

21世紀COEプログラム 平成16年度採択拠点事業結果報告書

機 関 名	東京大学	学長名	濱田 純一	拠点番号	K07	
1. 申請分野	K〈革新的な学術分野〉					
2. 拠点のプログラム名称 (英訳名)	言語から読み解くゲノムと生命システム—次世代バイオインフォマティクス拠点の創成— (Elucidation of Language Structure and Semantics behind Genome and Life Systems)					
研究分野及びキーワード	〈研究分野: 情報学〉 (バイオインフォマティクス) (システム生物学) (ゲノム科学) (計算言語学) (知能情報学)					
3. 専攻等名	大学院新領域創成科学研究科情報生命科学専攻、メディカルゲノム専攻、先端生命科学専攻、社会文化環境学専攻 (平成19年4月1日環境学専攻より改組)、大学院工学系研究科電気系工学専攻 (平成20年4月1日基盤情報学専攻から改組)					
4. 事業推進担当者	計 12 名					
ふりがなくローマ字) 氏 名	所属部局(専攻等)・職名	現在の専門 学 位	役割分担 (事業実施期間中の拠点形成計画における分担事項)			
(拠点リーダー) TAKAGI TOSHIHISA 高木 利久	大学院新領域創成科学研究科 情報生命科学専攻・教授	バイオインフォマティクス 工学博士	全体の総括、テキストからの知識抽出法及び生体内ネットワーク記述言語の開発、ゲノム辞書構築			
ASAИ KИYOSHИ 浅井 潔	大学院新領域創成科学研究科 情報生命科学専攻・教授	ゲノム情報学 工学博士	ゲノム配列、RNA配列の文法解析、比較ゲノムのための計算モデル構築			
ARITA MASANORI 有田 正規	大学院新領域創成科学研究科 情報生命科学専攻・准教授	システム生物学 理学博士	相互作用マップの構築、生体内ネットワークの特徴抽出とその産業応用			
ITO TAKASHI 伊藤 隆司	大学院新領域創成科学研究科 情報生命科学専攻・教授	ゲノム科学 分子生物学 医学博士	分子間相互作用の定量化技術の開発、相互作用の規則性抽出			
NAKAYA AKIHIRO 中谷 明弘	大学院新領域創成科学研究科 情報生命科学専攻・准教授	情報生命科学 理学博士	異種データからの知識発見、生体内ネットワークの可視化			
MORISHITA SHINICHI 森下 真一	大学院新領域創成科学研究科 情報生命科学専攻・教授	情報生命科学 理学博士	ゲノム配列からの特徴抽出、機能ゲノム産業応用のためのプライマー、チップ等の設計			
KOBAYASHI ICHIZOU 小林 一三	大学院新領域創成科学研究科 メディカルゲノム専攻・教授	ゲノム進化学 薬学博士	ゲノムテキストの維持と再編の解析			
SUGANO SUMIO 菅野 純夫	大学院新領域創成科学研究科 メディカルゲノム専攻・教授	ゲノム医科学 医学博士	遺伝子発現の定量解析技術の開発、発現制御の文法解読			
FUJIIWARA HARUHIKO 藤原 晴彦	大学院新領域創成科学研究科 先端生命科学専攻・教授	分子生物学 ゲノム解析学 理学博士	ゲノム言語の誕生と進化の解析、昆虫ゲノムの産業応用			
YAMAMOTO KAZUO 山本 一夫	大学院新領域創成科学研究科 先端生命科学専攻・教授	分子生物学・ゲノム創薬 薬学博士	糖鎖解析、創薬への応用			
MINO TAKASHI 味埜 俊	大学院新領域創成科学研究科 社会文化環境学専攻・教授	環境微生物工学 工学博士	微生物ゲノムを利用した環境監視と保全			
IBA HITOSHI 伊庭 斉志	大学院工学系研究科 電気系工学専攻・教授	進化型システム 工学博士	遺伝子制御ネットワーク解析、進化的最適化手法の応用			
5. 交付経費 (単位: 千円) 千円未満は切り捨てる ( ) : 間接経費						
年 度(平成)	16	17	18	19	20	合 計
交付金額(千円)	130,000	96,000	90,680	86,000 (8,600)	87,000 (8,700)	489,680  (17,300)

## 6. 拠点形成の目的

ゲノムおよびそれが作り出す生命システムを解明し産業に役立てるには、従来の分子生物学的手法を適用するだけでは不十分であり、情報技術を駆使して膨大なデータを解析することが不可欠である。このような研究は一般にバイオインフォマティクス (BI) と呼ばれ、近年大きな注目を集めているが、ポストゲノム時代を迎え、BIの教育研究の重要性はますます高まっている。BIは、これまでどちらかと言えば、ゲノム科学・生命科学を計算機で支援するための脇役的学問分野としての色彩が強かったが、ここにきてシステム生物学という言葉にも代表されるように、生命を従来のように物質的存在としてとらえるのではなく情報的存在として、すなわち、情報システムとしてとらえることの重要性が強く認識されるようになり、また、バイオ産業に大幅な効率化をもたらすことが明らかとなり、BIは生命科学、バイオ産業の将来を担う中核的な存在になりつつある。

しかし、残念なことに、欧米には優に100を越えるBI教育研究拠点が形成されているにもかかわらず、我が国では本格的なBI教育研究拠点と呼べるものはこれまでほとんどなかった。BI拠点と名乗るものもあっても、その多くは、教員が少ない、BIの専門家がほとんどいない、研究だけで教育の体制が併設されていない、生物系か情報系どちらかに偏った組織である、ゲノム研究の進展に密接に連結した先端的なものでない、などの理由から本格的な拠点とは呼び難い。そのため、BI分野の人材不足は深刻で、その人材養成が総合科学技術会議、内閣府IT戦略会議等で最重要課題に取り上げられるまでになっている。

東京大学ではこのような状況を打破すべく平成15年度より大学院新領域創成科学研究科に情報生命科学専攻を、また、16年度よりメディカルゲノム専攻を設置し、その社会的な要請に応えようとしてきた。本拠点形成プログラムはこの動きをさらに拡大し加速させるためのものである。本プログラムでは現在緊急に養成が叫ばれているBI人材を社会に送り出すことはもちろんのこと、将来の生命科学、バイオ産業を担う人材の養成とそのための研究を目的とする。これまでBI研究は、どちらかと言えば、ゲノム配列などの1次元データの表層的な処理にとどまっていた。これからは機能ゲノム学やシステム生物学などの言葉に代表されるように、生命機能そのものを表す高次の情

