して必要なスキルを与えることで、バイオインフォマティクス分野への展開を奨励し、かつそうした研究を核として、大学院教育を実施している。本ユニットではのべ12人のポスドクを雇用し、現在7人を継続雇用中である。雇用した12人のうち5人が外国人であり、優秀な人材であれば外国人も積極的に雇用する方針である。5人の国籍も中国、インド、マレーシア、アルバニアと多彩である。郷グループではセミナーを英語で行うなど外国人を受け入れやすい環境作りを心がけた。

12人の博士研究員のうち、学位取得時の専攻は、化学、数学、生化学、免疫学、原子核物理、画像情報処理、パターン認識、機械学習など極めて多彩である。彼らは雇用後、教官の指導のもとバイオインフォマティクスのスキル習得を行い、彼らの本来の専門の分野を生かしながら、この分野の研究を進めている。

情報系の専門で学位を取得した博士研究員は、自らが専門とする手法やアルゴリズムの生物学の研究課題への適用、あるいはそのためのさらなる手法・理論の開発を目指す研究を行っている。例えば、統計的学習理論を用いた遺伝子発現解析法の研究、土居グループの博士研究員が行っている教師あり分類器（サポートベクターマシン）を応用した生物学の文献から蛋白質名の抽出などの研究がこれに相当する。これらは、「アルゴリズム志向」あるいは「手法開発志向」の研究であり、バイオインフォマティクスの発展には情報処理

技術の革新が不可欠であるため、重要な研究の方向性である。もちろん、そのためには対象となるデータの理解が不可欠であり、情報系出身の研究者は、本ユニットで生物学の基本知識の学習を行っている。

一方、化学、生化学、免疫学で学位を取得した研究者には、既存の標準的なバイオインフォマティクス的手法を生物学データに適用し、その幅広い生物学的知識を生かし、そこから生物学的な意味を読み取る研究を行っている。例えば、化学系の研究者が行った「ゲノムからの新しい型の銅結合蛋白質の発見とその進化系統解析に関する研究」は、標準的なデータベースと配列解析ソフトウェアを使いながらも、無機化学に関する豊富な知識から、銅結合蛋白質に対する新たな知見が得られている。また、生化学系の背景を持つ研究者が行った「立体構造の大域的類似と酵素の化学反応に関する研究」では、抽象的な酵素の立体構造比較の結果に、豊富な生化学的知識を生かした考察を加えることで、独自性の高い研究となった。これらは「マテリアル志向」あるいは「データマイニング志向」の研究であり、新たな手法の開発よりはデータから生物学に興味深い事実を読み取るのが主たる関心になる。もちろん、標準的なソフトウェアの操作も、実験系の生物学者にとっては困難である場合が多く、その教育においては情報系の研究者も大きく貢献した。

これら「アルゴリズム志向」、「データマイニング志向」の二通りのアプローチはどちらも重要であり、幅広い背景の博士研究員を雇うことで、お互いに刺激しあい協力しあうような研究環境が実現しつつあるといえる。

現在在職中の博士研究員の一人は、バイオインフォマティクス分野の研究実績を評価され、近隣の大学からの要請を受けて、毎土曜日に非常勤講師としてバイオインフォマティクスの講義を担当している。

【想定外の成果、困難について】

当初の計画では、人材養成ユニットの教官で構成される「蛋白質機能予測学講座」において5年間で修士取得者20名、博士取得者5名を養成することを目標とした。しかし、実際には、本講座に大学院学生を集めることは予想外に難しいことがわかった。理由の一つとして、新設の研究室は先輩学生等が不在なため、一般に学生が敬遠する傾向があることがあげられる。もう一つの、より大きな理由として、特に後期課程への進学可能性を考える前期課程学生にとって、期限付きの研究室には不安を覚えることが挙げられる。通常、博士前期、後期を通じた教育には5年が必要だが、本ユニットには形式上4年の教育期間しか確保できていない印象がある。本大学院大学としては、そうした形式にとらわれず、課程途中のユニット終了で所属学生が不利益をこうむることのない方策を採ることになっているが、十分に不安感を払拭しきれていない。一方で、本講座に所属している前期課程2年の一人は、来年度からの後期課程への進学を決めている。質の高い研究をアクティブに進め、実質上で魅力のある研究室運営を行うことで、より多くの博士後期課程学生を確保する努力を継続する。