報の解析に切り込んで行かなければならない。本プログラムで提案する研究教育はそのためのものである。上述したように、我が国にはBIを先端的な科学として教育研究する本格的な拠点はほとんどない。仮に近いものがあるとしても、それらの多くはバイオに特化した情報技術 (リテラシー) の習得に力点を置いたものであるか、もしくは、情報系あるいは生命系に偏った組織である。これからのBI教育で重要なことは与えられた問題を効率よく解くだけでなく、自分で問題を見つけてそれを定式化し、さらに、その解決に必要なデータ収集を自分で行ったり、そのための実験をデザインしたりできる能力をもった人材の育成である。そのためには、生命科学と情報科学を基礎からきちんと学べるようなカリキュラムの提供と、それを教えるための、情報系生命系が融合し一体となった体制が必要である。本拠点形成プログラムでの教育はまさにこれを目指している。

一方、研究面においては、本プログラムは、これまでのものよりも意味的に深い解析を目指す。具体的には、ゲノム配列はどのような文法で書かれているか、ゲノム配列の言語的な特徴 (冗長性、多義性、曖昧性など) は何か、遺伝子やタンパク質はどのような文脈で働くのか、代謝やシグナル伝達のような生体ネットワークはどのような文法に支配されているのか、また、これらのネットワークを計算機の中に表現しシミュレーションによりその振る舞いを明らかにするにはどのような計算機記述言語が必要なのか等々、ゲノムと生命システムに内在する言語的な側面を明らかにすることを目指す。また、遺伝子発現や分子間相互作用などのデータの解釈には文献に書かれた知識の抽出と体系化が重要であるが、このための言語処理も本研究の目指すところである。これら「言語」の意味するところは[ゲノムを規定している言語]、[生命システムに内在する言語]、[生命システムの記述とシミュレーションのための計算機言語]、[生命学者が知識を記述するための自然言語]とそれぞれ異なるが、相互に密接に関連しており、本プログラムではこれらを情報系、生命系の研究者が融合し一体となった体制を構築し、総合的に研究する。

## 7. 研究実施計画

本プログラムでは、ゲノムと生命システムに内在する言語構造を読み解くという共通の目標のもとに、生命系情報系が連携・融合し、研究を実施する。ここではおもに以下の4つの言語解明の視点からそれぞれ研究を展開する。

### ・ゲノムを規定している言語の解明

ゲノム配列はATGCの4種類のアルファベットからなる1次元の文字列データであり、まさに言語と呼ぶのに相応しい特徴を備えている。これまでその言語構造の解明に多くの努力が払われてきたが、まだ全貌の解明には至っていない。本プログラムでは、言語理論、確率文法、高速文字列照合、機械学習、並列処理などの先進的な情報処理技術を駆使して、また、これとゲノムの進化的解析とを統合して、この問題にアプローチする。具体的には、膨大なゲノム配列断片をアセンブルして1本の配列にするための技術開発、異なる生物のゲノムを比較したりmRNAのデータを網羅的に集めて解析したりすることによる遺伝子制御領域の解明、DNAチップに利用できるゲノム配列断片の設計、機能をもったRNA配列の発見手法、などの研究を行う。

### ・生命システムに内在している言語的構造の解明

複雑で多様な生命現象は、いくつもの遺伝子やタンパク質の相互作用がおりなすネットワークにより実現されている。これを生命システムと呼ぶ。このネットワークの解明こそが、ポストゲノム時代の中心的な課題である。この問題にアプローチするには、まず、タンパク質間相互作用やタンパク質DNA相互作用の網羅的かつ定量的なデータ収集とそのための高精度な計測技術開発が必要である。また、遺伝子の系統的条件的な破壊などにより細胞形態や体形成などの表現型（画像データとして表される）にどのような変化が起きるかを調べることも必要である。これと並行して、ゲノム配列データからこれらの相互作用を予測する、相互作用データから全体のネットワークを推定する、などの情報技術の開発も欠かせない。このような研究はまだ方法論が確立していないが、本プログラムでは、実験系、情報系一体となった体制で研究を進める。

### ・生命学者が知識を記述する自然言語の解析

上記の生命システムの解明には、これまで生命学者が生命現象をどのようにとらえ記述してきたかを

調べるが大いに参考になる。システムや機能にどの程度の階層性やモジュール性などがあるのか、それが論文や教科書の中にどのように記述されているのかを調べ、それを取り出し計算機に辞書や概念階層の形で格納することができれば、上記の相互作用やネットワークとつぎ合わせることにより、生命システムの解明に役立てられる。生命科学は物理学などの数式で表現される学問とは異なり、自然言語や図、絵で記述することにより完結するという側面をもっている。しかしながら、これまで生物種毎や研究コミュニティ毎にばらばらに知識の記述がなされてきた。本プログラムではこれらを自動的に抽出し、かつ、統一的に扱うための技術開発を自然言語処理、知識処理の観点から展開する。

### ・生命システムを記述しシミュレートするための計算機言語の開発

ゲノムがどのようにして生命システムを規定しているのか、我々人間が生命を理解するとはどういうことか、などの生命の謎に迫るには、上記の3つの言語（ゲノムの言語、生命システムの言語、生命科学者の言語）の間に成り立つ関係を明らかにしなければならない。そのためには、この3つの言語に関するさまざまな情報や知識を計算機に入れ、それらがどのような関係になっているか、どのようにして冗長性、多義性、階層性などを生み出しているかを調べる必要がある。ゲノム配列のような1次元の文字列データであれば、計算機に入れるのは容易であるが、ネットワーク、生命現象や機能、などとなるとそのための計算機記述言語とその解析技術を新たに考案する必要がある。本プログラムでは、グラフ理論、知識表現技術、並列プロセス記述言語、推論技法などを駆使して、この問題にアプローチする。

なお、このような4つの研究の成果は論文にするだけでなく、ソフトウェアやデータベースの形で世界に発信する予定である。

## 8. 教育実施計画

本プログラムでは、将来の生命科学、BI、バイオ産業を担う人材（情報系、生命系両方のセンスを身に付けた）の養成・輩出を目指して、大学院教育において、生命科学、情報科学、BIをそれぞれ基礎から、かつ、系統的に教育する仕組みを実現する。現時点では、学部レベルのBI教育体制はどの大学でも整備されていない。そこで、生物系、情報系、物理系、工学系を問わず、どのような教育的背景をもつ学生でも本プログラムに参画して学べるように入試やカリキュラムを工夫する。なお、BI教育では、単に講義を聴くだけでなく、計算機実習、実験演習を通してその内容を体得することがとくに重要であるが、これを特任助教、RA、TAの採用により格段に充実させる。

本プログラムでは、BI分野ですでに確立した配列解析技術やゲノムデータベース技術などはもちろんのこと、情動的・言語的な視点で生命をとらえ、それを解明するための情報技術とその基礎理論の教育を行う。具体的には、言語理論、確率モデル、知識処理、自然言語処理、文献検索、データマイニング、機械学習、グラフ理論、アルゴリズム設計およびこれらを基盤とした遺伝子多型解析、遺伝子発現情報解析、相互作用解析、システム生物学、生命情報表現論、パスウェイ解析、オントロジー構築などの講義を行う。これらはおもに情報生命科学専攻が担当する。

一方、前述したように、これからのBI教育で重要なことは与えられた問題を計算機で効率よく解くだけでなく、自分で問題を見つけそれを定式化し、さらに、その解決に必要なデータ収集を自分で行ったり、そのための実験をデザインしたりできる能力をもった人材の育成である。そのためには、情報技術やその基礎理論の習得だけでは不十分である。分子生物学、細胞生物学はもちろんのこと最先端のゲノム科学、タンパク質科学、発生学、進化学やそれを支える計測技術（具体的には、遺伝子発現解析、タンパク質相互作用解析、糖鎖解析など）の教育が必要である。本プログラムではこれらの教育にも力を入れる。なお、これらはおもにメディカルゲノム専攻や先端生命科学専攻が担当する。

このような基礎的な教育の他に、以下のような取り組みをあわせて実施する。

- ・世界で通用する研究者を育成するために、学生主導の国際ワークショップを企画・実施する。
- ・最先端の研究に触れさせるために、第一線の研究者による特別講義や特別セミナーを実施する。
- ・最先端の研究プロジェクトに参加させ、実践力を身につけさせるために、東京大学以外のバイオインフォマティクスの研究機関とも連携した研究教育体制を構築し、そこでの研究プロジェクト（多くは国際共同研究）に積極的に参加させる。

最先端の研究に従事していて、かつ、多様な経歴や教育的背景をもつ教員集団を研究科や大学の枠を超えて組織することにより、最先端のBIおよび関連分野の教育を実施する。それにより、情報系、生命系両方のセンスを身に付けた新しいタイプの学生を5年間で100人程度産業界に輩出することを目指す。また、日本から革新的なBI技術を世界に向けて発信できる能力を備えた研究者を30人程度養成することを目指す。そのために博士課程の学生だけでなく、特任研究員や企業からの研究員も積極的に受け入れ、最先端の教育研究環境を提供する。なお、そのための基本的な計算機設備や実験設備はすでにほぼ整っている。また、今後必要となるものは競争的資金の獲得により導入を図る予定である。

## 9. 研究教育拠点形成活動実績

### ①目的の達成状況

#### 1) 世界最高水準の研究教育拠点形成計画全体の目的達成度

研究面でも教育面でも本プログラムで当初計画した目的は十分に達成できた。これにより、我が国におけるバイオインフォマティクスの研究教育拠点としてはトップの位置を占める拠点になることができた。ただ、世界的には、この分野の進展・拡大は激しく、本プログラムにおいていくつもの世界的成果や何人もの世界に通用する人材が出てはいるが、教員数、博士課程の学生数やポスト数、教育プログラムの充実度、計算機やゲノム解読装置などの設備において若干不足するところがあり、世界有数の研究教育拠点と呼べるまでには至っていない。これは今後の課題とした。

本プログラムの具体的な成果としては、研究面では、ゲノム解読やそのための情報技術開発、分子間相互作用の網羅的取得法、機能性RNA配列の発見手法、などにおいていくつもの世界的成果が出ている。また、いち早く次世代の超高速ゲノム解読装置を導入し、これを活用した研究を我が国で先駆けて立ち上げた。今後、生命科学が次世代超高速ゲノム解読装置の活用を中心に展開して行くことを考慮すると、これは特筆すべきことである。

教育面では、とくに奇をてらったことはしていないが、基礎的な教育や演習を地道に実施するとともに、学生主導の国際ワークショップを企画したり、国際共同研究に学生を参加させるなどの教育研究活動を充実させたりすることにより、世界に通用する若手研究者や産業界で活躍できる人材を何人も育成することができた。

これらにより、本プログラムの目的は十分に達成したと考えている。

#### 2) 人材育成面での成果と拠点形成への寄与

バイオインフォマティクスの研究者・技術者を育てるには、単にツールやデータベースの使い方を教えるだけでは十分ではない。生命科学、情報科学の基礎を系統立って教える必要がある。また、そのための演習（プログラミングなど）を充実させることが不可欠で

ある。本プログラムでは、この点に十分な配慮をした基礎的教育を実施した。これ以外にも第一線で活躍できる研究者、世界に通用する人材を育てるために、特色ある教育プログラムをいくつか実施した。その一部を以下に紹介する。

- ・本COEプログラム事業推進担当者も含め、最先端の研究者による特別講義やセミナーを毎年十数回開催。
- ・TV 会議システムを活用した、外部研究機関における最先端の研究セミナーの配信を実施。
- ・毎年40名から70名程度のリサーチアシスタント(RA)を博士課程学生から選抜。RAによる研究成果発表会・交流会を毎年1回開催。
- ・この分野の重要性と面白さを広く認知してもらうための公開シンポジウムを関東と関西で年2回開催。
- ・学生に国際会議の運営方法を伝授し、会議のプログラム策定から招待講演者の選定、投稿される発表アブストラクトの査読までを学生が主体となって行う日本・台湾の国際ワークショップを2回開催した。
- ・この分野の最先端の成果の普及教育のための実験医学(羊土社)増刊「バイオデータベースとソフトウェア最前線」の企画、出版。
- ・ゲノム解析アルゴリズムに関する英語の教科書を出版。

これらの取り組みにより21世紀COE で雇用した特任助教およびRAが第一著者でScienceおよびNatureに論文を報告しており、世界トップレベルの研究者を養成することに成功した。本プログラムの当初の計画目標であった、バイオインフォマティクスに関して100人規模の技術者・開発者および30人規模の研究者を養成できた(現在博士課程在学中のものやポストも含む)。

なお雇用した特任助教の中からは、准教授や講師に昇格して他大学や本学においてPI(研究室の責任者)として独立したものが3名出ており、この分野の発展を支える人材供給源としても本拠点は重要な役割を果たした。

### 3) 研究活動面での新たな分野の創成や、学術的知見等

- ・我が国で最も早期に次世代の超高速ゲノム解読装置を導入し、その活用法をいくつも開発した。また、そこから産出される膨大なデータを超並列計算機により解析する技術を開発した。これらは、次世代の超高速ゲノム装置の活用を軸に展開される、新たな生命科学の進展に大きく寄与するものと期待される。
- ・言語という視点で統一的融合的に研究を展開することにより、膨大なデータを知識体系化するための新たなバイオインフォマティクス分野を開拓した。

これらの他にも、文献からの生体内パスウェイ図の取得法の開発、ゲノム進化過程の再構築手法の開発とそれによるパスウェイ進化過程の解明、タンパク質非コードRNA探索手法の開発、代謝に関する情報の共有のための情報基盤開発、分子間相互作用定量化のための基盤技術開発、相互作用の規則性抽出、遺伝的アルゴリズムによる遺伝子ネットワーク推定法の開発、近縁複数ゲノム配列比較によるゲノム再編過程の再構築、新しいフォールド（基本立体構造）をとるDNA結合タンパク質の発見、小麦胚芽由来の無細胞タンパク質発現系で得た試料におけるタンパク質の結晶立体構造決定、プロモータ部位の同定、次世代シーケンサーを利用したゲノムワイドな転写開始点マッピング法の開発、体内の脂肪などの量的表現型情報のモデル化と遺伝子座探索手法の開発、レトロトランスポゾンLINEの転移機構の解明とその応用、ゲノム情報を利用した環境監視、環境保存法の開発、ゲノムアセンブル法の開発とそれを用いたメダカゲノム解読、脊椎動物の染色体進化解明、クロマチン構造の進化への影響解明、レクチンに関する種々の機能解明、などに関して研究の進展があった。なお、これらの成果は世界的な一流雑誌に論文の形で出版するだけでなく、その中のいくつかはデータベースやソフトウェアの形で公開している。これにはインターネットを介して世界中から多くのアクセスがある。

### 4) 事業推進担当者相互の有機的連携

本プログラムにおいては事業推進担当者およびその研究室に所属する学生の間では日常的に共同研究や緊密な連携が行われている。以下はその中で、成果としてまとめたものを紹介する。

- ・メダカ、カイコを中心にしたゲノムのアノテーションや進化解析の推進および次世代超高速ゲノム解読装置を活用した転写解析の方法論開発（森下・菅野、2007-2008年だけでもNature Scienceを含む5件の論文）
- ・出芽酵母の完全長cDNAデータベースをゲノムブラウザー上に構築し、遺伝子の再アノテーションを行い公開した（伊藤・森下 2006年 PNAS）
- ・次世代シーケンサーによるDNAメチル化の網羅的高分解能解析技術を開発し、アカパンカビゲノムのメチル化に関するデータベースを公開した（伊藤・森下）

### 5) 国際競争力ある大学づくりへの貢献度

- ・国際競争力があるとは、著名な国際誌に論文が掲載されること、データベース等に海外から多くのアクセスがあることに加え、国際共同研究の申し込みがあることなどが挙げられよう。発表論文は別途記載したので、ここでは国際共同研究等の活動を紹介する。
- ・本COEプログラムの活動（具体的には学生主導の日本台湾国際ワークショップ）を母体として結成されたバイオインフォマティクス若手の会は、バイオインフォマティクス業界で最大の国際組織ISCBIにおける下部組織として認定され、若手を中心にした活動を続けている。
- ・事業推進担当者の森下を中心に、スタンフォード大学、北京ゲノム研究所、英国サンガー研究所等とゲノム解読、進化、クロマチン構造解析で現在共同研究中。毎年平均5件の国際会議運営に携わっている。
- ・浅井は、麹菌ゲノム解析でMIT等の米国の主要なゲノム研究機関と共同研究を行った。
- ・伊藤は、国際的システム生物学グループYSBNのメンバーで、ワークショップの講師として若手研究者を育成している。
- ・中谷は英国ウェールズ大学とロボットを使った表現型画像解析のための完全自動化システムの共同研究を開始している。
- ・菅野は、全長cDNA解読やHUGOで様々な国際共同研究をまとめた。現在はメキシコ自治大学（メキシ

コノ東大)と共同で、有コウ条虫のゲノム解読と遺伝子分析を進めている。

- ・高木は、英国マンチェスター大学付属のテキストマイニングセンターとともに生物医学文献を自動的に解析し辞書とコーパス等の知識ベースを構築する共同研究を行っている。
- ・伊庭は、韓国・ソウル大学およびイギリス・ロンドン大学の研究グループと協力して遺伝子ネットワーク解析システムの開発にとりこんでいる。
- ・藤原は、トリボリウム、カイコなどの国際ゲノム解読チームの一員としてゲノム解読・遺伝子解析を成功させ、さらにアゲハなどの鱗翅目ゲノムについて共同研究を進めつつある。
- ・味埜は、中国科学院・北京工業大学と連携して排水処理にかかわる微生物群集構造の解析に関する共同研究を進めつつある。
- ・小林は、ヨーロッパのタンパク質立体構造予測のグループとの共同研究で、新しい立体構造(ハーフパイプ)をとるDNA結合タンパク質の発見に成功した。アメリカのグループとも共同でこの方向の研究を進めている。ヨーロッパの若手の研究者の養成システムであるMarie Curie Research Training Network “DNA Enzymes”では、講師を務めた。

#### 6) 国内外に向けた情報発信

特筆すべきことではないが、日本語と英語によるホームページを開設するとともに、年2回(関東、関西)の公開シンポジウム等を開催し、本プログラムの紹介、成果公開に務めた。これ以外にも、前記の国際共同研究の際や個々の研究発表の場で、本プログラムの成果を積極的に紹介した。

#### 7) 拠点形成費等補助金の使途について(拠点形成のため効果的に使用されたか)

本プログラムの経費の7割以上をRAの雇用、特任助教などの人件費にあてることにより、人材育成に重点をおいてプログラムを運営してきた。最先端の研究に触れさせたり、研究プロジェクトに参加させたりするために必要な大型測定装置や大規模計算機は外部資金もしくは東京大学からの支援により手当てした。これにより拠点形成費等補助金を本プログラムの一番の目的である人材育成に効果的に活用できた。

#### ②今後の展望

本プログラムで構築した、研究科や大学の枠を超えた融合型教育研究体制(理学部生物情報科学科、経済産業省産業技術総合研究所生命情報科学研究センターおよび理化学研究所ゲノム科学総合研究センター(平成20年3月末に解散、その後はその後継部門と連携)との連携体制)は今後も維持する。また、融合研究のさらなる促進、次世代の超高速ゲノム解読装置の活用、国内外の共同研究促進のために平成20年度に設置したオーミクス情報センターの強化を図る。そのための外部資金の獲得を目指す。

また、これまでの体制に、超大容量データ解析のための並列処理技術面等の教育研究体制を加えて、グローバルCOE拠点の形成を目指す。

これに加え、これまで行ってきた国内外の共同研究、ワークショップ(台湾日本国際ワークショップなど)を継続する。さらに北京ゲノム研究所等との共同シンポジウム開催を企画予定である。

#### ③その他(世界的な研究教育拠点の形成が学内外に与えた影響度)

- ・本プログラムにより、バイオインフォマティクスの重要性、有用性が東京大学においても広く認知されるようになり、東京大学理学部にバイオインフォマティクスを教育する生物情報科学科が平成19年度より設置された。この学科のバイオインフォマティクス教育は、本プログラムの中核メンバーである情報生命科学専攻の教員が中心となって実施している。
- ・細分化された知識の統合と俯瞰を目指した東大総長直轄プロジェクト「学術統合化プロジェクト(ヒトと地球)」が平成17年度より6年計画で立ち上がったが、そのときすでに本COEプログラムが高木を拠点リーダーに立ち上がっていたことより、学術統合化プロジェクト(ヒト)のリーダーも、高木が務めることとなった。これにより、両者のプロジェクトの連携が図られている。
- ・本プログラムで構築した融合型研究教育体制が評価され、これを強化するためにオーミクス情報センターが平成20年度に設置された。東京大学より、教授1名、超高速ゲノム解読装置1台が措置されている。

21世紀COEプログラム 平成16年度採択拠点事業結果報告書

機 関 名	東京大学			拠点番号	K07
拠点のプログラム名称	言語から読み解くゲノムと生命システム —次世代バイオインフォマティクス拠点の創成—				
<p>1. 研究活動実績</p> <p>①この拠点形成計画に関連した主な発表論文名・著書名【公表】</p> <div style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <ul style="list-style-type: none"> <li>・事業推進担当者（拠点リーダーを含む）が事業実施期間中に既に発表したこの拠点形成計画に関連した主な論文等〔著書、公刊論文、学術雑誌、その他当該プログラムにおいて公刊したもの〕</li> <li>・本拠点形成計画の成果で、DP（ディスカッション・ペーパー）、Web等の形式で公開されているものなど速報性のあるもの</li> <li>※著者名（全員）、論文名、著書名、学会誌名、巻(号)、最初と最後の頁、発表年（西暦）の順に記入</li> <li>波下線（<u>          </u>）：拠点からコピーが提出されている論文</li> <li>下線（<u>          </u>）：拠点を形成する専攻等に所属し、拠点の研究活動に参加している博士課程後期学生</li> </ul> </div> <ul style="list-style-type: none"> <li>・Koike, A. and Takagi, T. :Classifying biomedical figures using combination of bag of keypoints and bag of keywords, Proc. of 2nd Int. Workshop on Intell. Inform. in Biol. and Med., 848-853 (2009).</li> <li>・Iwasaki, W. and Takagi, T. :Rapid Pathway Evolution Facilitated by Horizontal Gene Transfers across Prokaryotic Lineages, PLoS Genetics, 5(3), e1000402 (2009).</li> <li>・Ishii, N., Koike, A., Yamamoto, Y. and Takagi, T. :Figure Classification in Biomedical Literature towards Figure Mining, 2008 IEEE International Conference on Bioinformatics and Biomedicine, 263-269 (2008).</li> <li>・Iwasaki, W. and Takagi, T. :Reconstruction of highly heterogeneous gene-content evolution across the three domains of life, Bioinformatics, 23 (13), i230-9 (2007).</li> <li>・Noguchi, H., Park, J. and Takagi, T. :MetaGene: prokaryotic gene finding from environmental genome shotgun sequences, Nucleic Acids Res., 34 (19), 5623-30 (2006).</li> <li>・Machida, M., Asai, K. et al. : Genome sequencing and analysis of Aspergillus oryzae. Nature, 438, 1157-1161 (2005)</li> <li>・Tabei, Y., Tsuda, K., Kin, T. and Asai, K. : SCARNA: fast and accurate structural alignment of RNA sequences by matching fixed-length stem fragments. Bioinformatics, 22(14), 1723-1729 (2006).</li> <li>・Okada, K. and Asai, K. : Retention of genes involved in the adenylation-mediated endocrine system in early vertebrates. Gene, 412, 71-83 (2008).</li> <li>・Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T. N., Siomi, M. C. and Siomi, H. : Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature, 453, 793-797 (2008)</li> <li>・Asai, K., Kiryu, H., Hamada, M., Tabei, Y., Sato, K., Matsui, H., Sakakibara, Y., Terai, G. and Mituyama, T. : Software.ncrna.org: web servers for analyses of RNA sequences, Nucleic Acids Research, 36, W75-W78 (2008).</li> <li>・Kusano, M., Fukushima, A., Arita, M., Jonsson, P., Moritz, T., Kobayashi, M., Hayashi, N., Tohge, T. and Saito, K. : Unbiased characterization of genotype-dependent metabolic regulations by metabolomic approach in Arabidopsis thaliana BMC Systems Biology, 1, 53 (2007).</li> <li>・Nagasaki, H., Arita, M., Nishizawa, T., Suwa, M. and Gotoh, O. : Automated classification of alternative splicing and transcriptional initiation and construction of visual database of classified patterns Bioinformatics, 22(10), 1211-1216 (2006).</li> <li>・Nagasaki, H., Arita, M., Nishizawa, T., Suwa, M. and Gotoh, O. : Species-specific variation of alternative splicing and transcriptional initiation in six eukaryotes. Gene, 364, 53-62 (2005).</li> <li>・Kito, K., Kawaguchi, N., Okada, S., Ito, T. : Discrimination between stable and dynamic components of protein complexes by means of quantitative proteomics. Proteomics 8(12), 2366-2370, (2008).</li> <li>・Kito, K., Ota, K., Fujita, T., Ito, T. :A synthetic protein approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. J. Proteome Res. 6(2), 792-800, (2007).</li> <li>・Miura, F., Kawaguchi, N., Sese, J., Toyoda, A., Hattori, M., Morishita, S., Ito, T. :A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc. Natl. Acad. Sci. USA 103(47), 17846-17851, (2006).</li> <li>・Miura, F., Uematsu, C., Sakaki, Y., Ito, T. :A novel strategy to design highly specific PCR primers based on the stability and uniqueness of 3' -end subsequences. Bioinformatics 21(24), 4363-4370, (2005).</li> <li>・Isobe, S., Nakaya, A., Tabata, S. :Genotype Matrix Mapping (GMM): Searching for QTL interactions in genetic variation in complex traits. DNA Res. 14(5), 217-225, (2007).</li> <li>・Watanabe, A., Toyota, T., Owada, Y., Hayashi, T., Iwayama, Y., Matsumata, M., Ishitsuka, Y., Nakaya, A., Maekawa, M., Ohnishi, T., Arai, R., Sakurai, K., Yamada, K., Kondo, H., Hashimoto, K., Osumi, N., Yoshikawa, T. : Fabp7 maps to a quantitative trait locus for a schizophrenia endophenotype. PLoS Biol. 5(11), 2469-2483, (2007).</li> <li>・Wei Qu, Hashimoto, S. Morishita, S. :Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing. Genome Research (in press).</li> <li>・Sasaki, S., Mello CC, Shimada, A., Nakatani, Y., Hashimoto, S., Ogawa, M., Matsushima, K., Gu SG, Kasahara, M., Ahsan, B., Sasaki, A., Saito, T., Suzuki, Y., Sugano, S., Kohara, Y., Takeda, H., Fire, A., Morishita, S. : Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. Science 323(5912), 401-404 (2009).</li> <li>・Nakatani Y, Takeda H, Kohara Y, Morishita S. :Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates," Genome Research 17(9), 1254-1265 (2007).</li> <li>・Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., Shimizu, N., Hashimoto, S., Yang, J., Lee, Y., Matsushima, K., Sugano, S., Sakaizumi, M., Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K., Morishita, T., Endo T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y. :The medaka draft genome and insights into vertebrate genome evolution," Nature 447(7145), 714-719 (2007).</li> <li>・Kasahara, M., Morishita, S. : Large-scale genome sequence processing. Imperial College Press, 248pp. (2006).</li> <li>・Ishikawa, K., Handa, N., Kobayashi, I. : Cleavage of a model DNA replication fork by a Type I restriction endonuclease, Nucleic Acids Research, Advance access published.</li> </ul>					



- Fukuda, E., Kaminska, K.H., Bujnicki, J.M., Kobayashi, I.: Cell death upon epigenetic genome methylation: a novel function of methyl-specific deoxyribonucleases. *Genome Biology*, 9:R163 (2008).
- Tsuru, T., Kobayashi, I.: Multiple genome comparison within a bacterial species reveals a unit of evolution spanning two adjacent genes in a tandem paralog cluster. *Molecular biology and evolution*, 25: 2457–2473 (2008).
- Miyazono, K., Watanabe, M., Kosinski, J., Ishikawa, K., Kamo, M., Sawasaki, T., Nagata, K., Bujnicki, J.M., Endo, Y., Masaru Tanokura, M., Kobayashi, I.: Novel protein fold discovered in the PabI family of restriction enzymes. *Nucleic Acids Research*, 35: 1908–1918 (2007).
- Mochizuki, A., Yahara, K., Kobayashi, I., Iwasa, Y.: Genetic addiction: selfish gene's strategy for symbiosis in the genome. *Genetics*, 172: 1309–1323 (2006).
- Hashimoto S, Qu W, Ahsan B, Ogoshi K, Sasaki A, Nakatani Y, Lee Y, Ogawa M, Ametani A, Suzuki Y, Sugano S, Lee CC, Nutter RC, Morishita S, Matsushima K. High-resolution analysis of the 5'-end transcriptome using a next generation DNA sequencer. *PLoS ONE*. 2009;4(1):e4108. Epub 2009 Jan 1.
- Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang TH. The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* 24:167–177, (2008).
- Tsuritani K, Irie T, Yamashita R, Sakakibara Y, Wakaguri H, Kanai A, Mizushima-Sugano J, Sugano S, Nakai K, Suzuki Y. :Distinct class of putative "non-conserved" promoters in humans: Comparative studies of alternative promoters of human and mouse genes. *Genome Res.* 17: 1005–1014, (2007).
- Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, Yamashita R, Yamamoto J, Sekine M, Tsuritani K, Wakaguri H, Ishii S, Sugiyama T, Saito K, Isono Y, Irie R, Kushida N, Yoneyama T, Otsuka R, Kanda K, Yokoi T, Kondo H, Wagatsuma M, Murakawa K, Ishida S, Ishibashi T, Takahashi-Fujii A, Tanase T, Nagai K, Kikuchi H, Nakai K, Isogai T, Sugano S. :Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res.* 16:55–65 (2006).
- Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sugang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, James K, Quiles M, Madan Babu M, Saito T, Buchrieser C, Wardroper A, Felder M, Thangavelu M, Johnson D, Knights A, Loulseged H, Mungall K, Oliver K, Price C, Quail MA, Urushihara H, Hernandez J, Rabbinowitsch E, Steffen D, Sanders M, Ma J, Kohara Y, Sharp S, Simmonds M, Spiegler S, Tivey A, Sugano S, White B, Walker D, Woodward J, Winckler T, Tanaka Y, Shaulsky G, Schleicher M, Weinstock G, Rosenthal A, Cox EC, Chisholm RL, Gibbs R, Loomis WF, Platzer M, Kay RR, Williams J, Dear PH, Noegel AA, Barrell B, Kuspa A. :The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 435: 43–57, (2005).
- Kojima, K. K. and Fujiwara, H. : An extraordinary retrotransposon family encoding dual endonucleases, *Genome Res.*, 15, 1106–1117 (2005).
- Matsumoto, T. Hamada, M. Osanai, M. and Fujiwara, H. : Essential domains for Ribonucleoprotein complex formation required for retrotransposition of a telomere specific non-LTR retrotransposon SART1. *Mol. Cell. Biol.*, 26, 5168–5179 (2006).
- Maita, N., Aoyagi, H, Osanai, M., Shirakawa, M. and Fujiwara, H. Characterization of the sequence specificity of the R1Bm endonuclease domain by structural and biochemical studies. *Nucl. Acids Res.*, 35, 3918–3927 (2007)
- Futahashi, R. and Fujiwara, H: Juvenile hormone regulates butterfly larval pattern switches, *Science*, 319, 1061 (2008).
- Tribolium genome sequencing consortium (including Fujiwara, H. and Osanai, M.): The genome of the developmental model beetle and pest *Tribolium castaneum*. *Nature*, 452, 949–955 (2008).
- Suzuki, N., Su, T. H., Wu, S. W., Yamamoto, K., Khoo, K. and Lee, Y. C.: Structural analysis of N-glycans from gull egg white glycoproteins and egg yolk IgG. *Glycobiology* in press (2009).
- Kawasaki, N., Ichikawa, Y., Matsuo, I., Totani, K., Matsumoto, N., Ito, Y., Yamamoto, K.: The sugar-binding ability of ERGIC-53 is enhanced by its interaction with MCFD2. *Blood* 111, 1972–1979 (2008).
- Yamaguchi, D., Kawasaki, N., Matsuo, I., Totani, K., Tozawa, H., Matsumoto, N., Ito, Y. and Yamamoto, K.: VIPL has sugar-binding activity speccific for high-mannose-type N-glycans, and glucosylation of the  $\alpha$ 1,2 mannosyl branch blocks its binding. *Glycobiology* 17, 1061–1069 (2007).
- Nawa, D., Shimada, O., Kawasaki, N., Matsumoto, N. and Yamamoto, K.: Stable interaction of the cargo receptor VIP36 with molecular chaperone BiP. *Glycobiology* 117, 913–921 (2007).
- Kawasaki, N., Matsuo, I., Totani, K., Nawa, D., Suzuki, N., Yamaguchi, D., Matsumoto, N., Ito, Y. and Yamamoto, K.: Detection of weak sugar binding activity of VIP36 using VIP36-streptavidin complex and membrane-based sugar chains *J. Biochem.* 141, 221–229 (2007).
- Ito, M., Maruyama, T., Saito, N., Koganei, S., Yamamoto, K., and Matsumoto, N.: Killer cell lectin-like receptor G1 binds three members of the classical cadherins to inhibit NK cell cytotoxicity. *J. Exp. Med.* 203, 289–295 (2006)
- Wong, M-T., Mino, T., Seviour, R.J., Onuki, M., Liu, W-T. : In situ identification and characterization of the microbial community structure of full-scale enhanced biological phosphorous removal plants in Japan. *Water. Res.*, 39, 2901–2914 (2005).
- Michinaka, A., Arou, J., Onuki, M., Satoh, H., Mino, T.: Analysis of polyhydroxyalkanoate (PHA) synthase gene in activated sludge that produces PHA containing 3-hydroxy-2-methylvalerate. *Biotechnol. Bioeng.*, 96, 871–880 (2006).
- Otawa, K., Lee, S.H., Yamazoe, A., Onuki, M., Satoh, H., Mino, T. : Abundance, Diversity, and Dynamics of Viruses on Microorganisms in Activated Sludge Processes. *Microb. Ecol.*, 53, 143–152. (2007).
- Fukushima, T., Uda, N., Okamoto, M., Onuki, M., Satoh, H., Mino, T.: Abundance of Candidatus 'Accumulibacter phosphatis' in Enhanced Biological Phosphorus Removal Activated Sludge Acclimatized with Different Carbon Sources. *Microb. Environ.*, 22(4), 346–354(2007).
- Sueoka, K., Satoh, H., Onuki, M., Mino, T: Microorganisms involved in anaerobic phenol degradation in the treatment of synthetic coke-oven wasteWaterer detected by RNA stable-isotope probing. *FEMS Microbiol. Lett.*, 291, 169–174(2009)
- Noman, N., Iba, H. : Inferring Gene Regulatory Networks using Differential Evolution with Local Search Heuristics", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 4, pp. 634–647, (2007).
- Nikolaev, N., Iba, H. : Adaptive Learning of Polynomial Networks Genetic Programming, Backpropagation and Bayesian Methods", in Series: Genetic and Evolutionary Computation, Springer, (2006).

## ②国際会議等の開催状況【公表】

(事業実施期間中に開催した主な国際会議等の開催時期・場所、会議等の名称、参加人数(うち外国人参加者数)、主な招待講演者(3名程度))

- ・ 21世紀COEプログラム「言語から読み解くゲノムと生命システム」公開シンポジウム  
平成17年2月21日(月)、場所：東京大学弥生講堂 一条ホール  
参加人数(うち外国人参加者数)：161(5)名、主な招待講演者：招待講演者なし
- ・ e-バイオロジーイニシアティブ：新しい生物学の開拓  
平成17年3月11日(金)、場所：東京大学 武田ホール  
参加人数(うち外国人参加者数)：85(13)名  
主な招待講演者：Vivian Lee (European Bioinformatics Institute), Limsoon Wong (Institute for Infocomm Research, Singapore), Kevin B. Cohen (Center for Computational Pharmacology, University of Colorado Health Sciences Center)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成17年5月28日(土)、場所：新大阪イベントホール レルミエール  
参加人数(うち外国人参加者数)：63(0)名  
主な招待講演者：田畑哲之(かずさDNA研究所)、富田勝(慶應義塾大学先端生命科学研究所)
- ・ 「東大一産総研一理研連携によるバイオインフォマティクス教育インフラストラクチャー」  
平成18年5月8日(月) 場所：東京大学弥生講堂 一条ホール  
参加人数(うち外国人参加者数)：181(6)名  
主な招待講演者：榊 佳之(理化学研究所 ゲノム科学総合研究センター長)、  
秋山 泰(産業技術総合研究所 生命情報科学研究センター長)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成18年5月28日(日) 場所：新大阪イベントホール レルミエール  
参加人数(うち外国人参加者数)：49(0)名 主な招待講演者：中村 春木(大阪大学蛋白質研究所)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成19年4月15日(日) 場所：千里ライフサイエンスセンタービル5階 ライフホール  
参加人数(うち外国人参加者数)：29(2)名 主な招待講演者：池村 淑道(長浜バイオ大学)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成19年4月22日(日) 場所：東京大学 武田ホール  
参加人数(うち外国人参加者数)：79(1)名 主な招待講演者：宮尾 祐介(東京大学情報理工学系研究科)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成20年4月12日(土) 場所：東京大学 武田ホール  
参加人数(うち外国人参加者数)：74(0)名 主な招待講演者：長谷部 光泰(自然科学研究機構 基礎生物学研究所)
- ・ 21世紀COEプログラム 公開シンポジウム「言語から読み解くゲノムと生命システム」  
平成20年4月20日(日) 場所：千里ライフサイエンスセンタービル5階 ライフホール  
参加人数(うち外国人参加者数)：34(1)名  
主な招待講演者：林 哲也(宮崎大学フロンティア科学実験総合センター)、佐々木 顕(総合研究大学院大学)
- ・ 日本台湾若手科学者会議
  - \* 平成20年3月9日(日)～3月11日(火) 場所： 国立清華大学(台湾)  
参加人数(うち外国人参加者数)：47(30)名  
主な招待講演者：Wen-Hsiung Li University of Chicago, USA、Russell Cox (University of Bristol, UK)、Limsoon Wong (National University of Singapore)
  - \* 平成20年11月4日(火)～11月6日(木) 場所：産業技術総合研究所(お台場)  
主な招待講演者：Kenneth K. Kidd (Yale University School of Medicine, U.S.A.)、Aviv Regev (Massachusetts Institute of Technology)、Luonan Chen (Shanghai University)  
参加人数(うち外国人参加者数)：57(29)名

## 2. 教育活動実績【公表】

- 博士課程等若手研究者の人材育成プログラムなど特色ある教育取組等についての、各取組の対象（選抜するものであればその方法を含む）、実施時期、具体的内容
- ・ COE特別講義（平成16年度～平成20年度）  
 修士課程学生、博士課程学生、RA（リサーチアシスタント）等を対象として、バイオインフォマティクスやゲノム研究の第一線で活躍する研究者によるCOE特別講義を実施した。また、COE特別講義において距離的に離れたキャンパス（本郷キャンパス・白金台キャンパス・産業技術総合研究所）にいる大学院生向けに遠隔講義システムを整備し、遠隔講義をおこなった。  
 平成16年度12コマ、平成17年度24コマ、平成18年度24コマ、平成19年度24コマ、平成20年度12コマ
  - ・ 外国人研究者によるセミナーの開催（平成16年度～平成20年度）  
 バイオインフォマティクスの分野で世界的に活躍する研究者の講演会を白金キャンパス、柏キャンパスにて行った（計7回）
  - ・ RA・研究員研究成果発表会（平成16年度～平成20年度）  
 各年度の3月に研究成果発表会（ポスター発表）を柏キャンパスで行った。RA及びCOE研究員の各年度の活動はそれぞれ成果報告書にまとめた。
  - ・ RA雇用と研究プロジェクトへの参加（平成16年度～平成20年度）  
 平成16年度 博士課程66名、修士課程14名/平成17年度 博士課程69名/平成18年度 博士課程46名  
 平成19年度 博士課程42名/平成20年度 博士課程39名
  - ・ 学生の海外派遣 平成16年度 5名/平成17年度 13名
  - ・ COEウェブページによる情報交換と異分野交流（平成16年度～平成20年度）  
 学生、RA、若手研究員同士がディスカッションできるブログを作成し、情報交換や異分野交流を促進した。
  - ・ 情報系と実験系の共同研究への積極的参画を推進（平成16年度～平成20年度）  
 実験系学生が情報系研究室においてバイオインフォマティクス技術を習得し実験研究に活用する事例、情報系学生が実験系研究室においてバイオインフォマティクス技術の有効性を実験的に検証する事例など、具体的な共同研究が活発に行われた。
  - ・ ジュニアアワード  
 COEプログラムに関連する分野の研究について、優秀な能力を持ち、研究者としての将来が期待できる修士課程（前期課程）大学院生を表彰した。平成19年度 1名 平成20年度 1名
  - ・ RA 異分野交流合宿（平成17年8月31日～9月1日）  
 本COEで雇用している博士課程学生RA（45名）、特任研究員（1名）および特任助手（5名）が参加し、情報科学と実験科学の融合を目的とした学生主催の合宿を実施した。かずさDNA研究所の見学と小グループ討論会を通じ、ゲノム科学における情報系と実験系の関係のあり方、具体的な相互交流の可能性などについて議論した。この合宿の成果をまとめた冊子を発行した。
  - ・ Chemical Biology Programming Workshop(平成19年3月20日～3月22日)  
 柏キャンパスに隣接するさわやか千葉県民プラザにおいて、プログラミング演習を中心としたワークショップを開催した。対象は新領域創成科学研究科に限らない全国の大学院生に旅費を支給する形で参加してもらい、合宿形式でケミカルバイオロジーの集中講義およびプログラミング演習を行った。（URL <http://www.metabolome.jp/doc/workshop/past/chembio2007/>）  
 参加希望者は受講の抱負を書いてA4用紙に提出し、その中から十数人を選抜。
  - ・ インターンシップ制度の実施（平成17年度）  
 文部科学省「派遣型高度人材育成協同プラン」制度を用いて、修士1年の学生数名が日立ソフトウェアエンジニアリング、かずさDNA研究所、産業技術総合研究所等の関連機関に短期滞在し、バイオインフォマティクスの幅広い知識を学んだ。
  - ・ 日本台湾若手科学者会議（平成19年度、平成20年度）  
 計2回開催されたこの国際会議は、新領域創成科学研究科を含む日本と台湾の博士課程在籍中および博士課程終了直後の若手研究者により、プログラム策定、招待講演者のアレンジ、演題の募集、査読等がおこなわれ、若手研究者にとって会議開催のOn the Job Trainingとなる試みである。学生オーガナイザは希望者が役割分担する構成。
  - ・ 生命科学実習（平成19年度）  
 wetの実験をしたことのない情報系の学生を対象にして、2週間の生命科学実習を企画しタンパク質、DNA等の実習を夏休み（8月上旬）に行った。

21世紀COEプログラム委員会における事後評価結果

(総括評価)

設定された目的は十分達成された

(コメント)

生命科学における重要課題を据え、異分野を統括する総合的な体制作りを目指し、大学からも多くの支援を引き出し、研究・教育両面において、組織作りや設備の整備が進み、また、中間評価の指摘に対しても、適切な対応が行われており、その結果、当初の目標を上回る成果をあげていると評価できる。

人材育成面については、新分野に相応しい教育プログラムの設定など、特色ある成果をあげ、人材育成に努力し、多くの人材を輩出しており、評価できる。

研究活動面については、一流の研究成果を多く生み出しており、また、今後の展開に関しても大いに期待できることから、さらに理念を追い求め、独自性の高い拠点の形成に努めることが望まれる。しかしながら、「言語」という視点を取り入れて他に類を見ない特徴ある新分野が形成されるという期待には十分には応えておらず、今後、バイオインフォマティクスの分野で世界最高水準の拠点形成を目指すことが望まれる。